Organon F 24 (4) 2017: 456-486

Learning Conditional and Causal Information by Jeffrey Imaging on Stalnaker Conditionals

MARIO GÜNTHER[1]

ABSTRACT: We show that the learning of (uncertain) conditional and/or causal information may be modelled by (Jeffrey) imaging on Stalnaker conditionals. We adapt the method of learning uncertain conditional information proposed in Günther (2017) to a method of learning uncertain causal information. The idea behind the adaptation parallels Lewis (1973c)'s analysis of causal dependence. The combination of the methods provides a unified account of learning conditional and causal information that manages to clearly distinguish between conditional, causal and conjunctive information. Moreover, our framework seems to be the first general solution that generates the correct predictions for Douven (2012)'s benchmark examples and the Judy Benjamin Problem.

KEYWORDS: Causal dependence – Douven's examples – Imaging – Judy Benjamin Problem – Learning – Stalnaker conditional.

---

[1] Received: 21 January 2017 / Accepted: 28 June 2017

✉ Mario Günther
Ludwig-Maximilian-University Munich
Munich Center for Mathematical Philosophy
& Graduate School for Systemic Neurosciences
Chair of Logic and Philosophy of Language
Geschwister-Scholl-Platz 1, D-80539 Munich, Germany
e-mail: Mario.Guenther@campus.lmu.de

1. Introduction

"How do we learn conditional information?" Igor Douven et al. present this question for consideration in a series of papers (cf. Douven & Dietz 2011; Douven & Romeijn 2011; Douven 2012; Pfeifer & Douven 2014, especially section 6). Douven (2012) contains a survey of the available accounts that model the learning of conditional information. The survey comes to the conclusion that a general account of probabilistic belief updating by learning (uncertain) conditional and causal information is still to be formulated. Pfeifer & Douven (2014) analyses the state of the art even more pessimistically by writing that "no one seems to have an idea of what an even moderately general rule of updating on conditionals might look like," even if we restrict the scope of the account to indicative conditionals (Pfeifer & Douven 2014, 213). We aim to provide such a general account of updating that unifies the learning of (uncertain) conditional and causal information.

In Günther (2017), we proposed a method of learning conditional information. We have shown that the predictions of the proposed method align with the intuitions in Douven (2012)'s benchmark examples and can generate predictions for the Judy Benjamin Problem. In this paper, we adapt the method of learning conditional information to a method of learning causal information. The adapted method allows us to causally conceive of the information conveyed by the conditionals uttered in Douven's examples and the Judy Benjamin Problem.

It may come as a surprise that we propose an account of learning that involves (Jeffrey) imaging. After all, the standard view on learning that α is Bayesian updating on α, while David Lewis's imaging on α is widely conceived of as modeling the supposition of α. But learning a conditional may – according to the suppositional view on conditionals – be interpreted as learning what is true under a supposition (about which we may be uncertain). In particular, learning the conditional "If α, then γ" is thus equivalent to learning the conditional information that γ is the case under the supposition that α is the case.

Douven aims to provide an account of learning conditional information that is an empirically adequate account of human reasoning. Douven & Verbrugge (2010) submitted the thesis whether the acceptability of an indicative conditional 'goes by' the conditional probability of its consequent given the antecedent to empirical testing, and claim that the experiments speak against the thesis.[2] Their results indicate that conditional probabilities do not correspond to probabilities of conditionals, which was proved by Lewis (1976), if conditionals are understood as Stalnaker conditionals. Those formal and empirical results obviously provide a severe challenge for Bayesian analyses of learning conditionals, where conditional probabilities usually take center stage.

Moreover, Zhao et al. (2012) obtained empirical results that indicate a fundamental difference between supposing and learning. In particular, supposing a conditional's antecedent α seems to have less impact on the credibility of the consequent γ than learning that α is true. We will provide a framework that allows us both, to distinguish between the learning of 'factual' and conditional information and to generate empirically testable predictions.

In Section 2, we introduce Douven's desideratum for accounts of learning (uncertain) conditional information. His own proposal is based on the explanatory status of the antecedent. In Section 2.1, we sketch his argumentation against the method of imaging on the Stalnaker conditional as an account of learning conditional information. The reason for Douven's dismissal of the method is that the rationality constraints of Stalnaker models are not sufficient to single out a model, which may count as a representation of a belief state.

In Section 3, we review the method of learning (uncertain) conditional information proposed in Günther (2017), where we showed that Douven's dismissal is unjustified. We met Douven's challenge for possible worlds models by imposing two additional constraints: interpreting the meaning of a Stalnaker conditional in a minimally informative way and supplementing the analysis by a default assumption. Moreover, we generalised Lewis's imaging method in order to account for uncertain information as well.

In Section 4, we adapt the method of learning conditional information to a method of learning causal information. The adaptation is inspired by Lewis's notion of causal dependence and replaces the default assumption by the assumption that the antecedent makes a difference. In Section 4.1, we apply our adapted method of learning causal information to Douven's

---

2 The 'goes by' is Lewis's formulation that may be found in Lewis (1976, 297).

examples and the Judy Benjamin Problem. In Section 5, we formally implement Douven's idea concerning the explanatory status of the antecedent within our framework.

2. Douven's account of learning conditional information via the explanatory status of the antecedent

Igor Douven propounds a broadly Bayesian model of learning conditional information. As the standard Bayesian view of learning, Douven's account assumes that learning the unnested indicative conditional "If α, then γ" implies that the posterior degree of belief for γ given α is set to approximately 1, i. e. $P^*(\gamma \mid \alpha) \approx 1$. In contrast to standard Bayesian epistemology, explanatory considerations play a major role in his model of updating on conditionals.

Douven proposes a desideratum for any account of learning conditional information, viz. a criterion that determines whether an agent raises, lowers, or leaves unchanged her degree of belief $P(\alpha)$ for the antecedent upon learning a conditional.

He even writes that we "should be [...] dissatisfied with an account of updating on conditionals that failed to explain [...] basic and compelling intuitions about such updating, such as, in our examples" (Douven 2012, 3). Douven's methodology consists in searching for an updating model that accounts for our intuitions with respect to three examples, the Sundowners Example, the Ski Trip Example and the Driving Test Example. The three examples represent the classes of scenarios, in which $P(\alpha)$ should intuitively remain unchanged, be increased and decreased, respectively. He dismisses any method of learning conditional information that cannot account for all of the three examples. He emphasises that no single account of learning uncertain conditional and/or causal information is capable of solving all of his examples. Taking the examples as benchmark, he also dismisses the Stalnaker conditional as a tool to model the learning of conditional information.

The core hypothesis of Douven's account is that the change in explanatory quality or 'explanatory status' of the antecedent α during learning the information results in a change of the degree of belief for α. If the explanatory status of α goes up, that is α explains γ well, then the degree of belief after learning the conditional increases, i. e. $P^*(\alpha) > P(\alpha)$; if the explanatory status of α goes down, $P^*(\alpha) < P(\alpha)$; if the explanatory status remains the same, a variant of Jeffrey conditioning is applied that has the property that $P^*(\alpha) = P(\alpha)$. Following Richard Bradley, Douven calls this Jeffrey conditioning over a restricted partition 'Adams conditioning on $P^*(\gamma \mid \alpha) \approx 1$'.[3]

Douven and Romeijn proposed a solution to the Judy Benjamin Problem. The problem indicates that the revision method that minimises the Kullback-Leibler divergence leads to counterintuitive results for learning uncertain conditional information. Their solution uses the variant of Jeffrey conditioning mentioned above. However, their proposed method fails to account for examples where the probability of the antecedent is supposed to change, since it has the invariance property that $P^*(\alpha) = P(\alpha)$, for all α, and thus disqualifies as a general account of learning conditional information (cf. Douven & Romeijn 2011, 648-655; Douven 2012, 9-11).

2.1. Douven's dismissal of the Stalnaker conditional

Douven claims that Stalnaker conditionals are not suited to model the learning of conditional information. He argues for this claim by pointing out that a learning method based on the Stalnaker conditional "makes no predictions at all about any of our examples" (Douven 2012, 7). The cited reason is that we would not be able to exclude certain Stalnaker models as rational representation of a belief state.

Douven provides three possible worlds models for his point. Each model consists of four worlds such that all logical possibilities of two binary variables are covered. He observes that imaging on "If α, then β" interpreted as a Stalnaker conditional has different effects: in model I, the probability of the antecedent α, i. e. $P(\alpha)$ decreases; in model II, $P(\alpha)$ remains unchanged; and in model III $P(\alpha)$ increases. According to Douven this flexibility of the class of possible world models is a problem rather than an advantage, since there would be no rationality constraints to rule out certain models as rational representations of a belief state.

