Beliefs and Evidence in Justification Models

Baltag, A.; Fiutek, V.; Smets, S.

Publication date
2016

Document Version
Final published version

Published in
Advances in Modal Logic

Citation for published version (APA):
Baltag, A., Fiutek, V., & Smets, S. (2016). Beliefs and Evidence in Justification Models. In L. Beklemishev, S. Demri, & A. Máté (Eds.), *Advances in Modal Logic: AiML 11* (pp. 156-176). College Publications. http://www.aiml.net/volumes/volume11/Baltag-Fiutek-Smets.pdf

Beliefs and Evidence in Justification Models

Alexandru Baltag [1]

Institute for Logic, Language and Computation. University of Amsterdam
The Netherlands

Virginie Fiutek [2]

University of Lille 3
France

Sonja Smets [3]

Institute for Logic, Language and Computation. University of Amsterdam
The Netherlands

Abstract

In this paper we focus on the formal qualitative representation of an agent's evidence and justification in support of her beliefs and knowledge. Our formal setting is based on 'justification models', which we introduce as a generalization of the so-called 'evidence models' proposed by J. van Benthem and E. Pacuit in [18]. We use these structures to express how an agent's evidence supports her doxastic state, expressing as such an agent's justifiable beliefs. We study a number of specific classes of justification models as well as their relations. Overall, these structures are more general than the so-called plausibility models used to represent an agent's doxastic and epistemic states in [14,7,8]. We illustrate the models in this paper via examples and focus on the dynamics of justification models.

Keywords: Justification models, justifiable beliefs, doxastic logic, evidence logic

1 Introduction

We have gained inspiration from Keith Lehrer's informal analysis of defeasible knowledge in terms of "undefeated justified acceptance" [11]. Any formal representation of this epistemic concept will requires an intricate analysis that can link an agent's defeasible knowledge of a proposition to the sound justification that she has in support of it. We pursue this line of thought in this paper

1 A.Baltag@uva.nl

2 fiutek.virginie@gmail.com

3 S.J.L.Smets@uva.nl. The contribution of S. Smets to this work is funded by The European Research Council ERC-2011-STG No. 283963, part of the European Community's FP7.

and focus on the introduction of a formal system that can express the necessary relations between an agent's epistemic or doxastic state and the evidence or justification that supports it. The models we introduce in this paper are called "justification models" and we study these structures in the framework of Dynamic Epistemic Logic and its recent extensions that can deal with belief revision theory [14,7,8].

In the tradition of logics designed to handle evidence and beliefs, we follow in this paper the semantic account that was initiated in [18] in the context of neighborhood models while the work in [2,5,3,4] explores evidence and beliefs in the context of topological models. van Benthem and Pacuit's semantic approach to evidence in terms of neighborhood models allows them to deal with possibly false and possibly mutually inconsistent evidence. If we focus on the relation between evidence and the agent's beliefs as well as her belief dynamics, we observe that the belief revision policies modelled in the context of 'evidence models' [18] do not necessarily satisfy the AGM postulates of belief revision [1]. This is due to the fact that the preorder relation that can be induced on the possible worlds in these models is not total. van Benthem and Pacuit go on to show that (uniform) evidence models can be turned into partial plausibility models and they indicate how a partial plausibility model can be extended to an evidence model. In their later work, van Benthem, Pacuit and Fernandez-Duque [16,17] study different types of evidence models as well as their relations to plausibility models. Continuing this line of work, we study in this paper different types of justification models to represent evidence and beliefs and we explore the relations between these classes of models.

In this paper we first introduce our main formal system in section 2 and we explain how to enhance our structures with a plausibility relation over possible worlds in section 3. In section 4 we study a number of different classes of justification models. An important type of justification models is given by the class of (introspective) evidence models of [18]. In our overview we include plausibility models and show how they are related to justification models. We further indicate how a justification model can be mapped into a plausibility model which allows us to define a number of epistemic and doxastic attitudes of an agent. Next we introduce counting models and weighting models, proving that they can be considered as a special kind of justification models and we show that (introspective) evidence models match to a special kind of justification models.

After introducing different types of justification models in section 4, we study their dynamics in section 5. In particular we focus on the notion of update of a justification model. We study this update operation in each of the introduced classes of justification models.

Finally, we provide a language for a logical system containing a number of sound axioms that can establish a new logic of justifiable beliefs in section 6. Given this logical language, we can express a number of interesting philosophical concepts and properties about justifiable beliefs and defeasible knowledge. We conclude this paper with some final remarks in the last section.

2 Introduction to Justification Models

We start with the following example of a specific scenario:

Example 1. Our agent Alice, a biology student, investigates an animal which is unkown to her. Alice will form a belief about the animal in front of her on the basis of the evidence that she gathers from four different sources of information (from her colleagues). Her first information source tells Alice that 'the animal can swim'. The second source states that 'it is a non-flying bird'. The third source says 'it lays eggs' and the fourth source says 'it flies'.

In this example, the collection of evidence coming from the four sources is accessbile to Alice. Yet this doesn't mean that all the evidence that is accessible to Alice forms a consistent set. Indeed, in this example we assume that an animal either does or doesn't fly but can't do both, as such the second and fourth sources are contradicting each other. Alice's evidence is not conclusive, yet Alice can reason about her evidence and she knows the evidence that is accessible to her.

To provide a formal model of this example, we start with a possible worlds model in which a piece of *evidence* is represented as a set of possible worlds. Formally, we introduce a new type of model, called *justification model*, to capture the agent's evidence or justification as well as the agent's doxastic and epistemic state.

Definition 1 *A justification model $\mathcal{M}$ is a tuple $(S, E, \preceq, \|\cdot\|)$ consisting of a finite set S of states or so-called possible worlds, a family $E \subseteq \mathcal{P}(S)$ of non-empty subsets $e \subseteq S$ ($\emptyset \notin E$), called evidence (sets) such that S is itself an evidence set ($S \in E$). We call a body of evidence (or argument) any $F \subseteq E$ such that $\bigcap F \neq \emptyset$ and we denote by $\mathcal{E} \subseteq \mathcal{P}(E)$ the family of all bodies of evidence. Any justification model comes equipped with a standard valuation map $\|\cdot\|$ and a partial preorder $\preceq$ on $\mathcal{E}$ satisfying the following constraints:*

$$F \subseteq F' \Rightarrow F \preceq F'$$

$$F \preceq F', G \preceq G' \text{ and } F' \cap G' = \emptyset \Rightarrow F \cup G \preceq F' \cup G'$$

$$F \prec F', G \preceq G' \text{ and } F' \cap G' = \emptyset \Rightarrow F \cup G \prec F' \cup G'$$

where $F, F', G, G', F \cup G, F' \cup G'$ are bodies of evidence, i.e. consistent families of evidence sets.

Note that the empty family of evidence sets $\emptyset$ is still a body of evidence, since $\bigcap \emptyset = \{s \in S \mid \forall e \in E(e \in \emptyset \Rightarrow s \in e)\}$, $\emptyset$ is a consistent family of evidence sets.

The above introduced relation $\preceq$ is a partial preorder, connecting only the consistent families of evidence sets. Here we read $F \preceq G$ as the body of evidence G is (considered) to be at least as convincing or easier to accept (by some implicit agent) as the body of evidence F. Similarly, the strict version $F \prec G$ denotes that the body of evidence G is (considered to be) more convincing, easier to accept (by some implicit agent) than the body of evidence F. The conditions in definition 1 indicate that the introduced preorder $\preceq$ does not contradict the set-theoretic inclusion order on bodies of evidence. We can impose further conditions on $\preceq$ to obtain a total preorder by requiring that either $F \preceq F'$ or $F' \preceq F$. Such a justification model with a total preorder relation is called a *total justification model*. Note that in total justification models, all evidence sets are comparable.

