Natural deduction for belief *at most*

Johan W. Klüwer[1] and Arild Waaler[2]

[1] Dep. of Philosophy, University of Oslo, `johanw@filosofi.uio.no`
[2] Finnmark College and Dep. of Informatics, University of Oslo, `arild@ifi.uio.no`

Abstract. We present a new approach to the logic of *at most*, introducing the notion of *parametric propositions* to modal natural deduction proofs. We apply the method with a natural deduction formulation of the doxastic logic $\mathrm{N}^{Æ}$, a new system in the "only knowing" family. Using parametric proof rules, we give introduction and elimination rules for belief *at most* that directly match the natural second-order formulation of the concept. $\mathrm{N}^{Æ}$ is sound and complete with respect to the class of intended models. We conjecture that it weakly normalizes and that it satisfies the subformula property.

1 Introduction

The aim of this paper is to provide a new approach to modal logic with *at most* modalities. To the basis of a bimodal natural deduction system, we add introduction and elimination rules for an *at most* modality that directly match its natural definition as a second-order concept. In the rules, second-order quantifiers are reflected in an original use of *parametric propositions*, the role of which is reminiscent of the way parameters are used in the standard analysis of first-order quantifiers. The according pattern of *parametric proposition proof* enable us to directly translate second-order definitions of logical constants into proof rules.[3]

We do not here attempt a general analysis of proof with parametric propositions, but focus on the application to doxastic logic. For this purpose, we introduce the natural deduction system $\mathrm{N}^{Æ}$ of belief, conceivability, and belief *at most*. In $\mathrm{N}^{Æ}$ the modal rules for belief and necessity are standard, providing a well understood analysis of basic concepts, to which parametric rules for belief *at most* are added. We show that on this approach, the introspection properties we take to be valid for belief and conceivability are inherited in the appropriate way by belief *at most*. A consequence of this is that the introspection schemas adopted for belief *at most* in existing "all I know" logics are not just stipulations, but follow from conceptually adequate second-order assumptions that can readily be given intuitive interpretation.

Current approaches to *at most* modalities employ complementary logics, first addressed by Humberstone [3]. These logics adopt one modality which, in the

---

[3] Our deployment of parametric propositions in proof rules is inspired by Frank Pfenning's use of parametric judgments in his treatment of negation [5, 1078 f].

semantics, quantifies over a preferred set of points, and another modality which intuitively quantifies over a complementary set of points. The framework was, in particular, successfully applied by Levesque [4] to the study of default reasoning. However, a weakness of the standard approach to *at most* is that the systems have unintended models (see [3, p. 350]), in which the *at most* modality complementary to belief can not be fully accounted for by reference to strength of belief alone (cf. [8, section 3.2] for a detailed discussion). Perhaps this has not been seen as problematic because the default representation pattern introduced by Levesque employs only formulae that require for their truth that models are of the intended kind (see [8, lemma 2]).

There is however every reason to believe that the scope of *at most* modalities goes beyond the representation of defaults, so finding the correct logic is still of interest. We will in this paper remedy the situation by providing a logic of *at most* that (i) is based only on notions – belief and conceivability – that can be given solid common-sense interpretations, and (ii) whose natural models are of the intended kind.

The object language of our target system $\mathrm{N}^{\text{Æ}}$ is defined as follows: It contains a stock of propositional variables, the constants $\top$ and $\bot$, and connectives $\neg, \vee, \wedge, \supset$, and $\equiv$. Modal operators are $\Diamond$ (conceivability), B (belief) and $\breve{B}$, and their duals: $\Box$ (necessity) is $\neg\Diamond\neg$, b (plausibility) is $\neg B \neg$, and $\breve{b}\varphi$ is $\neg\breve{B}\neg\varphi$. The desired interpretation of $\breve{B}\varphi$ is that at most $\neg\varphi$ is believed. Giving conceptual priority to belief and strength of propositions, we take this to mean that φ is at least as strong as the strongest proposition that is believed. Observe that "φ is at least as strong as ψ" has a natural expression, using the notion of conceivability, in the formula $\Box(\varphi \supset \psi)$. This means that we can specify the meaning of $\breve{B}\varphi$, using a propositional quantifier, as follows.

$$\breve{B}\varphi =_{\text{def}} \forall p(Bp \supset \Box(\neg\varphi \supset p))\,. \tag{1}$$

$\breve{B}\neg\varphi$ is true whenever φ is a proposition that necessarily implies every believed proposition. Its negation $\breve{b}\varphi$ then means that some belief is held, the negation of which is compatible with φ. In terms of possible worlds, $\breve{b}\varphi$ expresses that there is a conceivable world in which φ is true, and which is believed not to be the actual world.