---

3 The partition is restricted according to the odds for the consequent of the learned conditional. For details, see Bradley (2005, 351-352); and Douven & Romeijn (2011, 650-653).

Consider a scenario of the class, where the antecedent remains unchanged (e.g. the Sundowners Example). The problem is, so Douven argues, that there are no criteria to exclude models I and III as rational representations of a belief state, in which $P(\alpha)$ should not change. In Douven's words:

> In fact, to the best of my knowledge, nothing said by Stalnaker (or Lewis, or anyone else working on possible worlds semantics) implies that, supposing imaging is the update rule to go with Stalnaker's account, models I and III [...] could not represent the belief state of a rational person; [...] In short, interpreting "If A, B" as the Stalnaker conditional and updating on it [...] by means of imaging offers no guarantee that our intuitions are respected about what should happen – or rather not happen – after the update [...]. Naturally, it cannot be excluded that some of these models – and perhaps indeed all on which [...] [the degree of belief in the antecedent] changes as an effect of learning [the conditional] – are to be ruled out on the basis of rationality constraints that I am presently overlooking, perhaps ones still to be uncovered, or at least still to be related to possible worlds semantics as a tool for modelling epistemic states. It is left as a challenge to those attracted to the view considered here to point out such additional constraints. (Douven 2012, 8-9)

In Günther (2017), we met the challenge Douven mentions in the quote. We discovered two constraints that singled out Stalnaker models that plausibly represent the belief states in Douven's benchmark examples. Imposing the two additional constraints amounts to interpreting the meaning of a Stalnaker conditional in a minimally informative way and supplementing the analysis by a default assumption.

3. Review of the Method of Learning Conditional Information by Jeffrey Imaging on Stalnaker Conditionals

Günther (2017) puts forward a method of learning conditional information by Jeffrey imaging on Stalnaker conditionals. The learning method may be summarised as follows. (i) We model an agent's belief state as a Stalnaker model. (ii) The agent learns conditional information by (ii).(a) interpreting the received conditional information as a Stalnaker conditional; (ii).(b) constraining the similarity order by the meaning of the Stalnaker conditional in a minimally informative way and respecting the default assumption; and (ii).(c) updating her degrees of belief by Jeffrey imaging on this Stalnaker conditional (together with further contextual information, if available).

We outline the method of learning conditional information by presenting its constituents, i. e. the semantics of the Stalnaker conditional, Jeffrey imaging and the meaning of 'minimally informative'. Afterwards, we put the constituents together.

3.1. The Stalnaker conditional

The idea behind a Stalnaker conditional may be expressed as follows: a Stalnaker conditional $\alpha > \gamma$ is true at a world w iff γ is true in the most similar possible world w' to w, in which α is true (cf. Stalnaker 1975).[4] We denote the set of possible worlds that satisfies a formula α by $[\alpha]$. Thereby, we identify the set $[\alpha]$ with the proposition expressed by α. In symbols, $[\alpha] = \{w \in W \mid w(\alpha) = 1\}$, where each w of the set of worlds under consideration W may be thought of as a Boolean evaluation.

A Stalnaker conditional is evaluated with respect to a Stalnaker model, i. e. a model of possible worlds where each world w is equipped with a total order such that w is the unique center of the respective order and, for non-contradictions α, it is guaranteed that there exists a unique most similar world $min_{\leq_w} [\alpha]$ from w that satisfies α. The accessibility relation of a Stalnaker model is reflexive and connective.

Let us state more precisely the meaning of a Stalnaker conditional using the notations just introduced. "If α, then γ" denotes according to Stalnaker's proposal the set of worlds (or equivalently the proposition) containing each world whose most similar α-world is a world that satisfies γ. In

4 Note that Stalnaker's theory of conditionals aims to account for both indicative and counterfactual conditionals. We set the complicated issue of this distinction aside in this paper. However, we want to emphasise that Douven's examples and the Judy Benjamin Problem only involve indicative conditionals.

symbols, $[\alpha > \gamma] = \{w \mid w \models \alpha > \gamma\} = \{w \mid \min_{\leq w} [\alpha] = \emptyset$ or $\min_{\leq w} [\alpha] \models \gamma\}$.[5]

Finally, note that any Stalnaker model validates the principle called 'Conditional Excluded Middle' according to which $(\alpha > \gamma) \vee (\alpha > \neg\gamma)$. The reason for the validity of Conditional Excluded Middle is that, for any $w \in W$, the single most similar α-world $\min_{\leq w} [\alpha]$ is either a γ-world, or else a $\neg\gamma$-world. This principle will come in handy when modeling the learning of uncertain information.

In the next section, we introduce Lewis's imaging method, which we will generalise in the subsequent section.

3.2. Lewis's imagining

David Lewis developed a probabilistic updating method called 'imaging' (cf. Lewis 1976). We introduce a notational shortcut: for each world w and each (possible) antecedens α, $w_\alpha = \min_{\leq w} [\alpha]$ be the most similar world of w such that $w_\alpha(\alpha) = 1$. Invoking the shortcut, we can then specify the truth conditions for Stalnaker's conditional operator > as follows:

(1) $w(\alpha > \gamma) = w_\alpha(\gamma)$, if α is possible.[6]

Definition 1. Probability Space over Possible Worlds

We call $\langle W, \wp(W), P\rangle$ a probability space over a finite set of possible worlds W iff

(i) $\wp(W)$ is the set of all subsets of W,

(ii) and $P : \wp(W) \mapsto [0, 1]$ is a probability measure, i.e.

(a) $P(W) = 1$, $P(\emptyset) = 0$, and

(b) for all $X, Y \subseteq W$ such that $X \cap Y = \emptyset$, $P(X \cup Y) = P(X) + P(Y)$.

As before, we conceive of the elements of $\wp(W)$ as propositions. We define, for each α, $P(\alpha) = P([\alpha])$. We see that W corresponds to an arbitrary tautology denoted by $\top$ and $\emptyset$ to an arbitrary contradiction denoted by $\bot$.

---

5 See Günther (2017) for a more thorough presentation of Stalnaker models. See Stalnaker & Thomason (1970) for Stalnaker and Thomason's original presentation of the Stalnaker semantics.

6 We assume here that there are only finitely many worlds. Note also that if α is -

Definition 1 allows us to understand a probability measure P as a probability distribution over worlds such that each w is assigned a probability $P(w) > 0$, and $\sum_w P(w) = 1$. We may determine the probability of a formula α by summing up the probabilities of the worlds at which the formula is true.[7]

$$(2) \qquad P(\alpha) = \sum_w P(w) \cdot w(\alpha)$$

Now, we are in a position to define Lewis's updating method of imaging.

Definition 2. Imaging (Lewis 1976, 310)

For each probability function P, and each possible formula α, there is a probability function P^α such that, for each world w', we have:

$$(3) \qquad P^\alpha(w') = \sum_w P(w) \cdot \begin{Bmatrix} 1 \text{ if } w_\alpha = w' \\ 0 \text{ otherwise} \end{Bmatrix}$$

We say that we obtain P^α by imaging P on α, and call P^α the image of P on α.

Intuitively, imaging transfers the probability of each world w to the most similar α-world w_α. Importantly, the probabilities are transferred, but in total no probability mass is additionally produced and no probability mass is lost. In formal terms, we have always $\sum_{w'} P^\alpha(w') = 1$. Any α-world w' keeps at least its original probability mass (since then $w_\alpha = w'$), and is possibly transferred additional probability shares of $\neg\alpha$-worlds w iff $\min_{\leq_w} [\alpha] = w'$. In other words, each α-world w' receives as its updated probability mass its previous probability mass plus the previous probability shares that were assigned to $\neg\alpha$-worlds w such that $\min_{\leq_w} [\alpha] = w'$. In this way, the method of imaging distributes the whole probability onto the α-worlds such that $P^\alpha(\alpha) = \sum_{w(\alpha)=1} P(w(\alpha)) = 1$, and each share remains 'as close as possible' at the world at which it has previously been located. For an illustration, see Figure 1.

---

7 We assume here that each world is distinguishable from any other world, i. e. for two arbitrary worlds, there is always a formula such that the formula is true in one of the worlds, but false in the other. In other words, we consider no copies of worlds.

Lewis proved the following theorem, which relates the semantics of the Stalnaker conditional and the method of imaging on its antecedent.

