

# On Fibring Semantics for BDI Logics

Guido Governatori<sup>1</sup> Vineet Padmanabhan<sup>2</sup> Abdul Sattar<sup>2</sup>

<sup>1</sup> School of Information Technology & Electrical Engineering  
The University of Queensland, Queensland, Australia  
guido@itee.uq.edu.au

<sup>2</sup> Knowledge Representation & Reasoning Unit (KRRU)  
School of Information Technology  
Griffith University, GoldCoast Campus, Queensland, Australia  
[vineet,sattar]@cit.gu.edu.au

**Abstract.** This study examines BDI logics in the context of Gabbay's *fibring* semantics. We show that *dovetailing* (a special form of fibring) can be adopted as a semantic methodology to combine BDI logics. We develop a set of interaction axioms that can capture static as well as dynamic aspects of the mental states in BDI systems, using Catach's *incestual* schema  $G^{a,b,c,d}$ . Further we exemplify the constraints required on fibring function to capture the semantics of interactions among modalities. The advantages of having a fibred approach is discussed in the final section.

## 1 Introduction

BDI based Agent-systems [1,7,4] have been well studied in the AI literature. The design of these systems hinges on mental attitudes like *beliefs* (B), *desires* (D) (or *goals* (G)) and *intentions* (I). The formalization of these three mental attitudes and their interactions have been captured to a great extend using Multi-Modal Logics (e.g., [15,17,3]). Moreover, some additional operators like *Capability*, *Opportunity*, *Obligation* [12,13,16] and several action constructs [18] have been introduced to improve the expressive power of the logics involved.

Much of the research in BDI is focused on improving the expressive power of the language with single operators and identifying formal properties of each of them. However, the general methodology for combining the different logics involved has been mainly neglected. For instance, any BDI system modelling rational agents consists of a combined system of logics of knowledge, belief, time and modal logic of actions. Each of these combined systems was presented and motivated by a different author for different reasons and different applications. We believe that investigating a general methodology for combining the component logics involved in BDI-like systems is an important research issue. This would result in a better understanding of the formal groundings of complex rational agent architectures and enable the designer to elegantly and easily incorporate new features of rational agency within her framework. Moreover the proposed general methodology should permit a modular treatment of the modal components, whereby, each component is analysed and developed on its own, with the most appropriate methodology for it, and is then reused in the combination. Furthermore each module has its own features but the framework remains unchanged among

the combined systems. Finally the combined system should offer preservation of some important logical properties of its elements.

In this study we investigate one such method, viz. *fibring* [5], and use it to reconstruct the logical account of BDI in terms of *dovetailing* (a special case of fibring) together with the multi-modal semantics of Catach [2]. In doing so we identify a set of interaction axioms for BDI, based on the incestual schema  $G^{a,b,c,d}$ , which covers many of the existing BDI axioms and also make possible the generation of a large class of new ones. Further we identify conditions under which completeness transfers from the component logics ( $L_1$  and  $L_2$ ) to their *fibred/dovetailed* composition ( $L_{1,2}^F/L_{1,2}^D$ ), with the help of canonical model structures. We also show completeness preservation in the case of interaction axiom of the form  $\Box_1\alpha \Rightarrow \Box_2\alpha$  ( $L_{1,2}^{F,D} \oplus \Box_1\alpha \Rightarrow \Box_2\alpha$ ). Our study differs from that of other combining techniques like *fusion* in terms of the interaction axiom. For instance, normal bimodal and polymodal logics without any interaction axioms are well-studied as *fusions* of normal monomodal logics in [19]. Property transfer for such logics has been dealt with in [9]. For a slightly different account on fusions of logics one can refer [11]. Moreover *fusions* of normal modal logics without interaction axioms is the same as *dovetailing*. But difficulty arises with the extra interaction axiom. Then we need a more general concept like *fibring*. Our study starts with the assumption that the combination of two complete logics need not be complete when we add interaction axioms [8]. We want to identify conditions under which completeness can be preserved when we include interaction axioms like above.