$$\breve{b}\varphi =_{\text{def}} \exists p(Bp \wedge \Diamond(\varphi \wedge \neg p))\,. \tag{2}$$

Overview of the paper: In section 2, we present a Fitch-style framework for natural deduction with modalities. The framework is applied in section 3, introducing the natural deduction system $\mathrm{N}^{\text{Æ}}$. Section 4 recounts the semantics of the system Æ [8], with restriction to models that can properly be said to be intended for the logic of *at most.* $\mathrm{N}^{\text{Æ}}$ is sound and complete with regard to these models, and proofs of this are given in sections 4, 5, and 6. Section 7 addresses normalization and the subformula property, which is essential for proof-search applications.

2 The natural deduction framework

We here present the framework of Fitch-style natural deduction by means of which the system $\mathrm{N}^{Æ}$ is defined in section 3. We assume that a standard set of rules for the propositional operators $\bot$, $\wedge$, $\vee$, and $\supset$ is in place (see e.g. [6]).

Parametric proofs. With the aim of representing second-order quantifiers, we introduce a new kind of subproof pattern, *parametric proof.* A parametric proof may be opened anywhere, and is annotated with a propositional variable which is not previously used in the proof (we will use p, q, r as names for variables). At the opening of a parametric proof, the formula p employed is said to be *parametric*, and should be seen as standing for any proposition. p remains parametric until used in an inference rule that requires it to have an unchanging value, after which it is said to be *fixed* (analogously to the binding of a new constant in elimination of a first-order existential quantifier). A subproof annotated with p is closed by an application of PP.

$$
\begin{array}{r|l}
 & \begin{array}{r|l} p & \vdots \\ & \varphi \qquad p \text{ not in } \varphi \end{array} \\
\rhd & \varphi
\end{array}
\quad \text{PP}
$$

PP is a sound rule: if φ follows from the mere assumption that p is a proposition, but does not involve p itself, then φ must be true.

Negation rules. The rules for negation match the second-order definition of negation as "explosion", $\neg\varphi$ iff $\forall\psi(\varphi \supset \psi)$. Explosion obtains if every proposition is provable. Thus, if we can prove that $\varphi \supset p$ is provable on the mere assumption that p is a proposition $\neg\varphi$ may be introduced.

$$
\begin{array}{r|ll}
 & \varphi \supset p & (p \text{ parametric}) \\
\rhd & \neg\varphi & \text{fixing } p
\end{array}
\quad \neg\mathrm{I}
\qquad\qquad
\begin{array}{r|l}
 & \neg\varphi \\
\rhd & \varphi \supset \psi
\end{array}
\quad \neg\mathrm{E}
$$

In order to obtain a classical system, we add the rule of excluded middle.

$$
\begin{array}{r|l}
\rhd & \varphi \vee \neg\varphi
\end{array}
\qquad \text{tnd}
$$

Modal rules. For reasoning with standard modalities, we add the device of *modal subproof.* The presentation here is intended as schematic, with L standing for any universal modality (here, $\Box$ or B) and M for any existential modality ($\Diamond$ or b). We take the rules for L as basic; rules for M can easily be derived.[4]

---

[4] The rules follow the format introduced by Fitch [e.g., 1]; cf. Siemens [7] and Fitting [2] for application to a variety of modal systems. In the terminology of Fitting, we have *A-style* rules as basic, and see *I-style* rules as abbreviations. In naming the rules, we have followed Siemens [7].

An L-subproof may be opened at any point in a proof. Introduction of $L\varphi$ requires φ to be proven in an L-modal subproof. Iteration into modal subproofs is restricted to formulae φ s.t. $L\varphi$.

$$\begin{array}{r|l} & \begin{array}{l|l} {\scriptstyle L} & \vdots \\ & \varphi \end{array} \\ \rhd & L\varphi \end{array} \qquad L\mathrm{I} \qquad\qquad \begin{array}{r|l} & L\varphi \\ & \begin{array}{l|l} {\scriptstyle L} & \vdots \\ & \varphi \end{array} \\ \end{array} \qquad L\mathrm{R}$$

Adding the rules LI and LR to the classical propositional logic yields the normal system K. Rules may be added to yield stronger systems, corresponding to taking axioms as valid. Here, we will be considering the rules LE (corresponding to the truth schema $L\varphi \supset \varphi$), WLE (corresponding to $L\varphi \supset M\varphi$), LLR (corresponding to $L\varphi \supset LL\varphi$), and MLR (corresponding to $M\varphi \supset LM\varphi$).