Theorem 1. (Lewis 1976, 311)
The probability of a Stalnaker conditional equals the probability of the consequent after imaging on the antecedent, i. e. $P(\alpha > \gamma) = P^{\alpha}(\gamma)$, if α is possible.

Note that α in Theorem 1 may itself be of conditional form $\beta > \delta$ for any formulas β, δ.

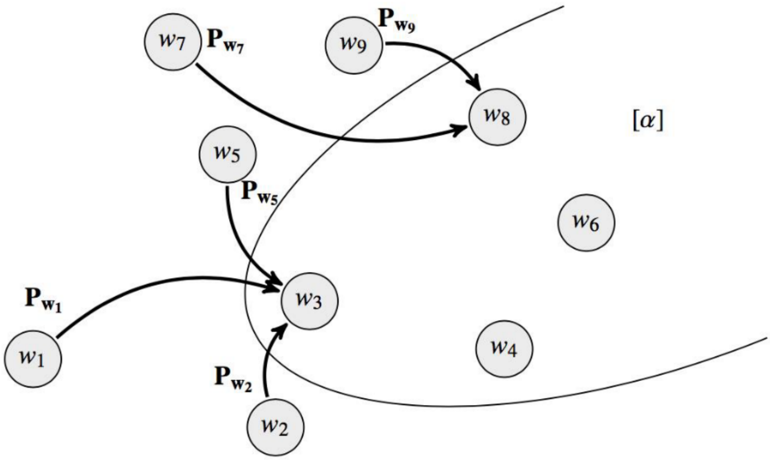

Figure 1: A set of possible worlds. The area delineated by the elliptical line represents the proposition or set of worlds $[\alpha] = \{w_3, w_4, w_6, w_8\}$. The thick arrows represent the transfer of probability shares from the respective $[\neg\alpha]$-worlds to their most similar $[\alpha]$-world. Similarity is graphically represented by topological distance between the worlds such that w_3, for instance, is the most similar or 'closest' $[\alpha]$-world to w_2.

3.3. Jeffrey imagining

The case of learning uncertain conditional information, i. e. $P(\alpha > \gamma) = k$ for $k \in [0, 1]$ but unequal to 0 or 1, requires to generalise Lewis's imaging method of Definition 2. In analogy to Jeffrey conditionalisation, we call the generalised method 'Jeffrey' imaging. Jeffrey imaging is based on Lewis's imaging and the fact that in a Stalnaker model the principle of Conditional Excluded Middle prescribes that $\neg(\alpha > \gamma)$ is equivalent to $\alpha > \neg\gamma$. We know, for all $w \in W$, presupposed $\alpha > \gamma$ is possible, both (I) that $\sum_w P^{\alpha > \gamma}(w)$ sums up to 1 and (II) that $\sum_w P^{\alpha > \neg\gamma}(w)$ sums up to 1. The idea is that if we form a weighted sum over the terms of (I) and (II) with some parameter $k \in [0, 1]$, then we obtain again a sum of terms $P_k^{\alpha > \gamma}(w)$ such that $\sum_w P_k^{\alpha > \gamma}(w) = 1$. Note, however, that we present the more general case $P_k^{\alpha}(w)$ in the definition below.

Definition 3. Jeffrey Imaging

For each probability function P, each possible formula α (possibly of conditional form $\beta > \delta$), and some parameter $k \in [0, 1]$, there is a probability function P_k^{α} such that for each world w' and the two similarity orderings centred on w_α and $w_{\neg\alpha}$, we have:

$$(4) \qquad P_k^{\alpha}(w') = \sum_w \left(P(w) \cdot \begin{Bmatrix} k & \text{if } w_\alpha = w' \\ 0 & \text{otherwise} \end{Bmatrix} + P(w) \cdot \begin{Bmatrix} 1-k & \text{if } w_{\neg\alpha} = w' \\ 0 & \text{otherwise} \end{Bmatrix} \right)$$

We say that we obtain P_k^{α} by Jeffrey imaging P on α, and call P_k^{α} the Jeffrey image of P on α. Note that in the case where $k = 1$, Jeffrey imaging reduces to Lewis's imaging.

Theorem 2. Properties of Jeffrey Imaging

(i) $\sum_{w'} P_k^{\alpha}(w') = 1$

(ii) $P_k^{\alpha}(\alpha) = k$

(iii) $P_k^{\alpha}(\neg\alpha) = (1-k)$

(iv) $P_k^{\alpha}(\gamma) = k \cdot P(\alpha > \gamma)$ [8]

---

[8] The proofs of the properties can be found in Günther (2017).

We see that in total the revision method of Jeffrey imaging does neither produce additional probability shares, nor destroy any probability shares. In contrast to Lewis's imaging, Jeffrey imaging does not distribute the whole probabilistic mass onto the α-worlds, but only a part thereof that is determined by the parameter k.

In particular, as compared to Lewis's imaging, Jeffrey imaging may be understood as implementing a more moderate or balanced movement of probabilistic mass between α- and $\neg\alpha$-worlds. For an illustration, see Figure 2.

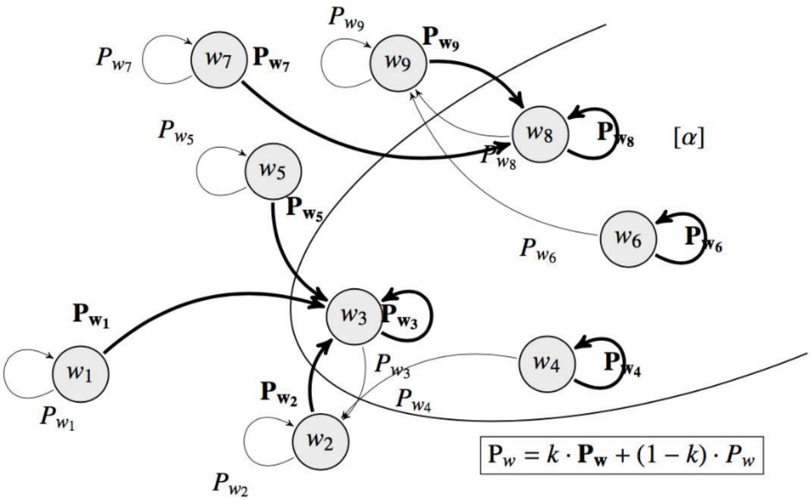

Figure 2: An illustration of the probability kinematics of Jeffrey imaging. The Jeffrey image P_k^α is characterised by a 'k-inertia' of the probabilistic mass from the respective α-worlds, and a '$(1-k)$-inertia' of the probabilistic mass from the respective $\neg\alpha$-worlds. Each thick arrow represents the transfer of the probability share $k\cdot P(w)$ to the closest α-world from w. Each thin arrow represents the transfer of the probability share $(1-k)\cdot P(w)$ to the closest $\neg\alpha$-world from w.

It is easy to show that P_k^α is a probability function. In a possible worlds framework, such a proof basically amounts to showing that the probability shares of all the worlds sum up to 1 after Jeffrey imaging. Therefore, property (i) of Theorem 2 provides minimal justification for applying Jeffrey imaging to probabilistic belief updating.

3.4. Putting the constituents together

Now we outline the method of learning conditional information put forward in Günther (2017). The method comprises three main steps:

(i) We model an agent's belief state as a Stalnaker model such that all and only those logical possibilities are represented as single worlds, which are relevant to the scenario under consideration. For instance, if only a single conditional "If α, then γ" is relevant and nothing else, then W contains exactly four elements as depicted in Figure 3.[9]

(ii) An agent learns conditional information "If α, then γ" iff (a) the agent interprets the received conditional information as a Stalnaker conditional $\alpha > \gamma$; (b) changes the similarity order $\leq$ by the meaning of $\alpha > \gamma$ in a minimally informative way and respecting the default assumption; and (c) updates her degrees of belief by Jeffrey imaging on the minimally informative meaning of $\alpha > \gamma$.

(iii) Finally, we check whether or not the result of Jeffrey imaging obtained in step (ii).(c) corresponds to the intuition associated with the respective example.

Step (ii) constitutes the core of the learning method and requires further clarification:

(a) In the agent's belief state, i.e. a Stalnaker model, the received information is interpreted. In the case of conditional information, the received information is interpreted as Stalnaker conditional. Hence, if the agent receives the information "If α, then γ", she interprets the information as meaning that the most similar

[9] In other words, we consider "small" models of possible worlds and do not allow for copies of worlds, i. e. worlds that satisfy the same formulas.