## 2 BDI & Multi-Modal Logics

The main advantage of using Multi-Modal Logics in BDI is their ability to express complex modalities, that can capture the inter-relationships existing among the different mental attitudes. This can be achieved by either composing modal operators of different types, or by using formal operations over modalities.

For instance the idea that an agent's goal is always supported by its belief is captured by the following BDI axioms:

$$\text{GOAL}^{KD}(\alpha) \Rightarrow \text{BEL}^{KD45}(\alpha) \quad (1)$$

$$\text{GOAL}^{KD}(\alpha) \Rightarrow \text{BEL}^{KD45}(\text{GOAL}^{KD}(\alpha)) \quad (2)$$

The axiom systems for each of the mental operators is written as a superscript and the justification for such a preference is given in [15]. For a further account of the different axiom systems and the semantics for BDI logics refer to [15]; accordingly the semantic conditions for (1) and (2) are:

$$\text{if } (w_x, w_y) \in \text{BEL} \text{ then } (w_x, w_y) \in \text{GOAL} \quad (3)$$

$$\text{if } (w_x, w_y) \in \text{BEL} \text{ and } (w_y, w_z) \in \text{GOAL} \text{ then } (w_x, w_z) \in \text{GOAL} \quad (4)$$

Condition (3) captures inclusion (containment) of a binary relation for beliefs in the relation for goals, whereas (4) captures the combined transitivity on two binary relations  $R_1$  and  $R_2$ . The above mentioned axioms and conditions together with a range

of additional axioms and constructs characterizes a typical BDI system. The properties of soundness, completeness etc are defined via canonical Kripke structures and for a further account see [14]

The point here is that the axiom systems for GOAL and BEL is a combination of other axiom systems and hence they are different. They can be considered as two different languages  $L_1$  and  $L_2$  with  $\Box_1$  (BEL) and  $\Box_2$  (GOAL) built up from the respective sets  $A_1$  and  $A_2$  of atoms and supported by the logics  $KD45$  and  $KD$ . Hence we are dealing with two different systems  $S_1$  and  $S_2$  characterized, respectively, by the class of Kripke models  $\mathcal{K}_1$  and  $\mathcal{K}_2$  and this fact should be taken into consideration while defining semantic conditions for interaction axioms like those given above. For instance, we know how to evaluate  $\Box_1\alpha$  (BEL( $\alpha$ )) in  $\mathcal{K}_1$  ( $KD45$ ) and  $\Box_2\alpha$  (GOAL( $\alpha$ )) in  $\mathcal{K}_2$  ( $KD$ ). We need a method for evaluating  $\Box_1$  (resp.  $\Box_2$ ) with respect to  $\mathcal{K}_2$  (resp.  $\mathcal{K}_1$ ).

The problem in its general form is how to construct a multi-modal logic containing several unary modalities, each coming with its own system of specific axioms. The fibring technique introduced in the next section allows one to combine systems through their semantics. The fibring function can evaluate (give a yes/no) answer with respect to a modality in  $S_2$ , being in  $S_1$  and vice versa. Each time we have to evaluate a formula  $\alpha$  of the form  $\Box_2\beta$  in a world in a model of  $\mathcal{K}_1$  we associate, via the fibring function  $\mathbf{F}$ , to the world a model in  $\mathcal{K}_2$  where we calculate the truth value of the formula. Formally

$$w \models_{m \in \mathcal{K}_1} \Box_2\beta \text{ iff } \mathbf{F}_2(w) \models_{m' \in \mathcal{K}_2} \Box_2\beta$$

$\alpha$  holds in  $w$  iff it holds in the model associated to  $w$  through the fibring function  $\mathbf{F}$ . Moreover  $\alpha$  could be a mixed wff consisting of operators from  $L_1$  and  $L_2$  (for instance  $\alpha$  can be  $\Diamond_1\Box_2q$ ). This is possible because the axiom systems of BEL and GOAL itself are combined systems. Then we have to say that the wff  $\alpha$  belongs to the language  $L_{(1,2)}$ . The existing BDI semantics fails to give adequate explanation for such formulas. The problem becomes even more complex when we allow the system to vary in time. Then we have to combine the BDI system with a suitable temporal logic. The fibring/dovetailing technique provides a general methodology for such combinations as shown in the next section.

