Some logics of belief and disbelief

Samir Chopra
schopra@cse.unsw.edu.au
Knowledge Systems Group
School of Computer Science and Engineering
University of New South Wales

Johannes Heidema
heidej@unisa.ac.za
Department of Mathematics
University of South Africa
Pretoria 0003,South Africa

Thomas Meyer
tmeyer@cs.up.ac.za
Department of Computer Science
University of Pretoria
Pretoria 0002,South Africa

Abstract

The introduction of explicit notions of rejection, or disbelief, into logics for knowledge representation can be justified in a number of ways. Motivations range from the need for versions of negation weaker than classical negation, to the explicit recording of classic belief contraction operations in the area of belief change, and the additional levels of expressivity obtained from an extended version of belief change which includes disbelief contraction. In this paper we present four logics of disbelief which address some or all of these intuitions. Soundness and completeness results are supplied and the logics are compared with respect to applicability and utility.

1 Introduction

The introduction of explicit notions of *belief rejection* or *disbelief* into logics for knowledge representation can be justified in a number of ways. One such justification is found in the research area of *belief change* [1, 3]. Classic belief change recognises two basic operations on the beliefs of an agent: *revision*, where the aim is to incorporate a new belief into an agent's belief state while still maintaining consistency, and *contraction*, where an agent needs to discard one of its beliefs while, at the same time, trying to retain as many of its remaining beliefs as possible. The argument for introducing explicit disbeliefs into the object language of choice is based on the conviction that belief contraction should be seen as a belief change operation on par with revision, and not just as an intermediate step for performing revision, as it is sometimes presented. The following example, based on an example in [9], but originally found in [4], illustrates the point.[1]

[1] Observe that this is an example of *base change* [2, 7, 11].

Example 1 *Consider the situation where a totally ignorant agent revises its current (empty) belief base with $p \rightarrow q$, then contracts with q, and then revises with p. In a framework which caters for the explicit representation of beliefs only, the principle of informational economy dictates that there is only one feasible result for this sequence of base change operations: the set $\{p \rightarrow q, p\}$. The argument against the claim that this is the only acceptable outcome is that, because the representational framework does not allow for the explicit recording of the contraction of q, the choice of permissible outcomes is skewed in favour of revision operations. Contrast this with a situation in which disbelieving a sentence ϕ can be represented in the object language as a sentence of the form $\overline{\phi}$. In this case the sets $\{p \rightarrow q, \overline{q}\}$ and $\{\overline{q}, p\}$ will be acceptable outcomes as well.*

In the example above, classical belief contraction is equated with disbelief *revision*. But the extended version of belief change in which it is also possible to perform contraction with *disbeliefs* has no equivalent in classical belief change. The contraction of disbeliefs provides for a particularly interesting class of operations if disbeliefs are thought of as *guards* against the introduction of certain beliefs, in the spirit of default logic [10]. Disbelief *contraction* corresponds to the removal of such a guard, which might then trigger the addition of previously suppressed beliefs into the current belief set of an agent. An adequate modelling of this type of extended (dis)belief change requires some level of interaction between beliefs and disbeliefs, something which has not been dealt with adequately in the literature. It also presupposes a consequence relation which handles disbeliefs appropriately, since it allows for contraction with disbeliefs which are *consequences* of previously asserted disbeliefs.

Another motivation for the introduction of disbeliefs is that it allows us to reason about rejected beliefs and their consequences [12, 5, 6]. In this scenario the argument is that in many practical situations it makes more sense to adopt a "negative stance": to reason about how to reject new beliefs on the basis of rejecting old ones. Under these circumstances the focus is on the provision of an appropriate consequence relation with regard to disbeliefs.

Yet another motivation for an explicit representation of disbeliefs is that it can be viewed as a weaker version of classical negation. The value of such an additional version of negation is that it enables us to resolve well-known paradoxes of belief such as the lottery paradox.

Example 2 *Given a finite number of lottery tickets issued in a particular week, I express doubt about the possibility of any individual ticket winning the lottery, although I do believe that exactly one of these tickets will win the lottery.*

The lottery paradox would be resolved by replacing classical negation (I believe the claim that ticket number 113 will not win the lottery) with the much weaker notion of disbelief (I disbelieve the claim that ticket number 113 will win the lottery). This should not be seen as an *ad hoc* response to a particular paradox of belief. To the contrary, such a weaker version of negation is prevalent in many practical situations. Here is another example.