In this paper we assume that the agent is introspective regarding her evidence. Informally this means that the agent knows what evidence is available to her. Without this assumption, it will be natural to replace the current family of evidence sets E by a relation $E \subseteq S \times \wp(S)$, which will coincides with the definition of the evidence relation by van Benthem and Pacuit in [18].

Example 1 continued. Figure 1 illustrates the justification model for the above example. In this figure we introduce the state space S and name the possible worlds s, t, u, v, w and x, each of which satisfies a given atomic proposition coming from the set {Whale, Pigeon, Goldfish, Pengiun, Emu, Bat}. The family E consists of four evidence sets $e_1 = \{s, t, u\}$, $e_2 = \{u, v\}$, $e_3 = \{u, v, t, w\}$ and $e_4 = \{w, x\}$ and S. The agent Alice is implicitly present and we assume that she has the different arguments (or bodies of evidence) e_1, e_2, e_3, e_4 at her disposal. The collection of Alice's arguments (or bodies of evidence) $\mathcal{E}$ includes besides $\emptyset$ a number of important arguments $\{e_1\}, \{e_2\}, \{e_3\}, \{e_4\}, \{e_1, e_2\}, \{e_1, e_3\}, \{e_3, e_4\}, \{e_1, e_2, e_3\}$. In particular, her body of evidence coming from sources 3 and 4 supports the hypothesis that the animal is a typical flying bird (e.g. pigeon), while Alice's body of evidence from sources 1, 2 and 3 supports the hypothesis that the animal is a penguin (non-fying bird, lays eggs and swims). Alice's arguments are ordered, e.g. $\{e_1\} \preceq \{e_1, e_2\}$, hence the argument that the animal is a whale is less convincing to her in the light of the evidence coming from sources 1 and 2.

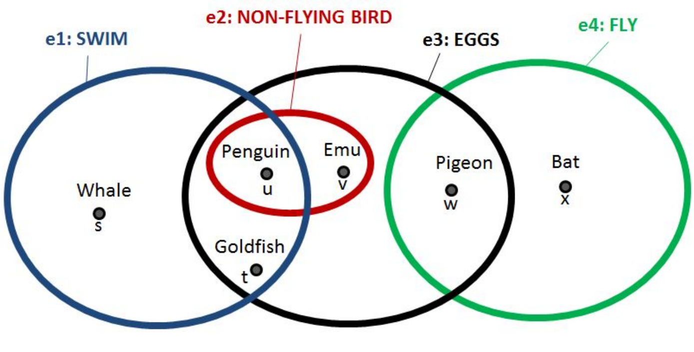

Fig. 1. Justification model for Example 1

3 Plausibility in Justification Models

In the literature on dynamic epistemic logic, a range of epistemic and doxastic attitudes of agents can be represented in the framework of so-called plausibilty models [7,8,14,19].Plausibility models are Kripke models in which the accessibility relation (or so-called plausibility relation) is given by a preorder $\leq \subseteq S \times S$ over the set of possible worlds. We will read the plausibility relation $s \leq t$ to capture that world s is at least as plausible as t. As argued for in [8,15,19], these type of models have a number of advantages over the well-known $KD45$ models in modal logic for the representation of doxastic states, especially in the context of belief dynamics and belief revision. Hence it is a natural step to investigate the options to introduce a plausibility relation over states within justification models. As we show next, this can be done in a canonical way.

We first introduce the notion of a *largest body of evidence consistent with a given state* $s \in S$ and denote it as

$$E_s := \{e \in E \mid s \in e\}$$

A plausibility relation on states can then be induced directly from the partial preorder on $\mathcal{E}$ as follows: For two states $s, t \in S$, we put

$$s \leq_E t \text{ iff } E_t \leq E_s$$

Example 1 continued. To illustrate this, we return to Figure 1 and observe that the largest body of evidence consistent with x is $E_x := \{e_4\}$ and the largest body of evidence consistent with w is $E_w := \{e_3, e_4\}$. Since $E_x \leq E_w$, we obtain $w \leq_E x$ in this example.

Epistemic and doxastic notions Given a justification model equipped with a plausibility relation, we can define all epistemic and doxastic notions usually defined on plausibility models [8,19]. This includes the notions of irrevocable knowledge (K), belief (B), conditional belief (B^{-}), strong belief (Sb) and defeasible knowledge (K_D), which we define below by using the plausibility order $\leq_E$ (and it's strict version $<_E$). In the following we use the notation $best_{\leq_E} P$ to denote the most plausible P-worlds in the plausiblity ordering, i.e. $best_{\leq_E} P = \mathrm{Min}_{\leq_E} P = \{s \in P |$ there is no $t < s$ for any $t \in P\}$. We abbreviate $best_{\leq_E} S$ as *best* to denote the set of most plausible states in the given state space S.

$$KP := \{s \in S : P = S\}$$
$$BP := \{s \in S : best_{\leq_E} \subseteq P\}$$
$$B^Q P := \{s \in S : best_{\leq_E} Q \subseteq P\}$$
$$SbP := \{s \in S : P \neq \emptyset \text{ and } t <_E w \text{ for all } t \in P \text{ and all } w \notin P\}$$
$$K_D P = \{s \in S : t \not>_E s \text{ implies } t \in P\}$$

Note that in case the plausibility order $\leq_E$ is a total relation, the following proposition holds:

$$s \models K_D P \text{ iff } s \models B^Q P \text{ for all } Q \text{ such that } s \models Q$$

In the last section we return to these epistemic and doxastic attitudes and investigate their link to an agent's arguments or bodies of evidence.

4 Special classes of justification models

Justification models provide a very general framework, subsuming a range of different existing settings. In this section we study the relations between different classes of justification models. In particular we show that partial and total plausibility models and the evidence models of [18] are a special class of justification models. We further introduce two other special classes of justification models, called counting models and weighting models. The following overview of the relations between justification models, counting models, weighting models, plausibility models and evidence models, illustrated in the following Figure 2, will be explained below:

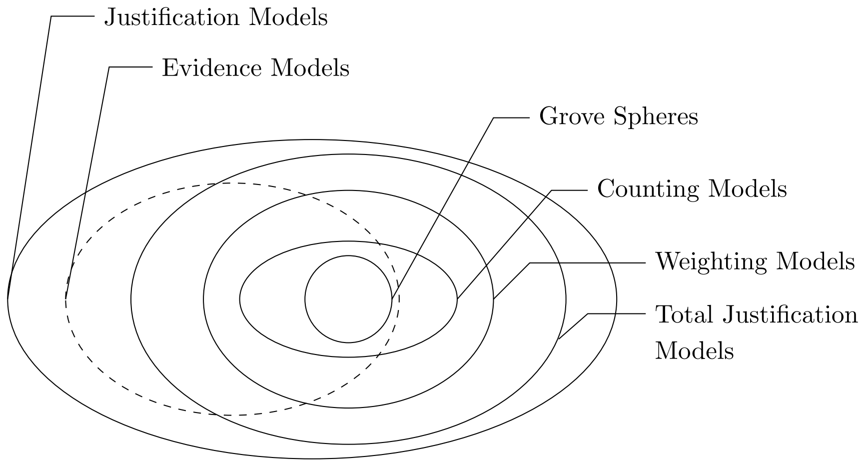

Fig. 2. Overview of Justification Models

4.1 Plausibility models

Any plausibility model $(S, \leq, \|\cdot\|)$, equipped with a preorder $\leq \subseteq S \times S$, can be viewed as a special kind of justification model $(S, E, \preceq, \|\cdot\|)$ in which S is the set of possible worlds and the set of evidence sets is given by $E = \{\downarrow w : w \in S\}$ where $\downarrow w = \{s \in S : s \leq w\}$. The preorder on bodies of evidence can be given by either one of the following options [4] :

(i) (Inclusion order) $F \preceq_1 F'$ iff $F \subseteq F'$ or

(ii) (Cardinality order) $F \preceq_2 F'$ iff $| F | \leq | F' |$.