α-world (from the respective actual world) is a world that satisfies γ (presupposed α is possible). Technically, the interpretation (i.e. the meaning) of $\alpha > \gamma$ (relative to the Stalnaker model) is the proposition $[\alpha > \gamma] = \{w \in W \mid \min_{\leq w} [\alpha] \in [\gamma]\}$, where w is the respective actual world.

(b) The similarity order(s) is/are changed upon receiving conditional information. The proposition $\{w \in W \mid \min_{\leq w} [\alpha] \in [\gamma]\}$ depends on the similarity order $\leq$. The learning method prescribes that $\leq$ is specified, or adjusted, such that from each world the most similar α-world is a γ-world whenever possible. In other words, the method demands a maximally conservative, or equivalently minimally informative, interpretation of the received information. This amounts to specifying or adjusting the orders $\leq_w$ such that as many worlds as possible satisfy the received information. On the one hand, we can describe this interpretation as maximally conservative in the sense that no worlds are gratuitously excluded. On the other hand, we may think of possible worlds as information states. Then the exclusion of possible worlds corresponds to a gain of information. If an agent interprets the received information in a maximally conservative way, then as few as possible worlds or information states are excluded. In this sense, her gain of information is minimal.

The learning method assumes that the agent changes her similarity order respecting a default assumption. This default assumption states that the most similar $\alpha > \gamma$-world from any excluded $\alpha > \neg\gamma$-world is an $\alpha \land \gamma$-world, if there is more than one candidate. Formally, the default assumption expresses that $\min_{\leq w(\alpha>\neg\gamma)=1} [\alpha > \gamma] \models \alpha \land \gamma$, if $\min_{\leq w(\alpha>\neg\gamma)=1} [\alpha > \gamma]$ is underdetermined.[10] A justification for the default assumption is provided in Günther (2017).

---

[10] Relying on the default assumption solves a well-known problem of underdetermination: it might well be that, for instance, in the Stalnaker model depicted in Figure 3 w_3 or w_4 is the more similar $\alpha > \gamma$-world to w_2 than w_1 is. However, we will see in the examples below that additional (contextual) information may sometimes fully determine the epistemic states under consideration such that we do not always need to rely on the default assumption.

(c) Jeffrey imaging is applied on the minimally informative meaning of the Stalnaker conditional $\alpha > \gamma$. The application of Jeffrey imaging determines a probability distribution after learning the (uncertain) conditional information.

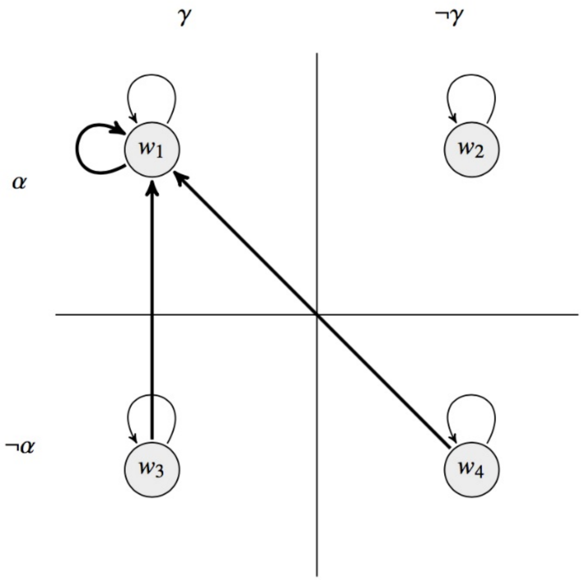

Figure 3: A four-worlds Stalnaker model for a case, in which the only received and relevant information is "If α, then γ". The reflexive thin arrows illustrate that each world w is the most similar to itself under the respective similarity order $\leq_w$. The thick arrows illustrate the change of the similarity order such that the received and interpreted information $[\alpha > \gamma]$ is minimally informative. Here, the minimally informative meaning of $\alpha > \gamma$ is $[\alpha > \gamma] = \{w \in W \mid w \models \alpha > \gamma\} = \{w_1, w_3, w_4\}$. Note that world w_2 is its own most similar α-world, but does not satisfy γ, i.e. $\min_{\leq w_2} [\alpha] \not\models \gamma$ and thus $\min_{\leq w_2} [\alpha > \gamma] \neq w_2$. Relying on the default assumption of step (ii).(b), $\min_{\leq w_2} [\alpha > \gamma] = w(\alpha \wedge \gamma) = w_1$. In words, the method prescribes that w_1 is the most similar $\alpha > \gamma$-world to w_2. This illustrates that the minimally informative meaning of $[\alpha > \gamma]$ implies that $\neg\gamma$ is excluded under the supposition of α. Hence, imaging on the minimally informative meaning of $\alpha > \gamma$ 'probabilistically excludes' w_2 and the probability share of w_2 will be fully transferred to w_1.

The proposed learning method has the following property that allows us to distinguish conditional and conjunctive information. If there is no further contextual information available to the agent receiving information, then learning the conditional information $\alpha > \gamma$ is less informative than learning the information $\alpha \land \gamma$. For, the proposition $[\alpha \land \gamma]$ is in the proposed framework always a strict subset of the minimally informative proposition $[\alpha > \gamma]$.

4. An adaptation of the method to the learning of causal information

In Section 2, we have seen that Douven invokes explanatory considerations in order to model the learning of conditional information. His account presupposes an explanatory reading of the learned conditional information, which may be of the form “If α, then γ”. While we are skeptical about the presupposition that any conditional can or should be read as (a part of) an explanation or causal dependence, we admit that conditionals often figure in explanations. Hence, the method of learning conditional information proposed in Günther (2017) should be able to account for the learning of causal information conveyed by conditionals; otherwise, the proposed method suffers a major drawback.

In this section, we sketch how the proposed method may be adapted to a method of learning causal information. The adaptation is inspired by Lewis’s analysis of causal dependence in terms of counterfactuals. Douven claims that, in any account of explanation that relies on a Stalnaker model, “to explain” means to “provide causal information”, where “causal” refers to a Lewis-style analysis.[11]

[11] Cf. Douven (2012, 8-9, especially footnote 7); and Lewis (1973c). Furthermore, Douven claims that Lewis’s and Stalnaker’s semantics for conditionals are “exactly the same” (Douven 2012, 8). However, there is a difference between Stalnaker’s and Lewis’s semantics. In a Stalnaker model, there is always a single most similar world (or no world) to the actual world, whereas Lewis’s semantics allows for a set of worlds (or no world) whose elements are equally similar to the actual world. A consequence of the difference is that Lewis’s ‘official’ semantics for conditionals, i.e. the system VC, does not validate the principle of Conditional Excluded Middle, whereas Stalnaker’s logic C2 for conditionals does. In Lewis’s nomenclature, system C2 is labelled by VCS. Cf.

We write $\alpha \Rightarrow \gamma$ for the causal reading of "If α, then γ". According to Lewis's idea of causal dependence, $\alpha \Rightarrow \gamma$ is satisfied iff $\alpha > \gamma$ and $\neg\alpha > \neg\gamma$. We may apply the proposed method of learning conditional information by taking the minimally informative meaning of $\alpha \Rightarrow \gamma$ into account (instead of the one of $\alpha > \gamma$), if we substitute the default assumption. We call the adaptation the 'method of learning causal information'.

The substitution of the default assumption to what we call 'causal difference assumption' runs as follows. Assume we have no further contextual knowledge. Then, the most similar $\alpha \Rightarrow \gamma$-world from any excluded $\alpha \Rightarrow \neg\gamma$-world is a $(\alpha \wedge \gamma)$-world, if the excluded $\alpha \Rightarrow \neg\gamma$-world satisfies α. Furthermore, the most similar $\alpha \Rightarrow \gamma$-world from any excluded $\alpha \Rightarrow \neg\gamma$-world is a $(\neg\alpha \wedge \neg\gamma)$-world, if the excluded $\alpha \Rightarrow \neg\gamma$-world satisfies $\neg\alpha$. In symbols,

(5) $$min_{\leq w_{\alpha \Rightarrow \neg\gamma}}[\alpha \Rightarrow \gamma] = \begin{Bmatrix} w_{\alpha \wedge \gamma} & if\ w_{\alpha \Rightarrow \neg\gamma}(\alpha) = 1 \\ w_{\neg\alpha \wedge \neg\gamma} & if\ w_{\alpha \Rightarrow \neg\gamma}(\alpha) = 0 \end{Bmatrix}$$

The causal difference assumption is justified, if we understand causal dependence as difference making à la Lewis (cf. Lewis 1973c). The antecedent α makes the difference as to whether γ or $\neg\gamma$. Hence, $\alpha \Rightarrow \gamma$ means that worlds in which α obtains are worlds in which γ obtains, and accordingly that worlds in which $\neg\alpha$ obtains are worlds in which $\neg\gamma$ obtains. It is built in the analysis of causal dependence, so to speak, that the difference making factors (α and $\neg\alpha$) are more dissimilar than the ensuing effects.