$$\begin{array}{r|l} & L\varphi \\ \rhd & \varphi \end{array} \qquad L\mathrm{E} \qquad\qquad \begin{array}{r|l} & L\varphi \\ \rhd & M\varphi \end{array} \qquad \mathrm{W}L\mathrm{E}$$

$$\begin{array}{r|l} & L\varphi \\ & \begin{array}{l|l} {\scriptstyle L} & \vdots \\ & L\varphi \end{array} \end{array} \qquad LL\mathrm{R} \qquad\qquad \begin{array}{r|l} & M\varphi \\ & \begin{array}{l|l} {\scriptstyle L} & \vdots \\ & M\varphi \end{array} \end{array} \qquad ML\mathrm{R}$$

3 The system $\mathrm{N}^{\text{Æ}}$

We propose a system for reasoning about belief B, conceivability $\Diamond$, and belief *at most* $\breve{B}\neg$: Using the naming patterns for rules given in section 2, the system $\mathrm{N}^{\text{Æ}}$, and its provability relation $\vdash$, is given by the following.

For the non-modal fragment, adopt a standard set of rules for $\bot$, $\wedge$, $\vee$, and $\supset$, to which is added PP, $\neg$I, $\neg$E, and tnd. Rules for B are BI, BR, BBR, and bBR, which yields a K45 modality. Rules for $\Box$ are $\Box$I, $\Box$R, $\Box$E, $\Box\Box$R and $\Diamond\Box$R, for an S5 modality. Rules for interaction between B and $\Box$ are defined below: they amount to $\Box B$E, weak introspection $\Box B$R and $\Diamond B$R, and strong introspection $B\Box$R and $b\Box$R. Parametric rules $\breve{B}$E and $\breve{B}$I for the $\breve{B}$ modality conclude the definition of the system.

The fundamental assumption about the relationship between belief and conceivability is that what is not conceivably false (i.e., what is necessary) is also believed. This is captured in $\Box B$E.

$$\begin{array}{r|l} & \Box\varphi \\ \rhd & B\varphi \end{array} \qquad \Box B\mathrm{E}$$

$\Box B$R and $\Diamond B$R are weak introspection principles, representing assumptions that what is necessary, resp. conceivable is believed to be so; they may be seen as weakened versions of $4_\Box$ resp. $5_\Box$.

$$\begin{array}{r|c|l} & \multicolumn{2}{l}{\Box\varphi} \\ & B & \vdots \\ \rhd & & \Box\varphi \end{array} \qquad \Box B\text{R} \qquad\qquad \begin{array}{r|c|l} & \multicolumn{2}{l}{\Diamond\varphi} \\ & B & \vdots \\ \rhd & & \Diamond\varphi \end{array} \qquad \Diamond B\text{R}$$

$B\Box$R and $b\Box$R are strong introspection principles, saying it is inconceivable for the doxastic subject that beliefs and non-beliefs could be different from what they are.

$$\begin{array}{r|c|l} & \multicolumn{2}{l}{B\varphi} \\ & \Box & \vdots \\ \rhd & & B\varphi \end{array} \qquad B\Box\text{R} \qquad\qquad \begin{array}{r|c|l} & \multicolumn{2}{l}{\neg B\varphi} \\ & \Box & \vdots \\ \rhd & & \neg B\varphi \end{array} \qquad b\Box\text{R}$$

The rules for belief *at most* directly reflect the modality's second-order definition (1). We take the pair of rules $\breve{B}$E and $\breve{B}$I to be primitive, and consider the rules $\breve{b}$E and $\breve{b}$I to be derived rules. In $\breve{B}$I and $\breve{b}$E q is assumed to be parametric.

$$\begin{array}{r|l} & \breve{B}\varphi \\ \rhd & B\psi \supset \Box(\neg\varphi \supset \psi) \end{array} \qquad \breve{B}\text{E} \qquad\qquad \begin{array}{r|l} & \breve{b}\varphi \\ \rhd & Bq \wedge \Diamond(\varphi \wedge \neg q) \end{array} \quad \text{fixing } q \qquad \breve{b}\text{E}$$

$$\begin{array}{r|l} & Bq \supset \Box(\neg\varphi \supset q) \\ \rhd & \breve{B}\varphi \end{array} \quad \text{fixing } q \qquad \breve{B}\text{I} \qquad\qquad \begin{array}{r|l} & B\psi \wedge \Diamond(\varphi \wedge \neg\psi) \\ \rhd & \breve{b}\varphi \end{array} \qquad \breve{b}\text{I}$$