In the second case, plausibility models are a special case of the counting models which we introduce below. In the first case, plausibility models $(S, \leq, \|\cdot\|)$ are a special kind of justification models $(S, E, \preceq_1, \|\cdot\|)$ in which the preorder on bodies of evidence is given by inclusion ($\preceq_1 = \subseteq$), the evidence sets are nested that is, $\forall e, e' \in E$ either $e \subseteq e'$ or $e' \subseteq e$ (i.e. the pre-order is a total

[4] Note that while the inclusion order is not necessarily total, the cardinality order is total.

pre-order). A body of evidence F corresponds to any family of spheres of a plausibility model [5] and $\mathcal{E}$ corresponds to all the families of spheres.

In line with the mentioned two cases, we can define two *plausibility maps* $Just_1$ and $Just_2$, mapping plausibility models to justification models:

- $(S, \leq, \|.\|) \overset{\text{Just1}}{\longmapsto} (S, E, \preceq_1, \|.\|)$
- $(S, \leq, \|.\|) \overset{\text{Just2}}{\longmapsto} (S, E, \preceq_2, \|.\|)$.

The plausibility map $Just_1$ corresponds to the case where the preorder on bodies of evidence is given by the inclusion order while the plausibility map $Just_2$ corresponds to the case where the preorder on bodies of evidence is given by the cardinality order. We call the justification models that can be obtained in one of these two ways (by applying $Just_1$ or $Just_2$ to a plausibility model), *sphere-based justification models.*

Reversly, any justification model $(S, E, \preceq, \|.\|)$ can be mapped into a type of plausibility model $(S, \leq, \|.\|)$. In order to do so, we define the *plausibility map* *Plau*, mapping justification models to plausibility models: $(S, E, \preceq, \|.\|) \overset{\text{Plau}}{\longmapsto} (S, \leq, \|.\|)$.

A justification model with a partial pre-order gives a partial plausibility model while a justification model with a total pre-order gives a total plausibility model. Following this, we observe that total justification models induce total plausibility models, i.e.

$$\forall F, F'(F \preceq F' \vee F' \preceq F) \Longleftrightarrow \forall s, s'(s \leq_E s' \vee s' \leq_E s)$$

Note that the map *Plau* mapping justification models to plausibility models is not an injective map. So, two different justification models can give rise to the same plausibility model. If we interpret a plausibility model $\mathcal{M}$ as a justification model $\mathcal{M}'$ and then apply the map *Plau*, we obtain the initial plausibility model $\mathcal{M}$. The converse is false since if we apply the map *Plau* on a justification model $\mathcal{M}'$ to obtain a plausibility model $\mathcal{M}$ and then interpret this plausibility model $\mathcal{M}$ as a justification model, we do not obtain the initial justification model $\mathcal{M}'$. Thus, we have both:

- $Plau(Just_1(\mathcal{M})) = \mathcal{M}$ for any plausibility model $\mathcal{M}$ and $Just_1(Plau(\mathcal{M}')) \neq \mathcal{M}'$ for any justification model $\mathcal{M}'$,
- $Plau(Just_2(\mathcal{M})) = \mathcal{M}$ for any plausibility model $\mathcal{M}$ and $Just_2(Plau(\mathcal{M}')) \neq \mathcal{M}'$ for any justification model $\mathcal{M}'$.

[5] We switch back and forth between the representation in terms of plausibility models and the equivalent setting in terms of Grove models, [10]. Semantically, the belief state of an agent can be modelled using families of sets of sets called spheres. This type of sphere model is built up from sets of possible worlds and so defines propositions as sets of possible worlds. The propositions believed by an agent (constituting her belief set) form the central sphere which is surrounded by concentric spheres, each of them representing a degree of similarity to the central sphere.

4.2 Counting models

We introduce the structure of *counting model* as follows. A counting model is a justification model $(S, E, \preceq, \|\cdot\|)$ in which the pre-order is given by the cardinality order, i.e. $F \preceq F'$ iff $| F |\leq| F' |$.[6]

In counting models, a body of evidence G is considered to be more convincing than a body of evidence F if and only if the number of evidence sets $e \in G$ is bigger than the number of evidence sets $e \in F$: $F \preceq G$ iff $| F |\leq| G |$. The intuition is that the more evidence the agent knows or has access too, the stronger is the evidence. This notion of evidence strength can be applicable to a number of scenarios, though one easily can provide examples where expressing strength of evidence in terms of counting will not be applicable.

Example 2. We represent an example of a counting model in Figure 3. In this example, we introduce four possible worlds namely s, t, v and w. We consider five evidence sets e_1, e_2, e_3, e_4 and e_5. Since the order on evidence is induced from the cardinality order, we have $F \preceq G$ iff $| F |\leq| G |$. As such we assign the following numbers to arguments: $| \{e_1\} |= 1$, $| \{e_3, e_4\} |= 2$, $| \{e_2, e_4\} |= 2$ and $| \{e_2, e_3, e_5\} |= 3$. This yields the following preorder over arguments, indicating their strenght $\{e_2, e_4\} \prec \{e_2, e_3, e_5\}$, $\{e_3, e_4\} \preceq \{e_2, e_4\}$, $\{e_1\} \prec \{e_3, e_4\}$.

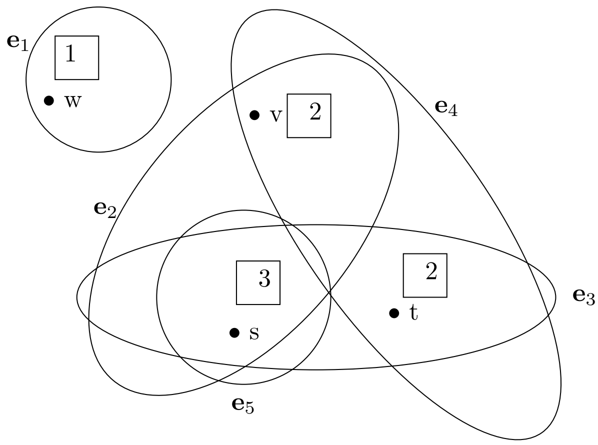

Fig. 3. Counting model

4.3 Weighting models

A special class of justification models is what we call 'Weighting models' these are structures $(S, E, f, \|\cdot\|)$ where we assign a weight (in terms of a natural number) to each piece of evidence $f : E \to \mathbb{N}$.

[6] Note that cardinality generates a total pre-order.

Let $(S, E, f, \|\cdot\|)$ be a weighting model. The function f can be extended to bodies of evidence $\mathcal{E}$ such that $f(E) = \sum_{e \in E} f(e)$. Given the weights of bodies of evidence, it is natural to introduce the preorder on bodies of evidence as follows:

$$E \preceq_f E' \text{ iff } f(E) \leq f(E')$$

As such any weighting model endowed with $\preceq_f$ is a justification model.

Example 3. We represent an example of a weighting model in Figure 4. In this example, there are four possible worlds namely s, t, v and w and five evidence sets e_1, e_2, e_3, e_4 and e_5. The model comes equipped with the function $f : E \to \mathbb{N}$ defined such that $f(e_1) = 1, f(e_2) = 2, f(e_3) = 1, f(e_4) = 3$ and $f(e_5) = 3$. We calculate that $f(e_1) = 1$, $f(\{e_3, e_4\}) = 4$, $f(\{e_2, e_4\}) = 5$ and $f(\{e_2, e_3, e_5\}) = 6$ and this then yields the following preorder over arguments $\{e_2, e_4\} \prec_f \{e_2, e_3, e_5\}$, $\{e_3, e_4\} \prec_f \{e_2, e_4\}$, $\{e_1\} \prec_f \{e_3, e_4\}$.