It is also important to note that since each mental operator (BEL, GOAL) itself is a combination of different axiom systems, the underlying multi-modal language ( $L_{BDI}$ ) should be such that we should be able to develop each single operator on its own within its own semantics so that in the later stage the operators and models can be glued together through fibring/dovetailing to form a comprehensive system. The multi-modal language should also be able to express multiple concepts like rational agents, actions, plans etc. The problem with the existing BDI-Language is that each time we want to incorporate a new concept we have to redefine the system and come up with characterization results. For instance, if we want to capture the notion of actions, plans, programs etc. in BDI, we need to come up with specific axiom systems for each of them and then show that they are characterized within the system. What we need is a set of interaction axioms that can generate a range of multi-modal systems for which there is a general characterization theorem so that we could avoid the need for showing it each time a new system is considered. To this end we adopt the class of interaction axioms  $G^{a,b,c,d}$  of Catach [2] that can account for a range of multi-modal systems.

### 3 Fibring of Modal Logics

In this section we present a general semantic methodology, called fibring, for combining modal (BDI) logics and a variant of it called dovetailing. Two theorems stating relationships between dovetailing and BDI logics and dovetailing and fibring are shown. It is shown that the existing BDI logic is a dovetailed system.

The Fibring methodology allows one to combine systems through their semantics. The method helps in combining arbitrary logical systems in a uniform way and gives a new insight on possible worlds semantics [5,6]. Its idea is essentially to replace the notion of a possible world as a *unit* by another Kripke structure. For instance, if we consider the earlier example of a mixed formula,  $\alpha = \diamond_1 \Box_2 q$ , the way the fibring function works can be shown as follows.  $\alpha$  can be considered as a formula of  $L_1$  (as the outer connective is  $\diamond_1$ ). From the point of view of language  $L_1$  this formula has the form  $\diamond_1 p$ , where  $p = \Box_2 q$  is atomic since  $L_1$  does not recognize  $\Box_2$ . The satisfaction condition for  $\diamond_1 p$  can be given considering a model  $\mathbf{m}^1 = (S^1, R^1, a^1, h^1)$  such that  $a^1 \models \diamond_1 p$  (where  $S^1$  is the set of possible worlds,  $a^1 \in S$  is the actual world and  $R \subseteq S \times S$  is the accessibility relation,  $h$  is the assignment function, a binary function, giving a value  $h(t, p) \in \{0, 1\}$  for any  $t \in S$  and atomic  $p$ ). For this to hold there should be some  $t \in S^1$  such that  $a^1 R^1 t$ , and we have  $t \models_1 p$ , i.e.,  $t \models_1 \Box_2 q$ . Since  $\Box_2$  is not in the language of  $L_1$  the normal evaluation is not possible. The basic idea of fibring is to associate with each  $t \in S^1$ , a model  $\mathbf{m}_t^2 = (S_t^2, R_t^2, a_t^2, h_t^2)$  of  $L_2$  and by evaluating  $\Box_2 q$  in the associated model, thus we have

$$t \models_1 \Box_2 q \text{ iff } a_t^2 \models_2 \Box_2 q.$$

If we take  $\mathbf{F}^1$  to be the fibring function associating the model  $\mathbf{m}_t^2$  with  $t$  then  $F^1(t) = \mathbf{m}_t^2$  and the semantics for the language  $\mathcal{L}_{(1,2)}$  has a model of the form  $(S^1, R^1, a^1, h^1, \mathbf{F}^1)$ .

**Fibring two semantics** Let  $I$  be a set of labels representing intentional states, and  $L_i, i \in I$  be modal logics whose respective modalities are  $\Box_i, i \in I$ .