4 Soundness

A *model* M is a quadruple (U, U^+, U^-, V), where U is a non-empty set of *points*, U^+ and U^- are subsets of U such that $U^+ \cup U^- = U$, and V is an valuation function which assigns a subset of U to each propositional variable in the language. A satisfaction relation can be defined for each point x in M,

$$\begin{aligned} &M \vDash_x p \quad \text{iff } x \in V(p) \text{ for a propositional variable } p \\ &M \vDash_x \Box\varphi \text{ iff } M \vDash_y \varphi \text{ for each } y \in U \\ &M \vDash_x B\varphi \text{ iff } M \vDash_y \varphi \text{ for each } y \in U^+ \\ &M \vDash_x \breve{B}\varphi \text{ iff } M \vDash_y \varphi \text{ for each } y \in U^- \ , \end{aligned}$$

and as usual for Boolean connectives. A formula is *satisfied* in a model if it is true at one of its points. If $M \vDash_x \varphi$ for all $x \in U$, we write $M \vDash \varphi$ and say that φ is *true in M.* The *truth set* of φ, denoted $||\varphi||$, is the set of points in U at which φ is true.

Observe that all points in a model agree on the truth value of every completely modalized formula. Hence, for such formulae the notions of satisfaction and truth in a model coincide. This justifies use of the notation $M \vDash \varphi$ whenever a completely modalized φ is satisfied in M.

A model of this kind, further exposed in [8, section 3.2], is intended to represent a doxastic subject, where U is the *space of conceivability*, i.e., the range of states of affairs that the subject can conceive of, and the subject has a *belief state*, modeled by U^+, and a *co-belief state*, modeled by U^-. Any point in U^+ is, on the perspective of the doxastic subject, a plausible candidate for being the actual world. U^- is intended to capture the set of *implausible* points, i.e, the set of points that are ruled out (by the subject's belief) as not actual.

Semantically, the intended interpretation requires that the models satisfy a notion of bisectedness to distinguish the belief state from the co-belief state. We will say a model M is *strongly bisected* if $U^+ \cap U^- = \emptyset$ and *weakly bisected* if the following condition holds for all formulae φ: if $||\varphi|| \subseteq ||\psi||$ for every ψ such that $U^+ \subseteq ||\psi||$, then $||\varphi|| \cap U^- = \emptyset$.

That is, in a weakly bisected model $U^+ \cap U^-$ can be non-empty, but there is no formula in the language which we can use to define these points. Although $\mathrm{N}^{\text{Æ}}$ is characterized by the class of strongly bisected models, it is also characterized by the set of weakly bisected models, which, in light of Lemma 1 below, seems to be the more natural class. The two notions of bisectedness coincide for finite models. For infinite models the syntax is in general not expressive enough to enforce a complete separation of the plausible points from the implausible ones; the reason being that the language is countable while the set of propositions in general is uncountable.

We can recast the stipulative definition of the $\breve{B}$ operator in section 1 into semantical terms. A model M satisfies the *$\breve{B}$-property* if for each formula φ, $\breve{B}\varphi$ is true in M iff $B\psi \supset \Box(\neg\varphi \supset \psi)$ is true in M for every formula ψ.

Lemma 1. *A model is weakly bisected if and only if it satisfies the $\breve{B}$-property.*

Proof. Assume that M is weakly bisected. If $\breve{B}\varphi$ is true in M, it follows from the model condition $(U \setminus U^-) \subseteq U^+$ that (†) $B\psi \supset \Box(\neg\varphi \supset \psi)$ is true for every ψ. Conversely, if the latter condition (†) holds, it follows from weak bisection that $||\neg\varphi|| \cap U^- = \emptyset$ and hence that $\breve{B}\varphi$ is true in M. Assume that M is not weakly bisected. There is then a formula φ such that (1) $||\varphi|| \subseteq ||\psi||$ for every ψ such that $U^+ \subseteq ||\psi||$ and (2) $||\varphi|| \cap U^- \neq \emptyset$. (1) implies that the (†) condition holds, while (2) implies that $\breve{B}\varphi$ is not true in M. M hence fails to satisfy the $\breve{B}$-property. □

Theorem 1. *Assume that there is a proof π of φ. Then φ is true at all points in all weakly bisected models.*

Proof. We argue by induction over π and show that all formulae in π are satisfied in all appropriate models at all appropriate points. Initially, all weakly bisected models and all points in any of these models are appropriate, but this may change as we inductively progress through the proof structure. When a hypothetical subproof is opened the appropriate set of points (wrt. the model at hand) is narrowed to those that satisfy the assumption of the subproof. When a $\Box$-subproof (B-subproof) is opened, the set of appropriate points is U (U^+). When a subproof of any of these three types is closed its constraint is suppressed and we reassume the notion of appropriateness that was operational immediately before the subproof was opened. When a parametric proposition is fixed by a $\neg$I ($\breve{B}$I) inference, the set of appropriate *models* is narrowed to those that assign $\emptyset$ (U^+) to p; this model constraint is suppressed when the parametric subproof annotated with p is closed by an application of PP. Given these notions the inductive soundness argument is straightforward. The only non-trivial cases are the $\breve{B}$E rule, for which we use Lemma 1, and the PP rule, in which case we can prove by a simple induction that removing the model constraint does not disturb the truth value of a formula which satisfies the side condition of the rule. $\Box$

A system in which there is an unprovable formula is *consistent*. It follows trivially from Theorem 1 that $\mathrm{N}^{\text{Æ}}$ (and all its subsystems) is consistent.