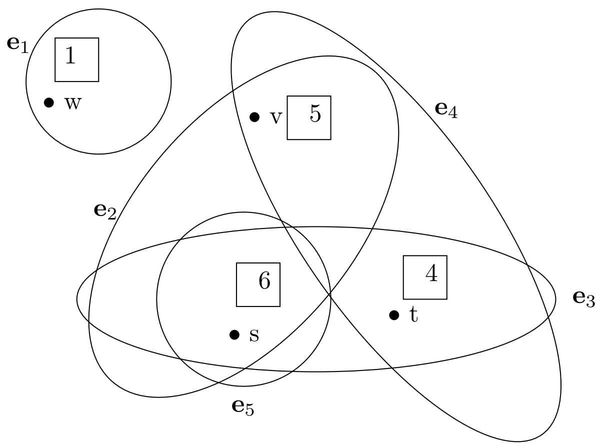

Fig. 4. Weighting model with $f(e_1) = 1$,$f(e_2) = 2$,$f(e_3) = 1$,$f(e_4) = 3$,$f(e_5) = 3$

One observes that counting models are a special case of weighting models in which $f(e) = 1$ for all $e \in E$.

Proposition 1. Every weighting model $(S, E, f, \|\cdot\|)$ is a justification model $(S, E, \preceq, \|.\|)$.

In order to prove this, we need to show that weighting models (and so counting models) satisfy the three constraints which the preorder must satisfy for it to be a justification model.

Proof. First note that since we use the order on natural numbers, transitivity follows.

- if $F \subseteq F'$ then $f(F) = \sum_{e \in F} f(e) \leq \sum_{e \in F'} f(e) = f(F')$ that is, $F \preceq F'$.

- Let $F \preceq F', G \preceq G'$ and $F' \cap G' = \emptyset$.
  $F \cup G = \sum_{e \in F \cup G} f(e) = \sum_{e \in F} f(e) + \sum_{e \in G} f(e) - \sum_{e \in F \cap G} f(e)$.
  Moreover $F' \cup G' = \sum_{e \in F' \cup G'} f(e) = \sum_{e \in F'} f(e) + \sum_{e \in G'} f(e) - \sum_{e \in F \cap G} f(e)$.
  Since $F' \cap G' = \emptyset$ then $F' \cup G' = \sum_{e \in F' \cup G'} f(e) = \sum_{e \in F'} f(e) + \sum_{e \in G'} f(e)$.
  Since $F \preceq F', G \preceq G'$ then $f(F) + f(G) - \sum_{e \in F \cap G} f(e) \leq f(F') + f(G')$.
  Hence $F \cup G \preceq F' \cup G'$.
- Let $F \prec F', G \preceq G'$ and $F' \cap G' = \emptyset$.
  $F \cup G = \sum_{e \in F \cup G} f(e) = \sum_{e \in F} f(e) + \sum_{e \in G} f(e) - \sum_{e \in F \cap G} f(e)$.
  Moreover $F' \cup G' = \sum_{e \in F' \cup G'} f(e) = \sum_{e \in F'} f(e) + \sum_{e \in G'} f(e) - \sum_{e \in F \cap G} f(e)$.
  Since $F' \cap G' = \emptyset$ then $F' \cup G' = \sum_{e \in F' \cup G'} f(e) = \sum_{e \in F'} f(e) + \sum_{e \in G'} f(e)$.
  Then $f(F) + f(G) - \sum_{e \in F \cap G} f(e) < f(F') + f(G')$.
  Hence $F \cup G \prec F' \cup G'$.

4.4 Evidence models

In [18], Johan van Benthem and Eric Pacuit introduce their so-called 'evidence models'. These models are based on the well-known neighbourhood semantics for modal logic in which the neighbourhoods are interpreted as evidence sets: pieces of evidence (possibly false, possibly mutually inconsistent) possessed by the agent. It is important to note that the plausibility relation that can be induced on possible worlds in evidence models is not a total preorder. Hence in these models, not all possible worlds are comparable.

Definition 2 *An evidence model M is a tuple $(S, E, ||.||)$ consisting of a non-empty set of worlds S. E is an evidence relation $E \subseteq S \times \mathcal{P}(\mathcal{S})$ and $||.||$ is a standard valuation function.*

The collection of evidence sets is defined as $E(s) = \{X | sEX, X \subseteq S\}$
we impose two constraints on the evidence function:
(Cons) For each state s, $\emptyset \notin E(s)$
(Triv) For each state s, $S \in E(s)$
These constraints ensure that no evidence set is empty and that the universe S is itself an evidence set. In this framework, the combination of different evidence sets does not necessarily yield consistent evidence. Indeed for any two evidence sets X and Y , X and Y may be disjoints sets that is, $X \cap Y = \emptyset$.

To investigate the relation between evidence models and justification models, we first introduce the notion of an *introspective evidence model*:

Definition 3 *An evidence model $\mathcal{M}$ is introspective iff we have sEX iff tEX for all $s, t \in S$ and for all $X \subseteq S$.*

Introspective evidence models are a special kind of justification models namely, they correspond exactly to those justification models in which the pre-order on bodies of evidence is given by the inclusion order.

In an introspective evidence model, the evidence relation E boils down to the concept of evidence used in justification models, that is, it becomes a family of evidence sets $E \subseteq \mathcal{P}(S)$ such that $E_s = \{e \mid s \in E\}$ for any $s \in S$. Moreover, the notions of irrevocable knowledge (K), belief (B) and conditional belief (B^-) defined in evidence models in [18] do exactly correspond to the notions we defined earlier in this paper.

4.5 Important notions in justification models

For a given argument $F \in \mathcal{E}$, an evidence set $e \in E$ and state $s \in S$ in a justification model, we can define a number of philosophical concepts indicating when an argument is sound, when an argument (conditionally) supports a proposition and when we have a (conditional) justification for a proposition:

Definition 4 *An argument F is sound at s iff $s \in \bigcap F$.*

Note that the empty argument $\emptyset$ is always sound at every state s since $s \in \bigcap \emptyset = S$.

Definition 5 *An argument F supports Q (or F is an argument for Q) iff $\bigcap F \subseteq Q$.*

Definition 6 *A justification for Q is an argument F such that all arguments at least as strong as F support Q, i.e. $\forall F'(F \preceq F' \Rightarrow \bigcap F' \subseteq Q)$.*

Definition 7 *An argument F supports Q conditional on P (or F is an argument for Q conditional on P) iff $\bigcap F \cap P \subseteq Q$.*

Definition 8 *A justification for Q given P is an argument F that is consistent with P such that all arguments at least as strong as F support Q conditional on P, i.e. $\bigcap F \cap P \neq \emptyset$ and $\forall F'(F \preceq F' \Rightarrow \bigcap F' \cap P \subseteq Q)$.*

5 Dynamics of Justification Models

In line with the work on dynamic epistemic logic, we will model the dynamics of justification models as a model transforming operation. Such a model transformation is taken to be triggered by an epistemic or doxastic event. In line with the above examples, we consider the case in which an agent is confronted with new incoming information and accommodates this new information into her epistemic or doxastic state.

While different types of events can be studied, ranging from public announcements [12] to private announcements including truthfull ones and untruthfull ones [6,15], in this paper we restrict ourselves to *updates* of justification models in which the agent receives new truthful information:

Definition 9 *Given a justification model $\mathcal{M} = (S, E, \preceq, \|\cdot\|)$ and a subset $P \subseteq S$, we define the relativization of the justification model $\mathcal{M}$ to P as $\mathcal{M}|P = (S', E', \preceq', \|\cdot\|')$ with:*

$$S' = P$$

$$E' = \{e \cap S' \mid \ e \in E,\ e \cap S' \neq \emptyset\}$$

$$F' \preceq' G' \text{ iff } \{e \in E \mid e \cap S' \in F'\} \preceq \{e \in E \mid e \cap S' \in G'\}$$
$$\|\cdot\|' = \|\cdot\| \cap S'$$

In the updated model, the new set of states S' is reduced to the set of states satisfying P. The new evidence set E' is taken to be the old evidence E that is consistent with the states surviving the update and the order $\preceq'$ on new bodies of evidence $F' \preceq' G'$ reflects the fact that the new evidence within G' is at least as strong as the new evidence in F'.