Note that causal dependence is more informative than conditional dependence. For, the minimally informative meaning of $[\alpha \Rightarrow \gamma]$ is always a strict subset of the minimally informative meaning of $[\alpha > \gamma]$. The reason is that causal dependence, by definition, conveys in addition to the indicative conditional information also the information $[\neg\alpha > \neg\gamma]$. In brief, if an agent learns $\alpha \Rightarrow \gamma$, our adapted method prescribes that the $\alpha \wedge \neg\gamma$-worlds

---

Lewis (1973b; 1973a); and, for details, Unterhuber (2013, especially chap. 3.2, 3.3.3 and 3.3.4). The non-identity of Lewis's and Stalnaker's semantics implies that the notion of causal dependence employed in our method of learning causal information is not equivalent to Lewis's notion of causal dependence. While the method relies on Lewis's idea, we stick to Stalnaker's semantics in this paper.

transfer their probability shares to the most similar $\alpha \wedge \gamma$-world, and the $\neg\alpha \wedge \gamma$-worlds transfer their probability shares to the most similar $\neg\alpha \wedge \neg\gamma$-world. In other words, if the antecedent α is a difference maker, then the probability mass of those worlds w that do not satisfy $\alpha \Rightarrow \gamma$ is shifted to the most similar $\alpha \Rightarrow \gamma$-world w' that agrees with the Boolean evaluation for α, i. e. $w(\alpha) = w'(\alpha)$.

4.1. Douven's examples and the Judy Benjamin Problem

We apply now our adapted method of learning causal information to Douven's examples and the Judy Benjamin Problem.

4.1.1. A possible worlds model for the Sundowners Example

Example 1. The Sundowners Example (Douven & Romeijn 2011, 645-646)

Sarah and her sister Marian have arranged to go for sundowners at the Westcliff hotel tomorrow. Sarah feels there is some chance that it will rain, but thinks they can always enjoy the view from inside. To make sure, Marian consults the staff at the Westcliff hotel and finds out that in the event of rain, the inside area will be occupied by a wedding party. So she tells Sarah:

(6) If it rains tomorrow, we cannot have sundowners at the Westcliff.

Upon learning this conditional, Sarah sets her probability for sundowners and rain to 0, but she does not adapt her probability for rain.

We model Sarah's belief state as the Stalnaker model depicted in Figure 4. W contains four elements covering the possible events of R, $\neg R$, S, $\neg S$, where R stands for "it rains tomorrow" and S for "Sarah and Marian can have sundowners at the Westcliff tomorrow".

Let us assume that Sarah interprets the conditional uttered by her sister Marian as conveying the causal information $R \Rightarrow \neg S$. As Douven himself points out, the intuition in the Sundowners Example derives from the verdict that whether or not it rains makes the difference as to whether or not they have sundowners, but not the other way around: having sundowners simply has no effect whatsoever on whether or not it rains (cf. Douven

2012, 8). Hence, the change of belief between R and $\neg R$ is more far-fetched than between S and $\neg S$. In other words, the worlds along the horizontal axis are more similar than the worlds along the vertical axis. Since $R \Rightarrow \neg S \equiv (R > \neg S) \wedge (\neg R > S)$, $R \Rightarrow \neg S$ expresses both that S is excluded under the supposition of R and $\neg S$ is excluded under the supposition of $\neg R$. By the causal difference assumption, we obtain $\min_{\leq w1} [R > \neg S] = w_2$ and $\min_{\leq w4} [\neg R > S] = w_3$. Lewis's imaging method results in a shift of probability shares along the horizontal axis of Figure 4.

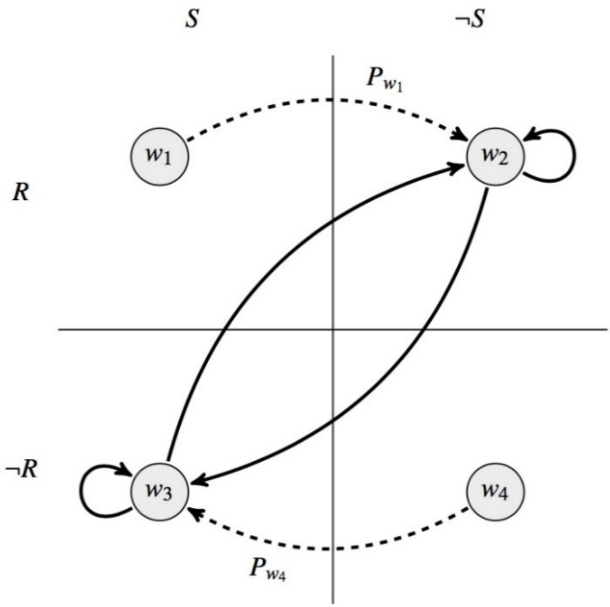

Figure 4: A Stalnaker model for Sarah's belief state in the Sundowners Example. The thick arrows illustrate the change of the similarity order such that the received information, causally understood as $R \Rightarrow \neg S$, is minimally informative. Here, the minimally informative meaning of $R \Rightarrow \neg S$ is $[R \Rightarrow \neg S] = [R > \neg S] \cap [\neg R > S] = \{w_2, w_3\}$. The dashed arrows represent the respective transfers of probability.

Imaging on the minimally informative proposition $[R \Rightarrow \neg S] = \{w_2, w_3\}$ results in $P^{R\Rightarrow\neg S}(w') = \sum_w P(w) \cdot \begin{cases} 1 & if\ w_{R\Rightarrow\neg S} = w' \\ 0 & \text{otherwise} \end{cases}$:

(7) $P^{R\Rightarrow\neg S}(w_1) = 0$

$P^{R\Rightarrow\neg S}(w_2) = P(w_1) + P(w_2)$

$P^{R\Rightarrow\neg S}(w_3) = P(w_3) + P(w_4)$

$P^{R\Rightarrow\neg S}(w_4) = 0$

We see immediately that both intuitions associated with the Sundowners Example are satisfied, viz. $P^{R\Rightarrow\neg S}(R) = P(R) = P(w_1) + P(w_2)$ and $P^{R\Rightarrow\neg S}(R \wedge S) = P(w_1) = 0$. We conclude that the method of learning causal information yields the intuitively correct results.[12]

4.1.2. A possible worlds model for the Ski Trip Example

Example 2. The Ski Trip Example (Douven & Dietz 2011, 33)

Harry sees his friend Sue buying a skiing outfit. This surprises him a bit, because he did not know of any plans of hers to go on a skiing trip. He knows that she recently had an important exam and thinks it unlikely that she passed. Then he meets Tom, his best friend and also a friend of Sue, who is just on his way to Sue to hear whether she passed the exam, and who tells him:

(8) If Sue passed the exam, her father will take her on a skiing vacation.

Recalling his earlier observation, Harry now comes to find it more likely that Sue passed the exam.

We model Harry's belief state as the Stalnaker model depicted in Figure 5. *W* contains eight elements covering the possible events of *E*, ¬*E*, *S*, ¬*S*,

[12] Note that the Sundowners Example seems to be somewhat artificial. It seems plausible that upon hearing her sister's conditional, Sarah would promptly ask 'why?' in order to obtain some more contextual information, before setting her probability for sundowners and rain to 0. After all, she 'thinks that they can always enjoy the view from inside'.

B, $\neg B$, where E stands for "Sue passed the exam", S for "Sue's father takes her on a skiing vacation", and B for "Sue buys a skiing outfit".

We assume that Harry interprets the conditional uttered by his friend Tom as conveying the causal information $E \Rightarrow S$. Furthermore, The Ski Trip Example assumes that Harry is equipped with the following contextual knowledge: Sue buying a skiing outfit may causally depend on the invitation of Sue's father to a skiing vacation, in symbols $S \Rightarrow B$. Finally, Harry observed Sue buying a skiing outfit, and thus has the factual information that B.

In total, Harry learns the minimally informative proposition $[(E \Rightarrow S) \wedge (S \Rightarrow B) \wedge B] = \{w_1\}$. Since w_1 is the only world that is not probabilistically excluded, we do not need to appeal to the causal difference assumption in this example.