Example 3 *I have two very good friends, Agnetha and Björn, who are the only suspects in the murder of my friends Annifrid and Benny. I refuse to believe that Agnetha was able to commit the murder (represented as the disbelief $\overline{a}$), and similarly for Bjorn (represented as the disbelief $\overline{b}$). Yet, the evidence compels me to believe that one of them committed the crime (represented as $a \vee b$).*

The challenge is to provide a formal definition of disbelief in such a way that sentences of the form $\overline{\phi}$ can be regarded as a legitimate way of negating ϕ, while at the same time ensuring the consistency of sets like $\{\overline{\phi}, \overline{\psi}, \phi \vee \psi\}$.

In this paper we present four different logics of belief and disbelief. Two of these logics, WBD and GBD, have been used implicitly in work on disbelief, but not in the form of a *logic* of belief and disbelief. The first one, WBD, has a very weak notion of consequence with regard to disbeliefs. It has been used for the explicit recording of belief contraction operations in [9] and [4]. This logic provides an extremely weak link between beliefs and disbeliefs. The relation between beliefs and disbeliefs is made only via a notion of consistency, and this is done for the sole purpose of providing an accurate model of classical belief contraction in terms of disbeliefs.

The second logic, GBD, was presented in [5, 6], albeit in a different form, and is a formalisation of the idea that it is useful to reason about rejected beliefs and their consequences. The notion of consequence, with regard to disbeliefs, associated with GBD is stronger than that of WBD. We argue in section 4 that it is *too* strong. On the other hand, the connection between beliefs and disbeliefs in GBD is just as weak as in WBD. We argue that it needs to be strengthened.

The third logic, BD, is designed with the express intention of obtaining a weaker version of classical negation. Its version of consequence with respect to disbeliefs seems to be pitched at just the right level. It is weaker than the version provided by GBD, but stronger than that of WBD. And unlike the cases for WBD and GBD, there is a strong and intuitively plausible connection between beliefs and disbeliefs. Indeed, one of the consequences of the belief in a sentence $\neg\phi$ is that ϕ will be disbelieved (that is, $\overline{\phi}$ will hold), as befits a view of disbelief as a weaker version of classical negation. It is our contention that BD is the most useful of the logics introduced in this paper.

With the introduction of BN, the fourth and final logic to be discussed, the aim is to take a matters a step further by obtaining new beliefs from existing disbeliefs as well. In the process of doing so, however, disbelief collapses into classical negation. That is, disbelieving $\overline{\phi}$ becomes equivalent to believing $\neg\phi$, and BN becomes equivalent to classical logic in its level of expressivity.

In section 2 we present the language on which all of our logics are based. In each of sections 3 to 6 we present and discuss a logic of belief and disbelief. Each logic is equipped with a proof-theoretic and a semantic version of consequence. It is shown that these logics are all sound and complete. In section 7 we draw conclusions and point to future research.

1.1 Formal preliminaries

We assume a classical propositional logic PL with a language L_{PL} generated from a (possibly countably infinite) set of propositional atoms, together with the the usual propositional connectives, and with $\perp$ and $\top$ as canonical representatives of the set of contradictions and tautologies, respectively. V is the set of classical valuations of PL. For $\phi \in L_{PL}$, $M(\phi)$ is the set of models of ϕ. Classical deduction for PL is denoted by $\vdash_{PL}$. Given a language L and a consequence relation $\vdash_X$ from $\mathcal{P}L$ to L we let $C_X(\Gamma) = \{\alpha \mid \Gamma \vdash_X \alpha\}$ be the consequence operation associated with $\vdash_X$. For example, C_{PL} denotes classical consequence for propositional logic. It is well-known that C_{PL} satisfies the properties of Inclusion: $X \subseteq C(X)$, Idempotency: $C(X) = C(C(X))$, Monotonicity: $X \subseteq Y$ implies $C(X) \subseteq C(Y)$, and Compactness: if $\phi \in C(X)$ then $\phi \in C(Y)$ for some finite subset Y of X.[2] For any set X we denote the powerset of X (the set of all subsets of X) by $\mathcal{P}X$. An operation C from $\mathcal{P}L$ to $\mathcal{P}L$ is a *Tarskian con-*

[2] Instead of taking PL as our "basic logic" we can work with any logic whose associated consequence relation satisfies these four properties.

sequence operation iff it satisfies Inclusion, Idempotency and Monotonicity.

2 A language for beliefs and disbeliefs

The language L on which we base all of the logics to be presented is a simple extension of L_{PL}. For each potential belief ϕ, expressed as a sentence of L_{PL}, there is a corresponding potential disbelief $\overline{\phi}$.