We define the restriction $F|\mathcal{M}'$ of the argument F to the justification model $\mathcal{M}'$ as follows: If F is an argument in a justification model $\mathcal{M}$, and if $\mathcal{M}' = \mathcal{M}|P$ is the relativization of the justification model $\mathcal{M}$ to a subset P, then $F|\mathcal{M}'$ is a body of evidence for the model $\mathcal{M}'$ (called the restriction $F|\mathcal{M}'$ of the argument F to model $\mathcal{M}'$), defined by

$$F|\mathcal{M}' = \{e \cap S' \mid e \in F\}$$

5.1 Plausibility models and Evidence Models

The dynamics of evidence models is studied in [18] and the relativization of these models coincides with the update defined by van Benthem and Pacuit. When restricting to sphere-based justification models, our update operation coincides with the usual update on plausibility models where the plausibility order $\leq'_E$ in the resulting model is given as:

$$\leq'_E = \leq_E \cap (S' \times S')$$

For $s \in S'$, we have $E_s = \{e \in E \mid s \in e\}$ and $E'_s = \{e \cap S' \mid e \in E_s\}$. Note that from $e \in E_s$ and $s \in S'$, we have $s \in e \cap S' \neq \emptyset$.

For $s, t \in S'$, we have:

$$\begin{array}{rcll} s \leq'_E t & \iff & E'_t \preceq' E'_s & \iff \{e \cap S' \mid e \in E_t\} \preceq' \{e \cap S' \mid e \in E_s\} \\ & & & \iff \{e \cap S' \mid e \in E, t \in e\} \preceq' \{e \cap S' \mid e \in E, s \in e\} \\ & & & \iff \{e \mid e \in E, t \in e\} \preceq \{e \mid e \in E, s \in e\} \\ & & & \iff E_t \preceq E_s \\ & & & \iff s \leq_E t \end{array}$$

5.2 Counting models

It is interesting to study the dynamics of specific classes of justification models. In particular we observe that the class of counting models is not closed under the update operation. When a given counting model is updated, the result of updating yields a justification model but not necessarily a counting model.

The following example illustrates the problem:

Example 4. Consider the counting model depicted in Figure 5 where there are three pieces of evidence e_1, e_2 and e_3 and 4 states s, t, u and v. The most plausible states are the states s and u since $E_s := \{e_1, e_2\}$, $E_t := \{e_1\}$, $E_u := \{e_1, e_3\}$, $E_v := \{e_3\}$ and so $|E_t| < |E_s|$, $|E_v| < |E_s|$, $|E_t| < |E_u|$ and $|E_v| < |E_u|$.

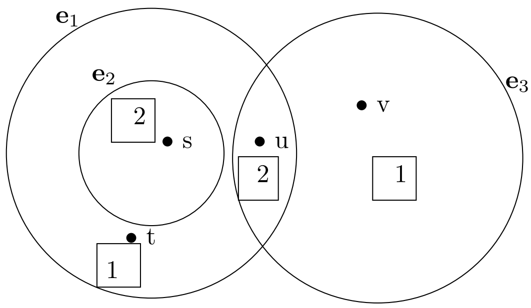

Fig. 5. Initial counting model

It is easily to see that a problem arises when dealing with certain updates of this model. Suppose that the implicit agent receives the hard information that P such that P is only true in s and v. Next, the model is updated with $!P$ and the states u and t are deleted as illustrated in Figure 6. Observe that s and v are equiplausible after the update action, since $E_v := \{e_3\}$, $E_s := \{e_4\}$ and so $\mid E_s \mid=\mid E_w \mid$.

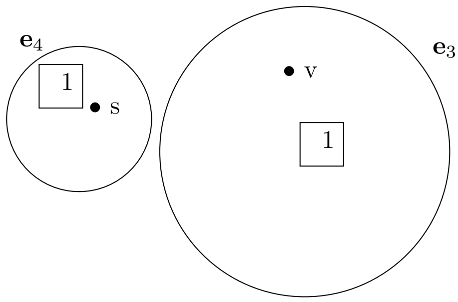

Fig. 6. Updated Counting Model

Due to the update we now lost the information that originally $\mid E_v \mid<\mid E_s \mid$, indeed one would have expected to obtain the justification model as depicted in Figure 7 yet no propositions are present to distinghuish e_1 and e_2 to make it possible that they can both be counted.

The problem is clearly visible if we work with plausibility models. We first provide the corresponding initial (total) plausibility model in Figure 8, and then represent the updated plausibility model after the update with P in Figure 9. After the update, the state s is still more plausible than the state v.

5.3 Weighting models

A solution to the problem in the previous subsection can be provided by working with weighting models instead of counting models.

We define the *map* Wei, mapping weighting models to justification models as follows: $(S, E, f, \|\cdot\|) \overset{\text{Wei}}{\longmapsto} (S, \leq, \|.\|)$.

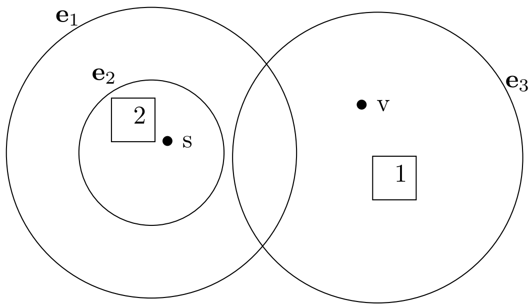

Fig. 7. Labelling solution

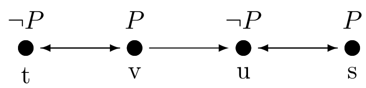

Fig. 8. Initial plausibility model

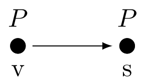

Fig. 9. Plausibility model updated

In contrast to counting models, we can see that the class of weighting models is better behaved:

Proposition 2. The class of weighting models is closed under the operation of updates

$$Wei(\mathcal{M})|\varphi = Wei(\mathcal{M}|\varphi)$$

Proof. Let $(S, E, f, \|\cdot\|)$ be a weighting model $\mathcal{M}$ where $f : E \to \mathbb{N}$. We define the result of updating this model with φ. The weighting model $\mathcal{M} = (S, E, f, \|\cdot\|)$ is changed to the weighting model $\mathcal{M}|\varphi = (S', E', f', \|\cdot\|')$ with:

$$S' = \|\varphi\|_{\mathcal{S}}$$

$$E' = \{e \cap S' \mid \; e \in E, \; e \cap S' \neq \emptyset\}$$

$$f'(e') = \sum \{f(e) \mid e \in E \text{ such that } e \cap S' \neq \emptyset\}$$

$$\|\cdot\|' = \|\cdot\| \cap S'$$

Coming back to Figure 5, let us put $f(e_1) = f(e_2) = f(e_3) = 1$. After the update with P, we obtain $v < s$ since as depicted in Figure 7, $f(e_1) + f(e_2) > f(e_3)$.

6 Justifiable Beliefs

Now that we have introduced a number of epistemic and doxastic notions and concepts such as evidence and arguments, we will study the link between them. In particular this ties in with philosophical debates about justifiable beliefs supported by arguments.

As our language to talk about justification models we work with an extension of the setting introduced in [18]. [7] This extension is necessary to capture the main interesting features of justification models.