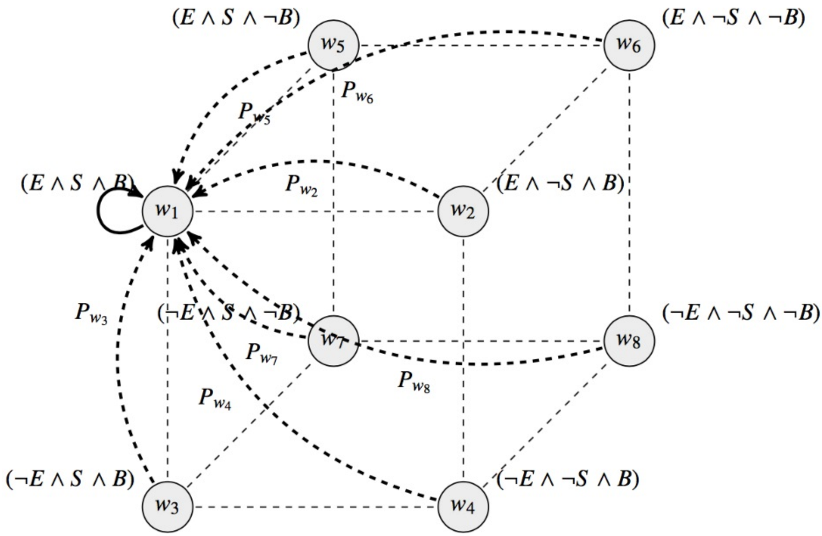

Figure 5: An eight-worlds Stalnaker model for Harry's belief state in the Ski Trip Example. Harry learns the minimally informative proposition $[(E \Rightarrow S) \wedge (S \Rightarrow B)] = \{w \in W \mid (\min_{w\leq} [E] \in [S]) \wedge (\min_{w\leq} [\neg E] \in [\neg S]) \wedge (\min_{w\leq} [S] \in [B]) \wedge (\min_{w\leq} [\neg S] \in [\neg B])\} = \{w_1, w_8\}$. Since Harry also obtains the factual information B, we can also exclude the $\neg B$-world w_8. (The arrows follow the convention of Figure 4.)

Imaging on the minimally informative proposition $[(E \Rightarrow S) \wedge (S \Rightarrow B) \wedge B] = \{w_1\}$ results in the following probability distribution, where we do not display the vanishing probabilities:

$$P^{(E\Rightarrow S)\wedge(S\Rightarrow B)\wedge B}(w') = P^*(w') = \sum_w P(w) \cdot \begin{cases} 1 & \text{if } w_{(E\Rightarrow S)\wedge(S\Rightarrow B)\wedge B} = w' \\ 0 & \text{otherwise} \end{cases}:$$

(9) $P^*(w_1) = 1$

The result meets the intuition associated with the Ski Trip Example: $P^*(E) > P(E)$, since $P^*(E) = P^*(w_1)$ and $P(E) = P(w_1) + P(w_2) + P(w_5) + P(w_6)$. Later on, we will see that the probabilities of the worlds w_2, w_3, w_4 would not have vanished entirely, if either $E \Rightarrow S$ or $S \Rightarrow B$ (or both) had conveyed only uncertain information.

In Günther (2017), we needed the default assumption to model the Ski Trip Example. If we appeal to the causal interpretation in the Ski Trip Example, we do neither need the default nor the causal difference assumption any more.

4.1.3. A possible worlds model for the Driving Test Example

Example 3. The Driving Test Example (Douven 2012, 3)

Betty knows that Kevin, the son of her neighbours, was to take his driving test yesterday. She has no idea whether or not Kevin is a good driver; she deems it about as likely as not that Kevin passed the test. Betty notices that her neighbours have started to spade their garden. Then her mother, who is friends with Kevin's parents, calls her and tells her the following:

(10) If Kevin passed the driving test, his parents will throw a garden party.

Betty figures that, given the spading that has just begun, it is doubtful (even if not wholly excluded) that a party can be held in the garden of Kevin's parents in the near future. As a result, Betty lowers her degree of belief for Kevin's having passed the driving test.

We model Betty's belief state as the Stalnaker model depicted in Figure 6. W contains eight elements covering the possible events of D, $\neg D$, G, $\neg G$,

S, $\neg S$, where D stands for "Kevin passed the driving test", G for "Kevin's parents will throw a garden party", and S for "Kevin's parents have started to spade their garden".

Assume Betty interprets the conditional uttered by her mother as the causal information $D \Rightarrow G$. Furthermore, Betty infers from her contextual knowledge that because Kevin's parents are spading their garden, they will not throw a garden party, in symbols $S \Rightarrow \neg G$. Finally, Betty knows that Kevin's parents have started to spade their garden, and thus has the factual information that S.

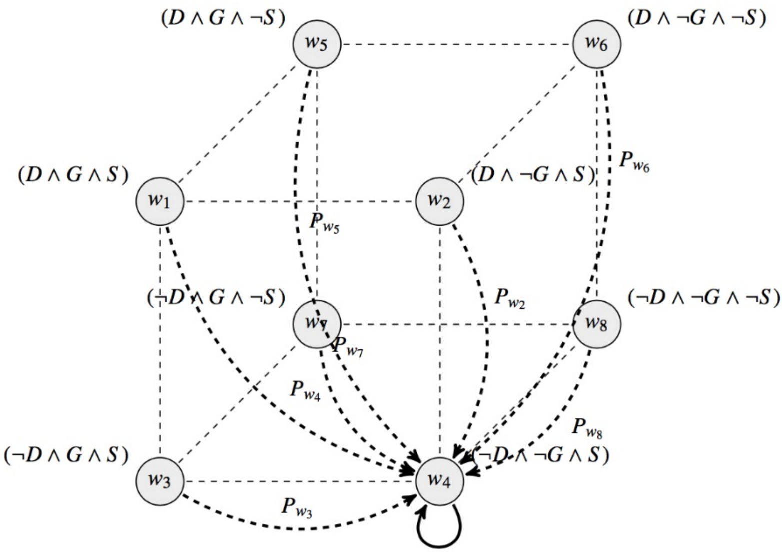

Figure 6: An eight-worlds Stalnaker model for Betty's belief state in the Driving Test Example.

In total, Betty learns the minimally informative proposition $[(D \Rightarrow G) \wedge (S \Rightarrow \neg G) \wedge S] = \{w_4\}$. In Figure 6, we see that the Driving Test Example is structurally similar to the Ski Trip Example.

Imaging on the minimally informative proposition $[(D \Rightarrow G) \wedge (S \Rightarrow \neg G) \wedge S] = \{w_4\}$ results in the following probability distribution, where we do not display the vanishing probabilities:

$$P^{(D\Rightarrow G)\wedge(S\Rightarrow\neg G)\wedge S}(w') = P^*(w') = \sum_w P(w)\cdot \begin{Bmatrix} 1 & if\ w_{(D\Rightarrow G)\wedge(S\Rightarrow\neg G)\wedge S} = w' \\ 0 & \text{otherwise} \end{Bmatrix}:$$

(11) $P^*(w_4) = 1$

Our method yields again the correct result regarding the intuition associated with the Driving Test Example: $P^*(D) < P(D)$, since $P^*(D) = 0$ and $P(D) = P(w_1) + P(w_2) + P(w_5) + P(w_6) > 0$.

The following Judy Benjamin Problem will illustrate that if Betty thinks that the conditionals $D \Rightarrow G$ or $S \Rightarrow \neg G$ (or both) convey uncertain information, then the probability shares for some other worlds will not reduce to zero. This fact fits nicely with the Driving Test Examples's remark that "given the spading that has just begun, it is doubtful [or uncertain] (even if not wholly excluded) that a party can be held in the garden of Kevin's parents". We will treat the application of our method to the learning of uncertain causal information in the next section.

4.1.4. A possible worlds model for the Judy Benjamin Problem

We apply now our method of learning causal information to a case, in which the received causal information is uncertain. We show thereby that the method may be generalised to those cases in which the learned causal information is uncertain, provided we use Jeffrey imaging. Following the presentation in Hartmann & Rad (2017), we consider Bas van Fraassen's Judy Benjamin Problem (cf. van Fraassen 1981, 376-379).

Example 4. The Judy Benjamin Problem (Hartmann & Rad 2017, 7))

A soldier, Judy Benjamin, is dropped with her platoon in a territory that is divided in two halves, Red territory and Blue territory, respectively, with each territory in turn being divided in equal parts, Second Company area and Headquarters Company area, thus forming four quadrants of roughly equal size. Because the platoon was dropped more or less at the center of the whole territory, Judy Benjamin deems it equally likely that they are in one quadrant as that they are in any of the others. They then receive the following radio message:

> (12) I can't be sure where you are. If you are in Red Territory, then the odds are 3 : 1 that you are in Second Company area.