Definition 1 *$L = L_B \cup L_D$, where $L_B = L_{PL}$ is the set of potential beliefs and $L_D = \{\overline{\phi} \mid \phi \in L_P\}$, the set of potential disbeliefs. An* information set *Γ is any subset of L. The* beliefs *in Γ are defined as $\Gamma_B = L_B \cap \Gamma$. The* disbeliefs *in Γ are defined as $\Gamma_D = L_D \cap \Gamma$.*

We use ϕ and ψ to denote potential beliefs and α and β to denote arbitrary sentences of L (potential beliefs and disbeliefs). The language L is fairly restrictive in the sense that $\overline{}$ is not viewed as a propositional connective. Thus, we cannot construct sentences of the form $p \vee \overline{q}$, for example. The reason for this is twofold. Firstly, in many (but not all) motivations for the introduction of disbeliefs, such a level of expressivity is simply not necessary. And secondly, this paper should be seen as a first step towards a description of logics of belief and disbelief. Our intention is to treat $\overline{}$ as a full-blown "propositional" connective in future research.

3 The logic WBD

The first logic we consider is relatively weak in terms of the consequences of explicit disbeliefs. The motivation for the introduction of WBD is primarily for the explicit expression of classical belief contraction. It was used as such in [4] and [9], albeit implicitly. A proof theory for WBD is obtained from the following three inference rules.

(B) If $\Gamma_B \vdash_{PL} \phi$ then $\Gamma \vdash \phi$

(D$\perp$) If $\phi \vdash_{PL} \perp$ then $\Gamma \vdash \overline{\phi}$

(WD) If $\overline{\psi} \in \Gamma_D$ and $\phi \vdash_{PL} \psi$ then $\Gamma \vdash \overline{\phi}$

(B) is a supraclassicality requirement. (D$\perp$) requires that contradictions always be disbelieved, analogous to the case of tautologies always being believed. (WD) requires that disbelieving a sentence $\overline{\psi}$ leads to a disbelief in all those sentences classically stronger than ψ.

Definition 2 *In WBD, α can be deduced from an information set Γ, written as $\Gamma \vdash_{\mathrm{WBD}} \alpha$, iff (Γ, α) is in the smallest binary relation from $\mathcal{P}L$ to L closed under (B), (D$\perp$), and (WD).*

The logic WBD has a well-behaved Tarskian consequence operation.

Proposition 1 *C_{WBD} satisfies Inclusion, Idempotency, Monotonicity and Compactness.*

Observe that beliefs and disbeliefs are completely decoupled in WBD. In particular, WBD satisfies the following two properties which show that disbeliefs and beliefs are only derivable from the set of disbeliefs and beliefs respectively:

(B$\not\to$D) $\forall \Omega \subseteq L_B$ and $\phi \in L_B$, $\Gamma \vdash \overline{\phi}$ iff $\Gamma_D \cup \Omega \vdash \overline{\phi}$

(D$\not\to$B) $\forall \Delta \subseteq L_D$ and $\phi \in L_B$, $\Gamma \vdash \phi$ iff $\Gamma_B \cup \Delta \vdash \phi$

Proposition 2 *$\vdash_{\mathrm{WBD}}$ satisfies (B$\not\to$D) and (D$\not\to$B).*

In fact, the only way in which beliefs and disbeliefs are related in WBD is by way of the notion of *WBD-inconsistency*.

Definition 3 *Γ is* WB-inconsistent *iff $\Gamma_B \vdash_{\mathrm{WBD}} \perp$. Γ is* WD-inconsistent *iff $\Gamma_D \vdash_{\mathrm{WBD}} \overline{\top}$. Γ is* WBD-inconsistent *iff $\Gamma \vdash_{\mathrm{WBD}} \phi$ and $\Gamma \vdash_{\mathrm{WBD}} \overline{\phi}$ for some $\phi \in L_B$.*

We have similar versions of inconsistency for the other logics too.

We now turn to the semantics for WBD. A useful intuition is to think of the beliefs in Γ as the information an agent has acquired on its own, and each disbelief $\overline{\phi}$ as information it has obtained from a different source, informing it that ϕ does *not* hold. The agent has more faith in its own capabilities than in those of its sources, and information obtained from its sources is therefore seen as less reliable. The information obtained from a specific source is independent of the beliefs of the agent and the information obtained from other sources. A model for WBD consists of a set of valuations, corresponding to the worlds that the agent regards as possible, together with a set of sets of valuations, with each element of this set corresponding to the worlds that a particular source of the agent regards as *impossible*. A potential belief is satisfied in a model if it is true in all the worlds that the agents regards as possible, and a potential disbelief $\overline{\phi}$ is satisfied in a model if at least one of the sources regards ϕ as impossible. Satisfaction is denoted by $\Vdash$.