Syntax. Formally, we build up the language $\mathcal{L}_{JB}$ as follows:

Definition 10 *Let Φ be a set of propositional atoms, we define the language in BNF format*

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \psi \mid K\varphi \mid K_D\varphi \mid sound \mid \forall^{ev}\varphi \mid [\preceq]\varphi$$

The language comes equipped with the standard Boolean operators of negation and conjuction and a number of modalities which come with the following intended interpretation: $K\varphi$ expresses that the agent knows that φ is the case. $K_D\varphi$ expresses that the agent defeasibly knows that φ is the case. The expression *sound* captures that the current argument F is sound (i.e. true) at the actual state s, i.e. the current pieces of evidence $e \in F$ are true. We use $\forall^{ev}\varphi$ to capture the expression that for every argument F, φ is the case. The dynamic construct $[\preceq]\varphi$ captures that for every argument F' at least as convincing as the current argument F, φ is the case.

Given the language $\mathcal{L}_{JB}$, we introduce the following abbreviations: $B\varphi := K\neg K_D\neg K_D\varphi$ is read as 'the implicit agent believes that φ'. In total justification models, this reduces to $B\varphi := \neg K_D\neg K_D\varphi$

In addition, we introduce the following abbreviations:

$$Supp\,\varphi := K(sound \to \varphi)$$

Here $Supp\,\varphi$ captures the fact that the current argument F supports φ.

$$Just\,\varphi := [\preceq]supp\,\varphi$$

The construct $Just\,\varphi$ expresses that the current argument is a justification for φ.

$$\boxplus\varphi := \exists^{ev} supp\,\varphi$$

Here $\boxplus\varphi$ captures the fact that there exists an argument in support of φ.

Semantics. The formulas of $\mathcal{L}_{JB}$ are interpreted at a state s and a body of evidence F such that F is the current argument. Given a justification model

[7] Note that the axiom system of van Benthem and Pacuit [18] holds for general justification models. In particular the reduction axioms in their dynamic account are sound in the class of all justification models.

$\mathcal{M}$, a semantics for $\mathcal{L}_{JB}$ is build up as follows:

$$
\begin{array}{ll}
s, F \models p & \text{iff } s \in V(p) \\
s, F \models \neg\varphi & \text{iff } s, F \not\models \varphi \\
s, F \models \varphi \wedge \psi & \text{iff } (s, F \models \varphi) \wedge (s, F \models \psi) \\
s, F \models K\varphi & \text{iff } t, F \models \varphi \text{ for every } t \in S \\
s, F \models K_D\varphi & \text{iff } t, F \models \varphi \text{ for every } t \in S \text{ such that } t \leq s \\
s, F \models \mathit{sound} & \text{iff } s \in \bigcap F \\
s, F \models \forall^{ev}\varphi & \text{iff } s, F' \models \varphi \text{ for every } F' \in \mathcal{E} \\
s, F \models [\preceq]\varphi & \text{iff } \forall F'(F \preceq F' \Rightarrow s, F' \models \varphi)
\end{array}
$$

The interpretation for the classical operators is standard. Following the above intented interpretation, note that irrevocable knowledge $K\varphi$ is expressed as truth of φ in all possible worlds while the concept of defeasible concept of knowledge $K_D\varphi$ is interpreted as truth of φ in all worlds that are at least as plausible as the point of evaluation. The interpretation of *sound*, refers to the current argument F at which the evaluation takes place. Guaranteeing that the current argument F is true at a state s, is captured by the fact that s has to be contained in $\bigcap F$. The expression $\forall^{ev}\varphi$ quantifies over all arguments in the collection $\mathcal{E}$ while the semantics of $[\preceq]\varphi$ uses the preorder over bodies of evidence to express that φ holds at every equally strong or stronger argument in the preorder.

A number of interesting axioms for the logic $\mathcal{L}_{JB}$ over the class of total justification models can be shown to be sound. [8] Besides the standard axioms of the logic for irrevocable and defeasible knowledge [8], this includes:

Necessitation Rules for both $\forall^{ev}$ and $[\preceq]$

$S5$-axioms for $\forall^{ev}$

$S4$-axioms for $[\preceq]$

$[\preceq]K\varphi \rightarrow K[\preceq]\varphi$

$\forall^{ev}K\varphi \rightarrow K\forall^{ev}\varphi$

$\forall^{ev}K_D\varphi \rightarrow K_D\forall^{ev}\varphi$

$\forall^{ev}\varphi \rightarrow [\preceq]\varphi$

In the context of total justification models, we add a *totality axiom* for bodies of evidence (arguments): $\forall^{ev}(\varphi \vee [\preceq]\psi) \wedge \forall^{ev}(\psi \vee [\preceq]\varphi) \rightarrow \forall^{ev}\varphi \vee \forall^{ev}\psi$

[8] Note that in the specific context of the logical system of [18], the axioms of van Benthem and Pacuit hold for general justification models. In particular also the reduction axioms in the dynamic approach of van Benthem and Pacuit are sound in the class of all justification models.

Justifiable beliefs. In the above logical setting we can encoded a number of important statements as follows:

Proposition 3. An agent believes Q iff every argument can be strengthened to a justification for Q, i.e.

$$\forall F \exists F' \succeq F (\forall F'' \succeq F' (\bigcap F'' \subseteq Q))$$

This fact can be captured by the following validity: $Bp \iff \forall^{ev} \langle \preceq \rangle just\, p$. Or more explicitly as: $Bp \iff \forall^{ev} \langle \preceq \rangle [\preceq] supp\, p$

To prove this proposition we first state and prove the following Lemma:

Lemma 1. An agent believes Q iff all maximal (in the sense of strength order) arguments supports Q, i.e. $\forall F \in Max_{\preceq}\mathcal{E}(\bigcap F \subseteq Q)$ where $Max_{\preceq}\mathcal{E} = \{F \in \mathcal{E} \mid F \not\prec F' \text{ for any } F' \in \mathcal{E}\}$.

Proof

- In the direction from left to right, we start from a given justification model $\mathcal{M}$ in which BQ is true at s. So we know that $bestS \subseteq Q$. Let $F \in Max_{\preceq}\mathcal{E}$ and $t \in \bigcap F$. Then $F \subseteq E_t$, so $F \preceq E_t$. Suppose $t \notin bestS$. Then $\exists w <_E t$, so $E_t \prec E_w$, so $F \prec E_w$. This contradicts $F \in Max_{\preceq}\mathcal{E}$. Then $t \in bestS$, so $\bigcap F \subseteq bestS$. Hence $\bigcap F \subseteq Q$.
- In the direction from right to left we assume as given a justification model $\mathcal{M}$ and a state s such that $\forall F \in Max_{\preceq}\mathcal{E}(\bigcap F \subseteq Q)$. Let $t \in bestS$. Suppose $E_t \notin Max_{\preceq}\mathcal{E}$. Then $\exists F' \in \mathcal{E}$ such that $E_t \prec F'$. Let $w \in \bigcap F'$, so $F' \subseteq E_w$, so $F' \preceq E_w$. Then $E_t \prec E_w$, so $w <_E t$. This contradicts $t \in bestS$. Hence, $E_t \in Max_{\preceq}\mathcal{E}$. Then, $t \in \bigcap E_t \subseteq Q$. Hence $t \in Q$, so $bestS \subseteq Q$. Hence, BQ is true at s.

Now we can prove Proposition 3.