5 Derivability in $\mathrm{N}^{\text{Æ}}$

For the purpose of comparison to existing doxastic logics, in particular Æ [8], let us look at the strength of the *at most* modality $\breve{B}$ in $\mathrm{N}^{\text{Æ}}$.

The following two lemmas establish the relationship between B, $\breve{B}$, and $\Box$, and the fact that $\breve{B}$ is a normal modality. These results can naturally be considered basic, as they are independent of any introspection schemas for B or $\Box$, relying only on their being normal modalities. (Justification for standard proof steps is abbreviated "ml" for "modal logic".)

Lemma 2. $\vdash \Box\varphi$ *if and only if* $\vdash B\varphi \wedge \breve{B}\varphi$

Proof. Assume $\Box\varphi$; $B\varphi$ follows by $\Box B$E. For $\breve{B}\varphi$, open a parametric subproof with p as parameter. $\Box\varphi$ implies $Bp \supset \Box(\neg\varphi \supset p)$ by ml. Then $\breve{B}\varphi$ may be introduced, and returned from the subproof by PP because p doesn't appear in $\breve{B}\varphi$. Assume $B\varphi$ and $\breve{B}\varphi$; from the latter, infer $B\varphi \supset \Box(\neg\varphi \supset \varphi)$ by $\breve{B}$E. Then $\Box\varphi$ follows by ml. $\Box$

Lemma 3. *$\breve{B}$ is a normal modality.*

Proof. Lemma 2 implies that the necessitation rule $\varphi/\breve{B}\varphi$ is valid. Assume $\breve{B}(\varphi \supset \psi)$ and $\breve{B}\varphi$. It suffices for normality of $\breve{B}$ that we can show $\breve{B}\psi$. Open a subproof with p as parameter; we want to show $Bp \supset \Box(\neg\psi \supset p)$ (with p parametric). Applying $\breve{B}$E twice, we infer $Bp \supset \Box(\neg(\varphi \supset \psi))$ and $Bp \supset \Box(\neg\varphi \supset p)$, and then $Bp \supset \Box(\neg\psi \supset p)$ by ml. Then $\breve{B}\psi$ follows by $\breve{B}$I, and $\breve{B}\psi$ may be returned from the subproof by PP. □

The introspection schemas for B and $\Box$ imply corresponding schemas for $\breve{B}$. $\breve{B}\neg\varphi$ expresses that no proposition stronger than φ is believed, and $\breve{b}\varphi$ that some proposition is believed, the negation of which is compatible with φ. This means that just as B is a universal belief modality and b an existential modality of non-belief, so $\breve{B}$ is a universal non-belief modality, and $\breve{b}$ an existential belief modality [cf. 8, section 3.3]. In $\mathrm{N}^{\text{Æ}}$, bBR and $b\Box$R are negative introspection rules, preserving non-belief, while BBR and $B\Box$R are positive introspection rules, preserving belief.

The following lemma demonstrates that, as intuitively desired, the rules preserving non-belief imply $\breve{B}$ iteration schemas, and the rules preserving belief imply $\breve{b}$ iteration. The rules bBR and BBR, together with weak introspection for $\Box$, validate introspection of $\breve{B}$ and $\breve{b}$ statements into B contexts. With S5 properties of $\Box$ plus strong introspection rules $b\Box$R and $B\Box$R, we gain full introspection for $\breve{B}$, i.e., the schemas $\breve{B}\varphi \supset \breve{B}\breve{B}\varphi$ and $\breve{b}\varphi \supset \breve{B}\breve{b}\varphi$.

Lemma 4.

1. *$\breve{B}\varphi \vdash B\breve{B}\varphi$ given bBR, $\Box BR$.*
2. *$\breve{b}\varphi \vdash B\breve{b}\varphi$ given BBR, $\Diamond BR$.*
3. *$\breve{B}\varphi \vdash \breve{B}\breve{B}\varphi$ given $b\Box R$, $\Box\Box R$.*
4. *$\breve{b}\varphi \vdash \breve{B}\breve{b}\varphi$ given $B\Box R$, $\Diamond\Box R$.*

Proof.