After this, the radio contact breaks down. Supposing that Judy accepts this message, how should she adjust her degrees of belief?

Douven claims that the probability of being in red territory should, intuitively, remain unchanged after learning the uncertain information. Furthermore, the probability distribution after hearing the radio message, i. e. P*, should take the following values:

(13) $P^*(R \wedge S) = \frac{3}{8}$ $\quad P^*(R \wedge \neg S) = \frac{1}{8}$

$P^*(\neg R \wedge S) = \frac{1}{4}$ $\quad P^*(\neg R \wedge \neg S) = \frac{1}{4}$

We model Judy Benjamin's belief state as the Stalnaker model depicted in Figure 7. W contains four elements covering the possible events of R, $\neg R$, S, $\neg S$, where R stands for "Judy Benjamin's platoon is in Red territory", and S for "Judy Benjamin's platoon is in Second Company area". The story prescribes that the probability distribution before learning the uncertain information is given by:

(14) $P(R \wedge S) = P(R \wedge \neg S) = P(\neg R \wedge S) = P(\neg R \wedge \neg S) = \frac{1}{4}$

In the previous examples, our agents implicitly learned Stalnaker conditionals of the form $\alpha > \gamma$ with certainty. According to Theorem 1, this amounts to the constraint that $P(\alpha > \gamma) = P^{\alpha}(\gamma) = 1$ (provided α is not a contradiction). Given this constraint and since P^{α} is a probability distribution, we have $P^{\alpha}(\neg\gamma) = 1 - P^{\alpha}(\gamma) = 0$. This means that we were able to probabilistically exclude any $\neg\gamma$-world under the supposition of α.

Now, our agent Judy Benjamin learns uncertain causal information, i.e. she implicitly learns Stalnaker conditionals with uncertainty. According to Theorem 1 and since $R \Rightarrow S$ is equivalent to $(R > S) \wedge (\neg R > \neg S)$, this amounts in the Judy Benjamin Problem to the constraint that $P(R \Rightarrow S) = \frac{3}{4}$.

By our method, we obtain $P(R \Rightarrow \neg S) = \frac{1}{4}$. In contrast to learning causal information with certainty, we cannot subtract the whole probabilistic mass from the $\neg S$ -worlds under the supposition of R, and accordingly from the S-worlds under the supposition of $\neg R$. However, Judy Benjamin is informed from an external source about the proportion to which she should gradually 'exclude' or downweigh the probability share of $R \Rightarrow \neg S$-worlds. Equivalently, we may say that the most similar $R \Rightarrow S$-world (from any $R \Rightarrow \neg S$)-world) obtains a gradual upweight of probability such that it receives $\frac{3}{4}$ of the probability shares of the $R \Rightarrow \neg S$-worlds; in turn, however, this $R \Rightarrow \neg S$-world then receives a probability share from the $R \Rightarrow S$-world weighed by $\frac{1}{4}$. Note that in Stalnaker models $R \Rightarrow \neg S$ is equivalent to $R > \neg S \wedge \neg R > S$.

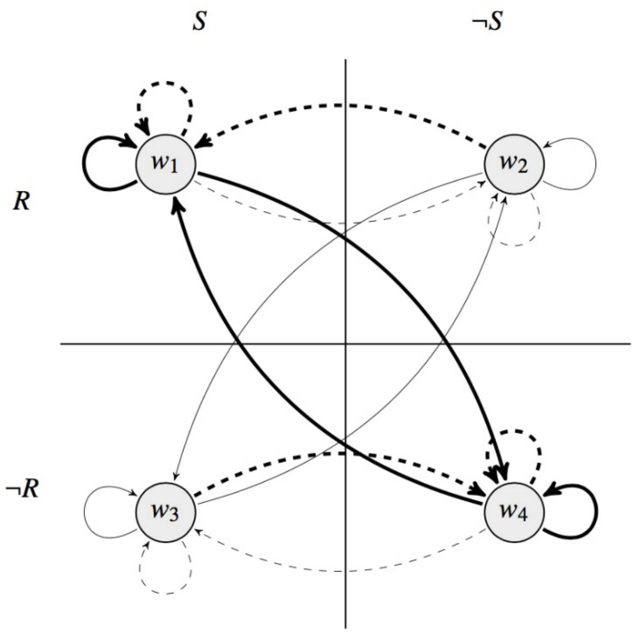

Figure 7: A Stalnaker model for Private Benjamin's belief state in the Judy Benjamin Problem. The thick arrows illustrate the specification of a similarity order $\leq'$ such that the received information $[R \Rightarrow S]$ is minimally informative. Note that each world having two outgoing thick arrows (one

for $R > S$ and one for $\neg R > \neg S$) satisfies $R \Rightarrow S$. The thin arrows illustrate the specification of another similarity order $\leq \neq \leq'$ such that the received information $[R \Rightarrow \neg S]$ is minimally informative. Each world having two outgoing thin arrows (one for $R > \neg S$ and one for $\neg R > S$) satisfies $R \Rightarrow \neg S$). In sum, the similarity orders are specified such that one makes $[R \Rightarrow S] = \{w_1, w_4\}$ a minimally informative proposition and the other makes the complement proposition $[R \Rightarrow \neg S] = \{w_2, w_3\}$ a minimally informative proposition. By the causal difference assumption, we obtain $min_{\leq'_{w2}}$ $[R \Rightarrow S] = w_1$ and $min_{\leq'_{w3}}[R \Rightarrow S] = w_4$. Furthermore, we obtain $min_{\leq_{w1}}$ $[R \Rightarrow \neg S] = w_2$ and $min_{\leq_{w4}}$ $[R \Rightarrow \neg S] = w_3$. The thick dashed arrows represent the transfer of $k \cdot P(w)$, while the thin dashed arrows represent the transfer of $(1 - k) \cdot P(w)$. The application of Jeffrey imaging on $[R \Rightarrow S]$ with $k = \frac{3}{4}$ leads to the following calculation for the probability distribution: $P^{R\Rightarrow S}_{3/4}(w_1) = 3/4 \cdot P(w_1) + 3/4 \cdot P(w_2)$, and $P^{R\Rightarrow S}_{3/4}(w_2) = 1/4 \cdot P(w_1) + 1/4 \cdot P(w_2)$, and $P^{R\Rightarrow S}_{3/4}(w_3) = 1/4 \cdot P(w_3) + 1/4 \cdot P(w_4)$, and $P^{R\Rightarrow S}_{3/4}(w_4) = 3/4 \cdot P(w_3) + 3/4 \cdot P(w_4)$.

We apply now Jeffrey imaging to the Judy Benjamin Problem, where a source external to Judy provides her with the information that $k = \frac{3}{4}$.

(15) $$P^{R\Rightarrow S}_k(w') = \sum_w \left(P(w) \cdot \begin{cases} k \text{ if } w_{R\Rightarrow S} = w' \\ 0 \quad \text{otherwise} \end{cases} + P(w) \cdot \begin{cases} 1-k \text{ if } w_{R\Rightarrow \neg S} = w' \\ 0 \quad \text{otherwise} \end{cases} \right)$$

Given the probability distribution before the learning process in Equation (14), Judy obtains the following probability distribution after being informed that $P(R \Rightarrow S) = \frac{3}{4}$:

(16) $$P^{R\Rightarrow S}_{3/4}(w_1) = P^{R\Rightarrow S}_{3/4}(R \wedge S) = \frac{3}{8}$$
$$P^{R\Rightarrow S}_{3/4}(w_2) = P^{R\Rightarrow S}_{3/4}(R \wedge \neg S) = \frac{1}{8}$$
$$P^{R\Rightarrow S}_{3/4}(w_3) = P^{R\Rightarrow S}_{3/4}(\neg R \wedge S) = \frac{1}{8}$$
$$P^{R\Rightarrow S}_{3/4}(w_4) = P^{R\Rightarrow S}_{3/4}(\neg R \wedge \neg S) = \frac{3}{8}$$

The probability distribution of (16) does not conform to Douven's intuitively correct distribution of (13), while the desideratum $P_{3/4}^{R>S}(R) = P(R) = \frac{1}{2}$ is met. Note that the learning of causal information results in $P_{3/4}^{R\Rightarrow S}(\neg R \wedge \neg S) = \frac{3}{8}$, which may be plausible for cases of causal dependence. However, we do not think that the conditional of the Judy Benjamin Problem is meant to express a causal dependence relation. In Günther (2017), we treated the received uncertain conditional as merely carrying uncertain conditional information. Applying the method of learning uncertain conditional information allowed us to offer a solution to the Judy Benjamin Problem that agrees with Douven's desired distribution of (13).