Proof

- In the direction from left to right, we start from a given justification model $\mathcal{M}$ in which BQ is true at s. Let $F \in \mathcal{E}$. Then F can be strengthened to a maximal argument F', i.e. $\exists F' \succeq F(F' \in Max_{\preceq}\mathcal{E})$. Indeed since S is finite, so is $\mathcal{E}$. By Lemma 1, since BQ is true at s and $F' \in Max_{\preceq}\mathcal{E}$, $\bigcap F' \subseteq Q$. So F' supports Q. Let $F'' \succeq F'$. Then $F'' \in Max_{\preceq}\mathcal{E}$ and by Lemma 1, $\bigcap F'' \subseteq Q$. So F'' supports Q. Then, F' is a justification for Q. Hence, F can be strengthened to a justification for Q.
- In the direction from right to left we assume as given a justification model $\mathcal{M}$ and a state s such that $\forall F \exists F' \succeq F(\forall F'' \succeq F'(\bigcap F'' \subseteq Q))$. Let $F \in Max_{\preceq}\mathcal{E}$. Then $F \succeq F'$. Take $F'' := F$. Hence, $\bigcap F \subseteq Q$. By Lemma 1, BQ is true at s.

Proposition 4. An agent believes Q conditional on P iff every argument consistent with P can be strengthened to a justification for Q given P, i.e.

$$\forall F(\bigcap F \cap P \neq \emptyset \Rightarrow \exists F' \succeq F(\bigcap F' \cap P \notin \emptyset \wedge \forall F'' \succeq F'(\bigcap F'' \cap P \subseteq Q)))$$

To prove Proposition 4, we first state and prove the following Lemma

Lemma 2. An agent believes Q conditional on P iff all maximal (in the sense of strength order) arguments consistent with P supports Q conditional on P, i.e. $\forall F \in \mathcal{E}(F \in Max^P_{\preceq}\mathcal{E} \Rightarrow \bigcap F \cap P \subseteq Q)$ where $Max^P_{\preceq}\mathcal{E} = \{F \in \mathcal{E} \mid \bigcap F \cap P \neq \emptyset$ and $F \not\prec F'$ for any $F' \in \mathcal{E}(\bigcap F' \cap P \neq \emptyset)\}$.

Proof

- In the direction from left to right, we start from a given justification model $\mathcal{M}$ in which $B^P Q$ is true at s. So we know that $bestP \subseteq Q$. Let $F \in Max^P_{\preceq}\mathcal{E}$ and $t \in \bigcap F \cap P$. Then $F \subseteq E_t$, so $F \preceq E_t$. Suppose $t \notin bestP$. Then $\exists w <_E t$, so $E_t \prec E_w$, so $F \prec E_w$. This contradicts $F \in Max^P_{\preceq}$. Then $t \in bestP$. So we proved that $\forall t(t \in \bigcap F \cap P \Rightarrow t \in bestP)$. Hence, $\bigcap F \cap P \subseteq bestP$, i.e. $\bigcap F \cap P \subseteq Q$.
- In the direction from right to left we assume as given a justification model $\mathcal{M}$ and a state s such that $\forall F \in \mathcal{E}(F \in Max^P_{\preceq}\mathcal{E} \Rightarrow \bigcap F \cap P \subseteq Q)$. Let $t \in bestP$. Then $\bigcap E_t \cap P \neq \emptyset$. Suppose $E_t \notin Max^P_{\preceq}\mathcal{E}$. Then $\exists F' \in \mathcal{E}$ such that $(\bigcap F' \cap P \neq \emptyset)$ and $E_t \prec F'$. Let $w \in \bigcap F' \cap P$, so $F' \subseteq E_w$, so $F' \preceq E_w$. Then $E_t \prec E_w$, so $w <_E t$. This contradicts $t \in bestP$. Hence, $E_t \in Max^P_{\preceq}\mathcal{E}$. Then, $t \in \bigcap E_t \cap P \subseteq Q$. Hence $t \in Q$. So we proved that $\forall t(t \in bestP \Rightarrow t \in Q)$. Hence, $bestP \subseteq Q$, i.e. $B^P Q$ is true at s.

Now we can prove Proposition 4.

Proof.

- In the direction from left to right, we start from a given justification model $\mathcal{M}$ in which $B^P Q$ is true at s. Let $F \in \mathcal{E}$ such that F is consistent with P, i.e. $\bigcap F \cap P \neq \emptyset$. Then F can be strengthened to a maximal argument F' consistent with P, i.e. $\exists F' \succeq F(F' \in Max^P_{\preceq}\mathcal{E})$. Indeed since S is finite, so is $\mathcal{E}$. By Lemma 2, since $B^P Q$ is true at s and $F' \in Max^P_{\preceq}\mathcal{E}$, $\bigcap F' \cap P \subseteq Q$. So F' supports Q conditional on P. Let $F'' \succeq F'$. Then $F'' \in Max^P_{\preceq}\mathcal{E}$ and by Lemma 2, $\bigcap F'' \cap P \subseteq Q$. So F'' supports Q conditional on P. Then, F' is a justification for Q given P. Hence, F can be strengthened to a justification for Q given P.
- In the direction from right to left we assume as given a justification model $\mathcal{M}$ and a state s such that $\forall F(\bigcap F \cap P \neq \emptyset \Rightarrow \exists F' \succeq F(\bigcap F' \cap P \notin \emptyset \wedge \forall F'' \succeq F'(\bigcap F'' \cap P \subseteq Q)))$. Let $F \in Max^P_{\preceq}\mathcal{E}$. Then $F \succeq F'$. Take $F'' := F$. Hence, $\bigcap F \cap P \subseteq Q$. By Lemma 2, $B^P Q$ is true at s.

In the remainder of this section we restrict ourselves to justification models with a total preorder. In such total justification models, every belief is a justified belief.

Proposition 5. In total justification models, an agent believes Q iff there exists a justification F for Q, i.e. $\exists F \forall F' \succeq F(\bigcap F' \subseteq Q)$
This fact can be captured by the following validity:

$$Bp \iff \exists^{ev} just\, p$$

Or writing it more explicitly: $Bp \iff \exists^{ev}[\preceq] supp\, p$

Proof.

- In the direction from left to right, we start from a given total justification model $\mathcal{M}$ in which BQ is true at s. By Proposition 3, every argument can be strengthened to a justification for Q. Take any argument and strengthen it, then we have a justification for Q.
- In the direction from right to left we assume as given a total justification model $\mathcal{M}$ and a state s such that there exists a justification F for Q. We have to show that $s \models BQ$. Since F is a justification for Q, by Definition 6, $\forall F' \in \mathcal{E}(F \preceq F' \Rightarrow \bigcap F' \subseteq Q)$. Take any argument G such that $F \preceq G$ or $G \preceq F$. If $F \preceq G$, $\bigcap G \subseteq Q$. If $G \preceq F$, G can be strengthened to a justification for Q since $G \preceq F$ and $\forall F' \in \mathcal{E}$ such that $F \preceq F', \bigcap F' \subseteq Q$. By Proposition 3, BQ is true at s.

Proposition 6. In total justification models, an agent defeasibly knows Q at s iff there exists a sound (true) justification F for Q at s, i.e.

$$\exists F(s \in \bigcap F \wedge \forall F' \succeq F(\bigcap F' \subseteq Q))$$

This fact can be captured by the following validity: $K_D p \iff \exists^{ev}(sound \wedge just\, p)$

Proof.

- In the direction from left to right, we start from a given total justification model $\mathcal{M}$ in which $K_D Q$ is true at s. Then $\forall t(t \leq_E s \Rightarrow t \in Q)$. So, $\forall t(E_s \preceq E_t \Rightarrow t \in Q)$. Take $F := E_s$. Since $s \in \bigcap E_s$, $s \in \bigcap F$. Let $F' \succeq F$ and $t \in \bigcap F'$. Then $F' \subseteq E_t$, so $F' \preceq E_t$. Since $E_s \preceq F'$, $E_s \preceq E_t$. Then $t \leq_E s$, so $t \in Q$. Hence, $\exists F(s \in \bigcap F \wedge \forall F' \succeq F(\bigcap F' \subseteq Q))$.
- In the direction from right to left we assume as given a total justification model $\mathcal{M}$ and a state s such that $\exists F(s \in \bigcap F \wedge \forall F' \succeq F(\bigcap F' \subseteq Q))$. Let $t \leq_E s$. We want to show that $t \in Q$. As we know, $s \in \bigcap F$. Then $F \subseteq E_s$, so $F \preceq E_s$. Since $t \leq_E s$, $E_s \preceq E_t$. Then $F \preceq E_t$. By assumption, $\bigcap E_t \subseteq Q$. So, $t \in Q$. Hence, $K_D Q$ is true at s.