Definition 4 *A* WBD-model *$\mathcal{M}$ is an ordered pair $(M, \mathcal{N})$ where $M \subseteq V$ and $\emptyset \subset \mathcal{N} \subseteq \mathcal{P}V$. For $\phi \in L_B$, $\mathcal{M} \Vdash \phi$ iff $M \subseteq M(\phi)$. For $\overline{\phi} \in L_D$, $\mathcal{M} \Vdash \overline{\phi}$ iff $\exists N \in \mathcal{N}$ such that $N \subseteq M(\neg\phi)$.*

In a WBD-model $\mathcal{M} = (M, \mathcal{N})$, M represents the models of the beliefs of the agent, and each $N \in \mathcal{N}$ represents the models of a sentence that the source associated with N

holds to be impossible. We require an agent to have at least one source of information. That is, we require that $\mathcal{N} \neq \emptyset$. Entailment for WBD (denoted by $\models_{\text{WBD}}$) is then defined in the normal model-theoretic fashion.

Definition 5 *$\mathcal{M} \Vdash \Gamma$ iff $\mathcal{M} \Vdash \alpha\ \forall\alpha \in \Gamma$. $\Gamma \models_{WBD} \alpha$ iff $\mathcal{M} \Vdash \alpha$ for every WBD-model $\mathcal{M}$ s.t. $\mathcal{M} \Vdash \Gamma$.*

It turns out that the logic WBD is sound and complete.

Theorem 1 *$\Gamma \vdash_{\text{WBD}} \alpha$ iff $\Gamma \models_{\text{WBD}} \alpha$, for all $\alpha \in L$.*

Based on the semantics of WBD we can also define appropriate notions of (un)satisfiability which, via theorem 1, can be shown to coincide with inconsistency. We postpone presentation of these details to the full version of this paper.

The logic WBD is, essentially, the logic on which the work in [4] and [9] is based. Both these papers argue for an explicit expression of disbeliefs in order to maintain a record of belief contraction. Although WBD seems adequate for this purpose, at least in simple settings, two types of criticisms can be levelled at it: the extreme weakness of the notion of consequence associated with disbeliefs, and the complete decoupling of beliefs and disbeliefs. These weaknesses become apparent in scenarios where it is appropriate to perform various belief change operations on beliefs as well as disbeliefs. So, while contracting with a belief might correspond to revising with a disbelief, contracting with a *disbelief* has no equivalent in classical belief change. And such an operation is useful if disbeliefs are thought of as *guards* against the introduction of certain information, in the spirit of default logic [10]. The removal of such a guard might then trigger the addition of beliefs into an agent's current belief set. An adequate modelling of this type of extended (dis)belief change requires some level of interaction between beliefs and disbeliefs. It also presupposes a consequence relation which handles disbeliefs appropriately, since it allows for contraction with disbeliefs which are *consequences* of previously asserted disbeliefs. In section 4 we consider an attempt to rectify the weakness of consequences derived from disbeliefs. In section 5 we address this issue as well, and we also consider the the provision of a link between beliefs and disbeliefs.

4 The logic GBD

One way in which to address the weakness of the consequences to be derived from disbeliefs is to regard consequence, with respect to disbeliefs, as the exact dual of consequence with respect to beliefs. This is the approach followed in [6] where, interestingly enough, the reason for defining such a logic is to define disbelief change.

Let $\overline{\Gamma}_D = \{\neg\phi \mid \overline{\phi} \in \Gamma_D\}$ and consider the following rule:

(GD) If $\overline{\Gamma}_D \vdash_{PL} \neg\phi$ then $\Gamma \vdash \overline{\phi}$

(GD) requires that the consequences of disbeliefs be the exact dual of classical consequence. The proof theory for the logic GBD is then obtained from WBD by replacing (WD) with (GD).

Definition 6 *In GBD, α can be deduced from an information set Γ, written as $\Gamma \vdash_{\text{GBD}} \alpha$, iff (Γ, α) is in the smallest binary relation from $\mathcal{P}L$ to L closed under (B), (D$\perp$), and (GD).*

GBD also has a well-behaved Tarskian consequence operation.

Proposition 3 *C_{GBD} satisfies Inclusion, Idempotency, Monotonicity and Compactness.*

The logic GBD satisfies the following property, which was proposed in [5].

(Rej) If $\Gamma \vdash \overline{\beta}$ and $\Gamma \vdash \overline{\neg(\alpha \to \beta)}$ then $\Gamma \vdash \overline{\alpha}$

(Rej) can be thought of as a version of the inference rule Modus Tollens. In fact, if disbelief is replaced with classical negation, (Rej) coincides exactly with Modus Tollens.

GBD also has a complete decoupling of beliefs and disbeliefs, as the following proposition shows.