1.

1		$\breve{B}\varphi$	
2	p	$Bp \supset \Box(\neg\varphi \supset p)$	1, $\breve{B}$E
3		$B(Bp \supset \Box(\neg\varphi \supset p))$	2, ml (bBR, $\Box B$R)
4	B	$Bp \supset \Box(\neg\varphi \supset p)$	3, BR
5		$\breve{B}\varphi$	4, $\breve{B}$I fixing p
6		$B\breve{B}\varphi$	BI
7		$B\breve{B}\varphi$	PP

2. Assume $\breve{b}\varphi$ and eliminate into a parametric subproof as Bp and $\Diamond(\varphi \wedge \neg p)$. These iterate by BBR and $\Diamond B$R into a B subproof, where $\breve{b}$I is applied; return by BI and PP.
3. Assume $\breve{b}\breve{b}\varphi$, and eliminate into a subproof parametric in p as Bp and $\Diamond(\breve{b}\varphi \wedge p)$. Open a subproof parametric in q, and a $\Diamond$-subproof with $\breve{b}\varphi$ as the opening formula. Apply $\breve{b}$E to obtain Bq and $\Diamond(\varphi \wedge \neg q)$, which are returned from the

$\Diamond$-subproof by appeal to $b\Box$R and $\Box\Box$R. Apply $\breve{b}$I to get $\breve{b}\varphi$, then PP twice to return this to the initial context.

4. As in 2. above, assume $\breve{b}\varphi$ and eliminate into a subproof parametric in p. Iterate by $B\Box$R and $\Diamond\Box$R into a $\Box$ subproof, apply $\breve{b}$I, then return using $\Box$I and PP.

$\Box$

The introduction of $\breve{B}\varphi$ in line 5 of lemma 4.1 takes place in a different subproof (indeed, in a different modal context) from that in which the parameter p is introduced. The elimination of a parameter (in PP) has been separated from its use in an inference rule that fixes it (here, with the implicit generalization of BI). Indeed, the need to allow such proof steps, which are clearly valid, is the rationale behind our approach of making parametric proof a separate proof rule.

A comparison to quantified logic helps illustrate how this is significant. One standard form of universal generalization can be formulated as, "if Gc has been established in a subproof from the assumption that Fc, where c is a new parameter, then the subproof may be closed and $\forall x(Fx \supset Gx)$ introduced". The employed parameter c is always dropped simultaneously with the introduction of $\forall$. If $\breve{B}$I were modeled directly on the standard form of $\forall$I from quantified logic, it would have a form such as the following: "if $\Box(\neg\varphi \supset p)$ has been proven in a subproof starting with the assumption that p is a fresh parameter s.t. Bp, then the subproof may be closed and $\breve{B}\varphi$ introduced". With this form of $\breve{B}$I, iteration principles for $\breve{B}$ would not be derivable from those for B and $\Box$.

The logic Æ [8] is defined as the smallest system which contains all tautologies, the following axioms,

Def$\Box$	$\Box\varphi \equiv (B\varphi \wedge \breve{B}\varphi)$	T	$\Box\varphi \supset \varphi$
K_B	$B(\varphi \supset \psi) \supset (B\varphi \supset B\psi)$	$K_{\breve{B}}$	$\breve{B}(\varphi \supset \psi) \supset (\breve{B}\varphi \supset \breve{B}\psi)$
$B_\Box$	$B\varphi \supset \Box B\varphi$	$\breve{B}_\Box$	$\breve{B}\varphi \supset \Box\breve{B}\varphi$
$\overline{B}_\Box$	$\neg B\varphi \supset \Box\neg B\varphi$	$\overline{\breve{B}}_\Box$	$\neg\breve{B}\varphi \supset \Box\neg\breve{B}\varphi\,,$

and is closed under all instances of the necessitation rule and modus ponens.[5] The axioms of Æ are all derivable in $\mathrm{N}^{Æ}$. K_B follows from BI and BE by a standard proof. $\breve{B}$I and $\breve{B}$E imply validity of the basic Def$\Box$ (lemma 2) and $K_{\breve{B}}$ (lemma 3). T, $B_\Box$, and $\overline{B}_\Box$ follow immediately from rules $\Box$E, $B\Box$R, and $b\Box$R, respectively. $b\Box$R implies $\breve{B}\varphi \supset B\breve{B}\varphi$ and $\breve{B}\varphi \supset \breve{B}\breve{B}\varphi$, and hence validity of $\breve{B}_\Box$ (lemma 4.1 and 4.2). $B\Box$R implies validity of $\overline{\breve{B}}_\Box$, from $\breve{b}\varphi \supset B\breve{b}\varphi$ and $\breve{b}\varphi \supset \breve{B}\breve{b}\varphi$ (lemma 4.3 and 4.4).