Proposition 7. An agent believes Q conditional on P iff there exists a justification F for Q given P, i.e. $\exists F(\bigcap F \cap P \neq \emptyset \wedge \forall F' \succeq F(\bigcap F' \cap P \subseteq Q))$

Proof.

- In the direction from left to right, we start from a given total justification model $\mathcal{M}$ in which $B^P Q$ is true at s. By Proposition 4, every argument consistent with P can be strengthened to a justification for Q given P. Take any argument consistent with P and strengthen it, then we have a justification for Q given P.
- In the direction from right to left we assume as given a total justification model $\mathcal{M}$ and a state s such that there exists a justification F for Q given P. We have to show that $s \models B^P Q$. Since F is a justification for Q given P, by

Definition 8, $\bigcap F \cap P \neq \emptyset$ and $\forall F' \in \mathcal{E}(F \preceq F' \Rightarrow \bigcap F' \cap P \subseteq Q)$. Take any argument G consistent with P ($\bigcap G \cap P \neq \emptyset$) such that $F \preceq G$ or $G \preceq F$. If $F \preceq G$, $\bigcap G \cap P \subseteq Q$. If $G \preceq F$, G can be strengthened to a justification for Q given P since $G \preceq F$ and $\forall F' \in \mathcal{E}$ such that $F \preceq F', \bigcap F' \cap P \subseteq Q$. By Proposition 4, $B^P Q$ is true at s.

Conclusion

We have provided a general setting of justification models which subsumes plausibility models, counting and weighting models as well as evidence models. With this line of work we enhance the investigations into different types of models that connect evidence and beliefs, relating it to the investigations in neighborhood structures in [18,16,17] and our recent work on topological models in [2,5,3,4]. From a logic-technical point of view, we have introduced a number of sound axioms for the logic of justifiable beliefs. In this context we were are able to express beliefs supported by a justification and knowledge supported by an agent's sound justification. Similar concepts are investigated in the context of topological models in [4] in which a complete axiom system is provided. An initial investigation of the relation between beliefs and arguments as studied in formal argumentation theory, ties in with the present research and is further explored in [13].

We have studied in this paper the dynamics of justification models under updates with new truthful information and indicated an interesting problem that we encountered in the context of updating counting models. There are of course a number of other ways in which a given justification model can be transformed into a new one. One can for instance add a new body of evidence to a given structure and give it a degree of plausibility in relation to the other bodies of evidence. This would correspond to the action of adding soft evidence of which the agent is not fully certain. We have left the theory of soft evidence upgrades for future work.

From a philosophical point of view, one can argue that the introduced setting in this paper is too restrictive as it can only deal with beliefs supported by genuine evidence and leaves no room for an agent's biases (or defaults). In a number of applications or scenarios, one may wantt to consider cases in which an agent may have some preferences that are not genuinely based on (external) evidence but based on the trustworthiness of her (internal) senses and reasons. To model such scenarios it would be necessary to introduce a refined type of justification model that contains evidence sets consisting of two types: genuine evidence and biases. In refined justification models, the definition of E can be given by:

Definition 11 *$E = E_0 \cup B$ such that E_0 is the family of evidence sets representing the genuine evidence the agent has while B is the family of evidence sets representing the prior biases of the agent.*

In such a refined setting it is natural to impose a condition on B that captures the fact that all biases $b \in B$ should strictly increase the strengh of a body of evidence. Formally, $F < F \cup \{b\}$ such that $F \cup \{b\} \in \mathcal{E}$. It is then interesting to note that in a refined justification model, we can consider a "softer" kind of support: an argument F *weakly supports* Q (or F is a *"soft" argument for* Q) conditional on some set $B' \subseteq B$ of biases iff $F \cup B'$ supports Q.

The (refined) justification models provide the technical underpinning that is needed to give a formal account of K. Leher's justification games, the first steps of which are provided [9].

References

[1] Alchourrón, C. E., P. Gärdenfors and D. Makinson, *On the logic of theory change: Partial meet contraction and revision functions*, J. Symb. Log. **50** (1985), pp. 510–530.

[2] Baltag, A., N. Bezhanishvili, A. Ozgun and S. Smets, *The topology of belief, belief revision and defeasible knowledge*, in: H. Huang, D. Grossi and O. Roy, editors, *Logic, Rationality and Interaction, Proceedings of the Fourth International Workshop, LORI 2013*.

[3] Baltag, A., N. Bezhanishvili, A. Ozgun and S. Smets, *The topological theory of belief*, Technical report, ILLC Pre-publication PP-2015-18. (2015).

[4] Baltag, A., N. Bezhanishvili, A. Ozgun and S. Smets, *Justified belief and the topology of evidence*, in: *WOLLIC 2016 proceedings, Springer*, 2016.

[5] Baltag, A., N. Bezhanishvili, A. Ozgun and S. Smets, *The topology of full and weak belief.*, in: *Forthcoming in the Proceedings of the 11th International Tbilisi Symposium on Language, Logic and Computation in Lecture Notes in Computer Science*, 2016.

[6] Baltag, A., L. Moss and S. Solecki, *The logic of public announcements, common knowledge, and private suspicions*, in: *Proceedings of the 7th conference on Theoretical aspects of rationality and knowledge*, Morgan Kaufmann Publishers Inc., 1998, pp. 43–56.

[7] Baltag, A. and S. Smets, *Dynamic belief revision over multi-agent plausibility models*, in: *Proceedings of the 7th Conference on Logic and the Foundations of Game and Decision (LOFT 2006), University of Liverpool.*, 2006.

[8] Baltag, A. and S. Smets, *A qualitative theory of dynamic interactive belief revision*, Logic and the Foundations of Game and Decision Theory (LOFT7) (2008), pp. 11–58.

[9] Fiutek, V., "Playing with Knowledge and Belief," ILLC Dissertation Series DS-2013-02 University of Amsterdam, 2013.

[10] Grove, A., *Two modellings for theory change*, Journal of Philosophical Logic **17** (1988), pp. 157–170.

[11] Lehrer, K., "Theory of Knowledge," Routledge, 1990.

[12] Plaza, J., *Logics of public communication*, in: *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems* (1989), pp. 201–216.

[13] Shi, C. and S. Smets, *Beliefs supported by arguments*, in: *Proceedings of the Chinese Conference on Logic and Argumentation (CLAR 2016), CEUR proceedings collection*, 2016.

[14] van Benthem, J., *Dynamic logic for belief revision*, Journal of Applied Non-Classical Logics **12(2)** (2007), pp. 129–155.

[15] van Benthem, J., "Logical dynamics of information and interaction," Cambridge University Press, 2011.

[16] van Benthem, J., D. F. Duque and E. Pacuit, *Evidence logic: A new look at neighborhood structures*, in: *Advances in Modal Logic 9, papers from the ninth conference on "Advances in Modal Logic,"*, 2012, pp. 97–118.

[17] van Benthem, J., D. Fernández-Duque and E. Pacuit, *Evidence and plausibility in neighborhood structures*, Annals of Pure and Applied Logic **165** (2014), pp. 106 – 133.

[18] van Benthem, J. and E. Pacuit, *Dynamic logics of evidence-based beliefs*, Studia Logica **99(1)** (2011), pp. 61–92.

[19] van Benthem, J. and S. Smets, *Dynamic logics of belief change.*, in: *Handbook of Logics for Knowledge and Belief*, College Publications, 2015 .