Proposition 4 *$\vdash_{\text{GBD}}$ satisfies (B$\not\to$D) and (D$\not\to$B).*

The three versions of inconsistency for GBD are defined as for WBD.

Definition 7 *Γ is* GB-inconsistent *iff $\Gamma_B \vdash_{\text{GBD}} \perp$. Γ is* GD-inconsistent *iff $\Gamma_D \vdash_{\text{GBD}} \overline{\top}$. Γ is* GBD-inconsistent *iff $\Gamma \vdash_{\text{GBD}} \phi$ and $\Gamma \vdash_{\text{GBD}} \overline{\phi}$ for some $\phi \in L_B$.*

Observe that beliefs and disbeliefs are related only by GBD-inconsistency.

A semantics for GBD is obtained by considering only those WBD-models in which $\mathcal{N}$ has a single element. Intuitively, all disbeliefs are obtained from a single source. This ensures that disbeliefs can be combined to obtain new disbeliefs, which allows for a stronger notion of consequence regarding disbeliefs. GBD-entailment are defined as for WBD.

Definition 8 *A* GBD-model *is an ordered pair $\mathcal{M} = (M, N)$ where $M, N \subseteq V$. For $\phi \in L_B$, $\mathcal{M} \Vdash \phi$ iff $M \subseteq M(\phi)$. For $\overline{\phi} \in L_D$, $\mathcal{M} \Vdash \overline{\phi}$ iff $N \subseteq M(\neg\phi)$. $\mathcal{M} \Vdash \Gamma$ iff $\mathcal{M} \Vdash \alpha\ \forall\alpha \in \Gamma$. $\Gamma \models_{GBD} \alpha$ iff $\mathcal{M} \Vdash \alpha$ for every GBD-model $\mathcal{M}$ s.t. $\mathcal{M} \Vdash \Gamma$.*

The logic GBD is sound and complete.

Theorem 2 $\Gamma \vdash_{\mathrm{GBD}} \alpha$ *iff* $\Gamma \vDash_{\mathrm{GBD}} \alpha$*, for all* $\alpha \in L$.

We postpone a full description of the appropriate versions of unsatisfiability for GBD, which coincide with the three versions of inconsistency for GBD, to the full version of this paper.

The crucial difference between WBD and GBD is that GBD combines disbeliefs to obtain new disbeliefs, which allows for a much stronger notion of consequence with respect to disbeliefs. In particular, GBD satisfies the following property.

(D$\vee$) If $\Gamma \vdash \overline{\phi}$ and $\Gamma \vdash \overline{\psi}$ then $\Gamma \vdash \overline{\phi \vee \psi}$

It is our contention that such a property is undesirable for a notion of disbelief. Since disbelief is *not* intended to be equivalent to classical negation, it is reasonable to require that it be possible to express notions not expressible in classical logic. One of these is *agnosticism*, in which an agent refuses to commit to a potential belief or its negation. Formally, this amounts to both $\overline{\phi}$ and $\overline{\neg\phi}$ being consequences of a consistent information set Γ. But if the consequence relation satisfies (D$\vee$), it means that the agent is forced to accept $\overline{\phi \vee \neg\phi}$, and therefore $\overline{\top}$ as well. Disbelieving a tautology amounts to GD-inconsistency. It is the analogue of classical inconsistency (believing the negation of the tautology), but with classical negation replaced by disbelief. Now, if agnosticism in the sense described above leads to inconsistency with respect to disbeliefs, it is an indication that the consequence relation for GBD is too strong to account for a proper treatment of disbeliefs. In section 5 we consider a logic which seems to be pitched at the right level in this regard.

5 The logic BD

In this section we introduce a logic in which there is interaction between beliefs and disbeliefs. Disbelief is seen as a weaker notion of classical negation. As a result, believing the negation of a sentence also results in disbelieving that sentence, although the converse relationship does not hold. Interestingly enough, this coupling of beliefs and disbeliefs comes about as a result of the introduction of the following inference rule, which looks like a simple strengthening of the consequence relation with respect to disbeliefs.

(D) If $\overline{\psi} \in \Gamma_D$ and $\Gamma_B \cup \{\phi\} \vdash_{PL} \psi$ then $\Gamma \vdash \overline{\phi}$

(D) requires that disbelieving a sentence ψ also leads to a disbelief in those sentences classically stronger than ψ, *but with respect to* Γ_B. Observe that (WD) is the special case of (D) where $\Gamma_B = \emptyset$.

Definition 9 *In BD,* α *can be deduced from an information set* Γ*, written as* $\Gamma \vdash_{\mathrm{BD}} \alpha$*, iff* (Γ, α) *is in the smallest binary relation from* $\mathcal{P}L$ *to* L *closed under (B), (D$\perp$), and (D).*

BD has a well-behaved Tarskian consequence operation as well.