**Definition 1** [5] A *fibred model* is a structure  $(W, S, R, \mathbf{a}, h, \tau, \mathbf{F})$  where

- $W$  is a set of possible worlds;
- $S$  is a function giving for each  $w$  a set of possible worlds,  $S^w \subseteq W$ ;
- $R$  is a function giving for each  $w$ , a relation  $R^w \subseteq S^w \times S^w$ ;
- $\mathbf{a}$  is a function giving the actual world  $\mathbf{a}^w$  of the model labelled by  $w$ ;
- $h$  is an assignment function  $h^w(q) \subseteq S^w$ , for each atomic  $q$ ;
- $\tau$  is the semantical identifying function  $\tau : W \rightarrow I$ .  $\tau(w) = i$  means that the model  $(S^w, R^w, \mathbf{a}^w, h^w)$  is a model in  $\mathcal{K}_i$ , we use  $W_i$  to denote the set of worlds of type  $i$ ;
- $\mathbf{F}$ , is the set of fibring functions  $\mathcal{F} : I \times W \mapsto W$ . A fibring function  $\mathcal{F}$  is a function giving for each  $i$  and each  $w \in W$  another point (actual world) in  $W$  as follows:

$$\mathcal{F}_i(w) = \begin{cases} w & \text{if } w \in S^{\mathbf{m}} \text{ and } \mathbf{m} \in \mathcal{K}_i \\ \text{a value in } W_i, & \text{otherwise} \end{cases}$$

such that if  $x \neq y$  then  $\mathcal{F}_i(x) \neq \mathcal{F}_i(y)$ .

Satisfaction is defined as follows with the usual truth tables for boolean connectives:

$$t \models p \text{ iff } h(t, p) = 1, \text{ where } p \text{ is an atom}$$

$$t \models \Box_i A \text{ iff } \begin{cases} t \in \mathbf{m} \text{ and } \mathbf{m} \in \mathcal{K}_i \text{ and } \forall s (tRs \rightarrow s \models A), \text{ or} \\ t \in \mathbf{m}, \text{ and } \mathbf{m} \notin \mathcal{K}_i \text{ and } \forall \mathcal{F} \in \mathbf{F}, \mathcal{F}_i(t) \models \Box_i A. \end{cases}$$

We say the model satisfies  $A$  iff  $w_0 \models A$ .

A fibred model for  $L_I^F$  can be generated from fibring the semantics for the modal logics  $L_i, i \in I$ . The detailed construction runs as follows: Let  $\mathcal{K}_i$  be a class of models  $\{\mathbf{m}_1^i, \mathbf{m}_2^i, \dots\}$  for which  $L_i$  is complete. Each model  $\mathbf{m}_n^i$  has the form  $(S, R, a, h)$ . The actual world  $a$  plays a role in the semantic evaluation in the model, in so far as satisfaction in the model is defined as satisfaction at  $a$ . We can assume that the models satisfy the condition  $S = \{x \mid \exists n a R^n x\}$ . This assumption does not affect satisfaction in models because points not accessible from  $a$  by any power  $R^n$  of  $R$  do not affect truth values at  $a$ . Moreover we assume that all sets of possible worlds in any  $\mathcal{K}_i$  are all pairwise disjoint, and that there are infinitely many isomorphic (but disjoint) copies of each model in  $\mathcal{K}_i$ . We use the notation  $\mathbf{m}$  for a model and present it as  $\mathbf{m} = (S^{\mathbf{m}}, R^{\mathbf{m}}, a^{\mathbf{m}}, h^{\mathbf{m}})$  and write  $\mathbf{m} \in \mathcal{K}_i$ , when the model  $\mathbf{m}$  is in the semantics  $\mathcal{K}_i$ . Thus our assumption boils down to  $\mathbf{m} \neq \mathbf{n} \Rightarrow S^{\mathbf{m}} \cap S^{\mathbf{n}} = \emptyset$ . In fact a model can be identified by its actual world, i.e.,  $\mathbf{m} = \mathbf{n}$  iff  $a^{\mathbf{m}} = a^{\mathbf{n}}$ . Then the fibred semantics can be given as follows:

- $W = \bigcup_{\mathbf{m} \in \bigcup_i \mathcal{K}_i} S^{\mathbf{m}}$ ;
- $R = \bigcup_{\mathbf{m} \in \bigcup_i \mathcal{K}_i} R^{\mathbf{m}}$ ;
- $h(w, q) = h^{\mathbf{m}}(w, q)$ , for the unique  $\mathbf{m}$  such that  $w \in S^{\mathbf{m}}$ ;
- $a^w = a^{\mathbf{m}}$  for the unique  $\mathbf{m}$  such that  $w \in S^{\mathbf{m}}$ .