---

[5] In the full version of Æ, the modalities B and $\breve{B}$ are indexed with priorities. We have chosen not to present a prioritized system here, on the belief that it would contribute little to the main issues of the paper.

6 Completeness

A *$N^{Æ}$-maximal* set s is a saturated set of formulae (i.e., a consistent set with every of its proper extensions inconsistent) which satisfies the following property for each formula φ: $\breve{B}\varphi \in s$ iff $B\psi \supset \Box(\neg\varphi \supset \psi) \in s$ for every formula ψ.

Lemma 5. *Let ρ be $N^{Æ}$-consistent. Then $\rho \in s$ for a $N^{Æ}$-maximal set s.*

Proof. Maximal consistent sets are, as usual, constructed on the basis of an enumeration $\varphi_1, \varphi_2, \ldots$ of formulae. We successively construct consistent sets $S_0, S_1, S_2 \ldots$ where $S_0 = \{\rho\}$ and S_{k+1} arises from S_k by adding φ_{k+1} to S_k whenever it is consistent to do so. We do, however, require that if φ_{k+1} is of the form $\neg\breve{B}\varphi$ and is consistent with S_k, there is a propositional variable q at hand which does not occur in any formula in S_k, and that S_{k+1} in this case is $S_k \cup \{\neg\breve{B}\varphi, Bq \wedge \Diamond(\neg\varphi \wedge \neg q)\}$. To see that S_{k+1} is consistent, assume not. We then have that (‡) $S_k, Bq \wedge \Diamond(\neg\varphi \wedge \neg q) \vdash \breve{B}\varphi$. Moreover, since q does not occur in S_k, it follows by an application of $\breve{b}$E that $S_k, \neg\breve{B}\varphi \vdash Bq \wedge \Diamond(\neg\varphi \wedge \neg q)$. Together with (‡) this gives $S_k \vdash \breve{B}\varphi$, contradicting the assumption. Clearly $s = \cup_{k \geq 0} S_k$ is a saturated set which contains ρ. If $\breve{B}\varphi \in s$, $\breve{B}$E along with saturation gives that $B\psi \supset \Box(\neg\varphi \supset \psi) \in s$ for every formula ψ. If $\neg\breve{B}\varphi \in s$, there is by construction of s a q such that $Bq \wedge \Diamond(\neg\varphi \wedge \neg q) \in s$, hence s is N^{Æ}-maximal. □

Let s be N^{Æ}-maximal. M_s, the *canonical model wrt. s*, is $(U_s^{\Box}, U_s^{B}, U_s^{\breve{B}}, V_s)$, where $U_s^{\Box}$ is the set of N^{Æ}-maximal sets such that $t \in U_s^{\Box}$ iff $\{\varphi \mid \Box\varphi \in s\} \subseteq t$. Likewise, $t \in U_s^{B}$ iff $\{\varphi \mid B\varphi \in s\} \subseteq t$ and $t \in U_s^{\breve{B}}$ iff $\{\varphi \mid \breve{B}\varphi \in s\} \subseteq t$. $V_s(p) = \{t \mid t \in U_s^{\Box}$ and $p \in t\}$. In the completeness proof below we make extensive use of the observation in the previous section to the effect that N^{Æ} extends Æ.

Theorem 2. *$N^{Æ}$ is complete wrt. the class of weakly bisected models.*

Proof. Let φ be consistent. By lemma 5, φ occurs in a N^{Æ}-maximal set s. By using the introspection principles of Æ (i.e., $B_{\Box}$, $\breve{B}_{\Box}$, $\overline{B}_{\Box}$ and $\overline{\breve{B}}_{\Box}$) it is straightforward to prove that for all t in $U_s^{\Box}$ the following three properties hold: (i) $U_s^{\Box} = U_t^{\Box}$, (ii) $U_s^{B} = U_t^{B}$ and (iii) $U_s^{\breve{B}} = U_t^{\breve{B}}$, while the T-axiom entails that (iv) $s \in U_s^{\Box}$. On the basis of these properties and the fact that the three modalities are all normal (this immediate for $\Box$ and B and established in lemma 3 for $\breve{B}$), it is straightforward to prove the Truth lemma: *$M_s \vDash_t \varphi$ iff $\varphi \in t$ for each $t \in U_s^{\Box}$*. The Truth lemma entails both that M_s satisfies φ and, together with lemma 5 and the construction of N^{Æ}-maximal sets, that M_s satisfies the $\breve{B}$-property. It remains to be shown that $U_s^{\Box} = U_s^{B} \cup U_s^{\breve{B}}$ and that M_s is weakly bisected. The former property follows easily from the Def$\Box$ scheme of Æ, while the latter follows from lemma 1. □