Proposition 5 C_{BD} *satisfies Inclusion, Idempotency, Monotonicity and Compactness.*

The manner in which disbeliefs are obtained from beliefs in BD can be expressed by the following property which links up a belief in $\neg\phi$ to a disbelief in ϕ. Observe that neither WBD nor GBD satisfies ($B \to D$).

(B$\to$D) If $\Gamma \vdash \neg\alpha$ then $\Gamma \vdash \overline{\alpha}$

Proposition 6 $\vdash_{\mathrm{BD}}$ *satisfies (B$\to$D) and does not satisfy (B$\not\to$D).*

The logic BD does not support any connection between beliefs and disbeliefs in the converse direction, though, as the following result shows.

Proposition 7 $\vdash_{\mathrm{BD}}$ *satisfies (D$\not\to$B).*

The different notions of inconsistency for BD are defined in the same way as for WBD and GBD.

Definition 10 *An information set* Γ *is* B-inconsistent *iff* $\Gamma_B \vdash_{\mathrm{BD}} \perp$, D-inconsistent *iff* $\Gamma_D \vdash_{\mathrm{BD}} \overline{\top}$*, and* BD-inconsistent *iff* $\Gamma \vdash_{\mathrm{BD}} \phi$ *and* $\Gamma \vdash_{\mathrm{BD}} \overline{\phi}$ *for some* $\phi \in L_B$.

Because there is interaction between beliefs and disbeliefs in BD, there is also a connection between the different notions of inconsistency for BD. In particular, we have the following results.

Proposition 8 *If* Γ *is B-inconsistent then it is also BD-inconsistent, but the converse does not hold.* Γ *is BD-inconsistent iff it is D-inconsistent.*

In the logic BD, then, BD-inconsistency collapses into D-inconsistency. Given the intuition of disbelief as a weaker version of classical negation, this is a particularly desirable state of affairs. In classical logic, asserting both ϕ and $\neg\phi$ is tantamount to the assertion that $\neg\top$ is the case, while in BD, asserting both ϕ and $\overline{\phi}$ amounts to the assertion that $\overline{\top}$ is the case.

Observe that BD-inconsistency (or D-inconsistency) amounts to disbelieving the tautology, which leads to a disbelief in every sentence in L_B. So, while B-inconsistency, like classical consistency, leads to the acceptance of *every* sentence in the language, BD-inconsistency leads only to the acceptance of every *disbelief* in the language. BD-inconsistency can thus be seen as a weaker version of classical inconsistency. We regard this as a particularly attractive feature of the logic BD. The logics WBD and GBD also have similar features, but the connection between classical inconsistency and the weaker version of inconsistency, based on disbeliefs, is not as intuitively appealing.

The semantics for BD is obtained by considering only those WBD-models for which every $N \in \mathcal{N}$ is a subset of M. That is, the worlds that the sources of an agent may regard as impossible have to be worlds that the agent itself regards as possible. BD-entailment are defined as for WBD-entailment and GBD-entailment.

Definition 11 *A* BD-model *is a tuple* $(M, \mathcal{N})$ *where* $M \subseteq V$ *and* $\emptyset \subset \mathcal{N} \subseteq \mathcal{P}M$. *For a BD-model* $\mathcal{M} = (M, \mathcal{N})$ *and* $\phi \in L_B$, $\mathcal{M} \Vdash \phi$ *iff* $M \subseteq M(\phi)$. *For* $\overline{\phi} \in L_D$, $\mathcal{M} \Vdash \overline{\phi}$ *iff* $\exists N \in \mathcal{N}$ *s.t.* $N \subseteq M(\neg\phi)$. $\mathcal{M} \Vdash \Gamma$ *iff* $\mathcal{M} \Vdash \alpha\ \forall\alpha \in \Gamma$. $\Gamma \vDash_{BD} \alpha$ *iff* $\mathcal{M} \Vdash \alpha$ *for every BD-model* $\mathcal{M}$ *s.t.* $\mathcal{M} \Vdash \Gamma$.

The logic BD is sound and complete.

Theorem 3 $\Gamma \vdash_{\mathrm{BD}} \alpha$ *iff* $\Gamma \vDash_{\mathrm{BD}} \alpha$, *for all* $\alpha \in L$.

Based on the semantics of BD we can define appropriate notions of (un)satisfiability which can be shown to coincide with inconsistency. We postpone presentation of these details to the full version of this paper.