**Dovetailing** Dovetailing is a special case of fibring in the sense that the dovetailed model must agree with the current world on the values of atoms. For instance, in the previous section we saw that the functions  $\mathcal{F}$  can be viewed as functions giving for each  $t \in S^1 \cup S^2$ , an element  $\mathcal{F}(t) \in S^1 \cup S^2$  such that if  $t \in S^i$  then  $\mathcal{F}(t) \in S^j, i \neq j$ . If  $L_1$  and  $L_2$  share the same set of atoms  $Q$  then we can compare the values  $h(t, q)$  and  $h(\mathcal{F}(t), q)$  for an atom  $q$  which need not be identical. If we require from the fibring functions that for each  $t \in S^i, \mathcal{F}_j(t) \in S^j$  and each  $q \in Q$  we want

$$h(t, q) = h(\mathcal{F}_i(t), q).$$

Then this fibring case is referred to as *dovetailing*. This means that the actual world of the model fibred at  $t, \mathcal{F}_i(t)$ , can be identified with  $t$ . The set of fibring functions  $\mathbf{F}$  is no longer needed, since we identified  $t$  with  $\mathcal{F}_i(t)$ , for every fibring function  $\mathcal{F}$ .

**Definition 2** [5] Let  $L_i$  be modal logics, with  $\mathcal{K}_i$  the class of models for  $L_i$ . Let  $L_I^D$  (the dovetailing combination of  $L_i, i \in I$ ) be defined semantically through the class of all (dovetailed) models of the form  $(W, R, a, h)$ , where  $W$  is a set of worlds,  $a \in W$ ,  $h$  is an assignment and for each  $i \in I, R(i) \subseteq W \times W$ . We require that for each  $i, (W, R(i), a, h)$  is a model in  $\mathcal{K}_i$ . It is further required that all  $t \in W$  be such that there exist  $n_1, \dots, n_k$  and  $i_1, \dots, i_k$  such that  $aR^{n_1}(i_1) \circ R^{n_2}(i_2) \cdots \circ R^{n_k}(i_k)t$  holds. The satisfaction condition  $w \models A$ , is defined by induction as

- $w \models q$  if  $w \in h(q)$  for  $q$  atomic;
- $w \models \Box_i A$  if for all  $y \in W$ , such that  $wR(i)y$  we have  $y \models A$ ;
- $\models A$  iff for all models and actual worlds  $a \models A$ .

Two theorems are given below, the proof of which can be found in [5].

**Theorem 1 (Dovetailing and Normal Modal Logic).** *Assume  $L_i$ ,  $i \in I$  all are extensions of  $K$  formulated using traditional Hilbert axioms and the rule of necessitation, then  $L_I^D$  can be axiomatized by taking the union of the axioms and the rules of necessitation for each modality  $\Box_i$  of each  $L_i$*

**Theorem 2 (Fibring = Dovetailing).** *If  $L_i$ ,  $i \in I$  admit necessitation and satisfy the disjunction property, then  $L_I^F = L_I^D$ .*

It is immediate to see that BDI logic without interaction axioms is nothing else but normal multi-modal logics —combinations of normal modal logics (e.g., the basic BDI logic proposed in [15] is the combination of a  $KD45$  modality for BEL,  $KD$  for GOAL and  $KD$  for INT)— hence, according to Theorem 1, dovetailing provides a general methodology for generating BDI-like systems.