The Judy Benjamin Problem illustrates quite vividly the main difference between learning conditional and causal information. A merely conditional understanding of the conditional in the Judy Benjamin Problem does not affect the (row of) $\neg\alpha$-worlds, whereas the difference-making or causal dependence interpretation of the conditional affects the (row of) $\neg\alpha$-worlds.

5. Stalnaker inferences to the explanatory status of the antecedent

The method of learning causal information provides a formally precise implementation for when and how Douven's explanatory status of the antecedent should change. Recall his idea from Section 2 that the explanatory power of the antecedent with respect to the consequent determines the probability of the antecedent after learning the conditional. The idea is related to abduction, nowadays more commonly referred to as 'inference to the best explanation', or at least to a good explanation. The schema of such an inference runs as follows: α explains γ (well), and γ obtains. Therefore, α is true, or at least more likely.

We may interpret a Stalnaker agent's learning of $\alpha \Rightarrow \gamma$ as inference to a good explanation. Suppose an agent believes the fact γ and receives the information $\alpha \Rightarrow \gamma$. Then the agent infers that α explains γ (well). For, $\alpha \Rightarrow \gamma$ implies that $\neg\gamma$ would be the case, if α were not the case. But γ is the case and thus indicates that α is the case as well. The Ski Trip Example is an instance of this type of reasoning. Harry learns $E \Rightarrow S$, $S \Rightarrow B$ and the fact B. He infers by our method of learning causal information that S explains B and, in turn, that E explains S. Consequently, $P^{(E \Rightarrow S) \wedge (S \Rightarrow B) \wedge B}(E) \geq P(E)$. In general, $P_k^{(\alpha \Rightarrow \gamma) \wedge \gamma}(\alpha) \geq P(\alpha)$, if $k > \frac{1}{2}$. In such a case, we call α the antecedent in a 'Stalnaker inference to a good explanatory status of the antecedent', or simply the antecedent in a 'Stalnaker inference to a good explanans'.

In the Driving Test Example, Kevin's passing the driving test (D) is at odds with the parent's spading their garden (S). D does not explain S (well). There is rather a tension between the occurrence of D and S. We can again formally implement the reasoning. Suppose S and $S \Rightarrow \neg G$, where G stands for "Kevin's parents will throw a garden party". Betty receives the information that $D \Rightarrow G$. S and $S \Rightarrow \neg G$ implies that G is not the case. By $D \Rightarrow G$, we may therefrom infer that D is not the case either. For, if D were the case, G would be the case. Consequently, $P^{(D \Rightarrow G) \wedge (S \Rightarrow \neg G) \wedge S}(D) \leq P(D)$. In general, $P_k^{(\alpha \Rightarrow \gamma) \wedge \neg \gamma}(\alpha) \leq P(\alpha)$, if $k > \frac{1}{2}$. In such a case, we call α the antecedent in a 'Stalnaker inference to a bad explanans'. Notice that our framework allows for a probabilification of the Stalnaker inferences, if uncertain causal information is learned.

6. Conclusion

We have seen that Douven's dismissal of the Stalnaker conditional as a tool to model the learning of conditional and causal information is unjustified. Rather, this type of learning may be modelled by Jeffrey imaging on the meaning of Stalnaker conditionals under the following condition: the similarity order of the Stalnaker model is changed in a way such that the meaning of the conditional is minimally informative. Both methods of learning information align with the intuitively correct results in Douven's benchmark examples. However, Douven's intuitions about the Judy Benjamin Problem are only met, if we understand the conditional Judy receives as conveying merely conditional information.

We have shown that the method of learning (uncertain) conditional information proposed in Günther (2017) may be adapted to a learning method of (uncertain) causal information. The adaptation is based on the Stalnaker conditional, for which Lewis's idea of causal dependence is implemented. The two methods come with two different assumptions, i.e. the default assumption and the causal difference assumption, respectively. The combination of the two methods provides a unified framework that manages to clearly discern between a merely conditional and a causal reading of the conditional "If α, then γ". Hence, the general method cannot be attacked for not being applicable to conditionals that (are supposed to) express causal dependences. In detail, if no further contextual information is available, conjunctive information is strictly more informative than causal information, which is in turn strictly more informative than conditional information. For, the minimally informative conjunctive, causal and conditional propositions stand in the following strict subset relation: $[\alpha \wedge \gamma] \subset [\alpha \Rightarrow \gamma] \subset [\alpha > \gamma]$.

The causal dependence reading can be used to formalise Douven's explanatory status of the antecedent. We thereby convey the explanatory status a precise formal meaning that may be used to operationalize Douven's idea that explanatory considerations play a core role in learning conditionals. Furthermore, the results suggest that we should distinguish between a merely conditional or suppositional interpretation and a causal dependence interpretation of a conditional. A supposition should not affect those cases, in which the antecedent is not satisfied, whereas a difference-making conditional should. Based on this distinction, we hope that the proposed framework can help psychologists of reasoning to provide an empirically adequate account of actual reasoning behaviour with respect to the learning of conditional and causal information.

The advantages of our unified framework of learning uncertain information, as compared to alternative accounts, will be assessed in a follow-up paper. We plan to compare our account in detail to Douven's account of learning conditional information and Bayesian accounts of learning conditionals. In particular, we will show that the Bayesian account of Hartmann & Rad (2017) – that minimizes the Kullback-Leibler divergence on a fixed Bayesian network – has severe problems to capture the merely conditional interpretation of conditionals. As a consequence the Judy Benjamin Problem remains troublesome for their account.

References

BRADLEY, R. (2005): Radical Probabilism and Bayesian Conditioning*. *Philosophy of Science* 72, No. 2, 342-364.

DOUVEN, I. (2012): Learning Conditional Information. *Mind & Language* 27, No. 3, 239-263.

DOUVEN, I. & DIETZ, R. (2011): A Puzzle about Stalnaker's Hypothesis. *Topoi* 30, No. 1, 31-37.

DOUVEN, I. & ROMEIJN, J.-W. (2011): A New Resolution of the Judy Benjamin Problem. *Mind* 120, No. 479, 637-670.

DOUVEN, I. & VERBRUGGE, S. (2010): The Adams Family. *Cognition* 117, No. 3, 302-318.

GÜNTHER, M. (2017): Learning Conditional Information by Jeffrey Imagining on Stalnaker Conditionals. *Journal of Philosophical Logic*, online first. DOI 10.1007/s10992-017-9452-z.

HARTMANN, S. & RAD, S. R. (2017): Learning Indicative Conditionals. *Unpublished manuscript*, available at: http://www.rafieerad.org/conditional.pdf.

LEWIS, D. (1973a): *Counterfactuals*. Blackwell Publishers.

LEWIS, D. (1973b): Counterfactuals and Comparative Possibility. *Journal of Philosophical Logic* 2, No. 4, 418-446.

LEWIS, D. (1973c): Causation. *Journal of Philosophy* 70, No. 17, 556-567.

LEWIS, D. (1976): Probabilities of Conditionals and Conditional Probabilities. *The Philosophical Review* 85, No. 3, 297-315.

PFEIFER, N. & DOUVEN, I. (2014): Formal Epistemology and the New Paradigm Psychology of Reasoning. *Review of Philosophy and Psychology* 5, No. 2, 199-221.

STALNAKER, R. C. (1975): A Theory of Conditionals. In: Sosa, E. (ed.): *Causation and Conditionals*. Oxford University Press, 165-179.

STALNAKER, R. C. & THOMASON, R. H. (1970): A Semantic Analysis of Conditional Logic. *Theoria* 36, No. 1, 23-42.

UNTERHUBER, M. (2013): *Possible Worlds Semantics for Indicative and Counterfactual Conditionals? A Formal Philosophical Inquiry into Chellas-Segerberg Semantics*. Ontos.

VAN FRAASSEN, B. C. (1981): A Problem for Relative Information Minimizers in Probability Kinematics. *The British Journal for the Philosophy of Science* 32, No. 4, 375-379.

ZHAO, J., CRUPI, V., TENTORI, K., FITELSON, B., & OSHERSON, D. (2012): Updating: Learning versus Supposing. *Cognition* 124, No. 3, 373-378.