We conclude this section by pointing out that BD is able to handle examples such as the lottery paradox and its variants (cf. examples 2 and 3) in a manner that is intuitively satisfactory. It can be shown that any information set of the form $\Gamma = \{\overline{\phi_1}, \ldots, \overline{\phi_n}, \bigvee_{i=1}^{i \leq n} \phi_i\}$ is neither B-inconsistent, nor BD-inconsistent.[3] BD thus allows us, for example, to disbelieve the fact that any particular ticket will win the lottery, while still believing that exactly one of the tickets will win the lottery, without collapsing into *any* kind of inconsistency.

6 The logic BN

We have seen that BD allows for the generation of disbeliefs from beliefs. An interesting question is whether it makes sense to do the opposite; that is, to generate beliefs from disbeliefs. We consider two ways of doing so. For the first one, note that the generation of disbeliefs from beliefs in BD is achieved by the inference rule (D). Now, it is easily verified that, in the presence of (B) and (D$\perp$), (D) is equivalent to the following property:

(D') If $\Gamma \vdash \overline{\psi}$ and $\Gamma \cup \{\phi\} \vdash \psi$ then $\Gamma \vdash \overline{\phi}$

[3]By proposition 8 this means that Γ is also not D-inconsistent.

On the basis of this equivalence we consider the following property:

(B') If $\Gamma \vdash \psi$ and $\Gamma \cup \{\overline{\phi}\} \vdash \overline{\psi}$ then $\Gamma \vdash \phi$

(B') asserts that if I currently believe ψ, and if the addition of ϕ as a disbelief leads me to disbelieve ψ, then ϕ should be one of my current beliefs. It is analogous to (D'), but with the roles of beliefs and disbeliefs reversed. It turns out, however, that (B') is a derived rule of the logic BD.

Proposition 9 $\vdash_{\mathrm{BD}}$ *satisfies the property* (B').

A more direct way to obtain new beliefs from current disbeliefs is to consider the converse of the property (B$\rightarrow$D).

(D$\rightarrow$B) If $\Gamma \vdash \overline{\phi}$ then $\Gamma \vdash \neg\phi$.

Observe that WBD, GBD and BD do not satisfy (D$\rightarrow$B). A proof theory for the logic we call BN is then obtained by adding (D$\rightarrow$B) to the inference rules of BD.

Definition 12 *For* $\alpha \in L$ *and* $\Gamma \subseteq L$, α *can be deduced from* Γ *in BN, written as* $\Gamma \vdash_{BN} \alpha$, *iff* (Γ, α) *is in the smallest binary relation from* $\mathcal{P}L$ *to* L *closed under (B), (WD), (D$\perp$) and (D$\rightarrow$B).*

The introduction of (D$\rightarrow$B) does indeed give us a logic that is stronger than BD. It turns out, however, that disbelief now collapses into classical negation

Theorem 4 *For every* $\phi \in L_B$, $\Gamma \vdash_{BN} \overline{\phi}$ *iff* $\Gamma \vdash_{BN} \neg\phi$.

BN is therefore exactly as expressive as PL.

7 Conclusion and future research

Of the four logics of belief and disbelief presented in this paper, the logic BD is the most deserving of such a label. BD covers all the motivations we have mentioned for the explicit introduction of disbeliefs. Contraction with a belief ϕ can be represented explicitly in BD as revision with the disbelief $\overline{\phi}$, just as it can in WBD and GBD. This ability has a useful side-effect as well. In classical belief change, the result of the pathological case in which a belief set is contracted with a tautology, is taken to be the belief set itself, primarily because it is unclear what else the result could be. In fact, it is the only case in which a contraction does not succeed. But in BD (as in WBD and GBD), contraction by ϕ corresponds to a revision by $\overline{\phi}$. So contraction by a tautology is equivalent to revising with the sentence $\overline{\top}$, a disbelief in the tautology. Furthermore, it can be verified that a disbelief in the tautology results in *disbelieving* all (propositional) sentences, but has no effect on the current *beliefs*. The explicit introduction of disbeliefs thus enables us to devise a result which is much more in line with our intuitions.

Like WBD and GBD, BD also provides a well-behaved notion of consequence with respect to disbeliefs. In BD, consequence with respect to disbeliefs is stronger than in WBD but weaker than in GBD. In particular, (D) is satisfied in BD but not in WBD, and BD does not satisfy (D$\vee$), a property satisfied by GBD. As a result of not satisfying (D$\vee$), consequence for BD is weak enough to allow for the expression of *agnosticism* (disbelieving both ϕ and $\neg\phi$) without collapsing into inconsistency. In this respect it is superior to the logic GBD. At the same time, there is a link between beliefs and disbeliefs in BD which is lacking in both WBD and GBD, and which ensures that BD is a suitable base logic for an extended version of belief change in which it is possible to contract with disbeliefs as well. The link between beliefs and disbeliefs also ensures that disbelief, as defined in BD, is an appropriate weaker version of classical negation, as is apparent from the fact that BD satisfies (B$\rightarrow$D) and (D$\not\rightarrow$B), as well as from the intuitively acceptable manner in which the lottery paradox and its variants are handled.