Corollary 1. *$N^{Æ}$ is characterized by the set of strongly bisected models.*

Proof. Since $\mathrm{N}^{Æ}$ is sound wrt. the class of weakly bisected models (Theorem 1) it is clearly sound wrt. the smaller class of strongly bisected models. Conversely, if φ is consistent, it has by Theorem 2 a weakly bisected model $M = (U, U^+, U^-, V)$. If we enrich the language with a new propositional variable q we can construct a new model M_q from M as follows: $U_q = U_q^+ \cup U_q^-$, $U_q^+ = \{\langle x, q\rangle \mid x \in U^+\}$, $U_q^- = \{\langle x, \overline{q}\rangle \mid x \in U^-\}$, $V_q(q) = U_q^+$, and $V_q(p) = \{\langle x, \cdot\rangle \mid x \in V(p)\}$ for $p \neq q$. M_q is a model of $\varphi \wedge Bq \wedge \breve{B}\neg q$ and is strongly bisected. $\square$

7 Conjectures

The system satisfies local reduction; in particular, the introduction of $\breve{B}$ formulae does not increase deductive power, which means that the elimination rule for $\breve{B}$ is sound. To see this, consider the following pattern, where the notation $\delta[\psi/p]$ refers to the derivation resulting from replacing every occurrence of the formula p in derivation δ by the formula ψ.

$$\begin{array}{ll}
\vdots \quad \delta & \\
Bp \supset \Box(\neg\varphi \supset p) & \\
\breve{B}\varphi & \breve{B}\mathrm{I}, \text{ fixing } p \\
B\psi \supset \Box(\neg\varphi \supset \psi) & \breve{B}\mathrm{E} \\
\vdots \quad \epsilon & \\
\chi &
\end{array}
\quad \Longrightarrow_R \quad
\begin{array}{l}
\vdots \quad \delta[\psi/p] \\
B\psi \supset \Box(\neg\varphi \supset \psi) \\
\vdots \quad \epsilon \\
\chi
\end{array}$$

Because introduction of $\breve{B}\varphi$ requires a fresh proposition symbol p, i.e., a proposition about which nothing is known, derivation of $B\psi \supset \Box(\neg\varphi \supset \psi)$ can not require greater deductive resources than the formula $Bp \supset \Box(\neg\varphi \supset p)$ on which the introduction of $\breve{B}\varphi$ is based. A similar pattern holds for $\breve{b}$. On the basis of this we conjecture interesting proof-theoretical properties.

Conjecture 1. $\mathrm{N}^{Æ}$ satisfies weak normalization.

Conjecture 2. If φ is $\mathrm{N}^{Æ}$-provable from no assumptions, it has a proof in which each instance of $\breve{B}$E is instantiated with a formula ψ which is a subformula of φ (and correspondingly for the derived rule $\breve{b}$I).

The natural way to prove the latter property is to exploit the structure of normal form natural deduction proofs. If it holds, $\mathrm{N}^{Æ}$ is an analytic system. From the point of view of automated deduction, this is of course a desirable property.

References

[1] Frederic Brenton Fitch. *Symbolic Logic: An Introduction.* The Ronald Press Company, New York, 1952.

[2] Melvin Fitting. *Proof Methods for Modal and Intuitionistic Logics.* Reidel, Dordrecht, 1983.

[3] I. L. Humberstone. Inaccessible worlds. *Notre Dame Journal of Formal Logic*, 24:346–352, 1983.

[4] Hector J. Levesque. All I know: A study in autoepistemic logic. *Artificial Intelligence*, 42:263–309, 1990.

[5] Frank Pfenning. Logical frameworks. In Alan Robinson and Andrei Voronkov, editors, *Handbook of Automated Reasoning*, pages 1065–1147. Elsevier Science Publishers, 2001.

[6] Dag Prawitz. *Natural Deduction: A Proof-Theoretical Study.* Almquist & Wiksell, Stockholm, 1965.

[7] David F. Siemens, Jr. Fitch-style rules for many modal logics. *Notre Dame Journal of Formal Logic*, 18:631–636, 1977.

[8] Arild Waaler, Johan W. Klüwer, Tore Langholm, and Espen H. Lian. Only knowing with degrees of confidence. *Journal of Applied Logic*, 2005. To appear. A preprint is available at `http://folk.uio.no/johanw/ok-doc.pdf`.