If disbelief is truly to be viewed as a weaker version of negation, it is necessary to conduct a proper investigation into the connection between disbelief, classical negation, and the other propositional connectives. In order to do so, it is necessary to treat $^-$ as full-blown propositional connective, in which sentences such as $p \vee \overline{q}$, $\neg\overline{p}$, and $\overline{\overline{p}}$ have a well-defined meaning.

There is an obvious connection between beliefs and disbeliefs in BD and the modal logic operators of necessity and possibility, specifically in the epistemic logic KD45 [8]. The statement $\Box\phi$ in KD45 corresponds to the assertion that ϕ is believed, while $\Diamond\neg\phi$ corresponds to the assertion that $\neg\phi$ is possible, and hence that ϕ is disbelieved. Observe, though, that statements in BD are object-level assertions, or assertions from a first person perspective (“The sky is blue”), while the corresponding statements in epistemic logics are meta-level assertions, or assertions from a third person perspective (“The agent believes the sky is blue”). However, the differences run slightly deeper and are easily illustrated. To compare BD with KD45 we need to restrict KD45 to only those sentences of the form $\Box\phi$ and $\Diamond\phi$ where Then KD45 satisfies (B), (D$\perp$) and (D). So it is at least as strong as BD. It doesn’t satisfy ($B \not\rightarrow D$), and satisfies ($D \not\rightarrow B$). It does not satisfy (Rej). It supports agnosticism (I disbelieve ϕ as well as $\neg\phi$). It does not satisfy ($D\vee$). So in these respects it has the same behaviour as BD. But it is stronger than BD. This can be seen by looking at what happens when the tautology is disbelieved. In KD45 (and in fact in any normal modal logic i.e., one containing the axiom schema **K**), if $\Diamond(p \wedge \neg p)$ can be deduced from $Gamma$, then so can $\Box\phi$ and $\Diamond\phi$ for every propositional sentence ϕ. In contrast disbelieving the tautology in BD leads to a disbelief in every propositional sentence, but our beliefs remain unaffected. This is a particularly desirable property—as pointed out in the discussion following Proposition 8.

8 Acknowledgements

The authors would like to thank Aditya Ghose for some insightful comments and suggestions.

References

[1] Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson, ‘On the logic of theory change: Partial meet functions for contraction and revision’, *Journal of Symbolic Logic*, **50**, 510–530, (1985).

[2] André Fuhrmann, ‘Theory contraction through base contraction’, *Journal of Philosophical Logic*, **20**, 175–203, (1991).

[3] Peter Gärdenfors, *Knowledge in Flux : Modeling the Dynamics of Epistemic States*, The MIT Press, Cambridge, Massachusetts, 1988.

[4] Aditya Ghose and Randy Goebel, ‘Belief states as default theories: Studies in non-prioritised belief change’, in *ECAI 98. 13th European Conference on Artificial Intelligence*, ed., Henri Prade, pp. 8–12, New York, (1998). Johm Wiley & Sons, Ltd.

[5] Anna Gomolińska, ‘On the logic of acceptance and rejection’, *Studia Logica*, **60**, 233–251, (1998).

[6] Anna Gomolińska and David Pearce, ‘Disbelief Change’, in *Electronic essays on the occasion of the fiftieth birthday of Peter Gärdenfors*, (2001).

[7] Sven-Ove Hansson, ‘Changes of disjunctively closed bases’, *Journal of Logic, Language and Information*, **2**(4), 255–284, (1993).

[8] G.E. Hughes and M.J. Cresswell, *An introduction to Modal Logic*, Methuen, 1972.

[9] Thomas Meyer, Aditya Ghose, and Samir Chopra, ‘Non-prioritised Ranked Belief Change’, in *Theoretical Aspects of Rationality and Knowledge: Proceedings of the Eighth Conference (TARK 2001)*, ed., Johan van Benthem, pp. 151–162, San Francisco, California, (2001). Morgan Kaufmann.

[10] Raymond Reiter, ‘A logic for default reasoning’, *Artificial Intelligence*, **13**, 81–132, (1980).

[11] Hans Rott, *Change, Choice and Inference*, Oxford University Press, 1996.

[12] J. Slupecki and G. Bryll and. U Wybraniec-Skardowska, 'Theory of rejected propositions II', *Studia Logica*, **30**, 97–145, (1972).