## 4 Semantics for Mental States

In the previous section we have seen how to provide a general semantics for BDI logics without interaction of modalities. However, mental states are very often connected to each other, for example the interaction axioms like (1) and (2); thus what is needed is a methodology to capture them. In this section we use Catach approach [2] to extend dovetailing in order to develop a general semantics that covers both the basic modalities and their interactions. Briefly Catach’s approach runs as follows:

Let  $I$  be a set of atomic labels; complex labels can be built from atomic ones using the neutral element “ $\lambda$ ”, the sequential operator “;”, and the union operator “ $\cup$ ”. If  $i$  is an atomic label and  $\alpha$  a well-formed formula, then the expression  $[i]\alpha$  corresponds to the modal formula  $\Box_i \alpha$ , and  $\langle i \rangle \alpha$  to  $\Diamond_i \alpha$ . Furthermore we assume that  $[\lambda] = \langle \lambda \rangle$ . The transformation of complex labels into modalities is governed by the following rules:

$$[\lambda]\alpha \Leftrightarrow \alpha; \quad [a;b]\alpha \Leftrightarrow [a][b]\alpha; \quad [a \cup b]\alpha \Leftrightarrow [a]\alpha \wedge [b]\alpha.$$

According to the above conditions we can identify, for example, the formula  $\Box_1 \Box_2 A \wedge \Box_3 A \wedge A$  with the expression  $[(1;2) \cup 3 \cup \lambda]$ .

Let us consider now the expression  $\langle a \rangle [b]\alpha \Rightarrow [c] \langle d \rangle \alpha$ , known as the  $a, b, c, d$ -incestuality axiom (we will use  $G^{a,b,c,d}$  to refer to it). It can be used to generate, among others, the well know  $D$ ,  $T$ ,  $B$ , 4 and 5 axioms of modal logic. For example, when  $a = b = \lambda$  and  $c = d = 1$  we obtain the symmetry axiom  $B$  for  $\Box_i$ .

It is then immediate to see that the above axiom schema covers many existing systems of multi-modal logic, including the BDI system and make the generation of a large class of new ones possible.

*Example 1.* Let  $\alpha$  be a formula and BEL, GOAL, INT, CAP, OPP and RES be the modal operators for the mental constructs; then the following are instances of  $G^{a,b,c,d}$ .

- C1**  $\text{GOAL}(\alpha) \Rightarrow \text{BEL}(\alpha)$  (Inclusion)  
**C2**  $\text{INT}(\alpha) \Rightarrow \{\text{GOAL}(\alpha) \Rightarrow \text{BEL}(\alpha)\}$  (Relative Inclusion)  
**C3**  $\text{RES}(e) \Leftrightarrow \text{CAP}(e) \wedge \text{OPP}(e)$  (Union)

The axioms C2 and C3 are possible additions to the existing BDI axioms. The above axioms (as well as others) can be used to represent various concepts such as rational agents, programs, actions etc. For instance C2 captures the fact that an agent's intention to achieve  $\alpha$  is supported by having a goal towards  $\alpha$  and this goal is based on its belief of  $\alpha$ . The existing BDI framework lacks such axioms.

As far as dovetailed models are concerned it is possible to define a mapping  $\rho$  between labels and the accessibility relations of dovetailed models.

**Definition 3** *Let  $a$  and  $b$  be labels,  $i$  an atomic label, and  $(W, R(i), a, h)$  a dovetailed model. Then*

$$\rho(i) = R(i); \quad \rho(\lambda) = \Delta; \quad \rho(a;b) = \rho(a)|\rho(b); \quad \rho(a \cup b) = \rho(a) \cup \rho(b);$$

where the operators  $\cup$  (union) and  $|$  (composition) are defined for binary relations, and  $\Delta$  is the diagonal relation over  $W$

**Definition 4** *Let  $a, b, c$ , and  $d$  be labels. A dovetailed model  $D = (W, R(i), a, h)$  enjoys the  $a, b, c, d$ -incestuality property iff the following condition holds for  $D$ .*

$$\rho(a)^{-1}|\rho(c) \subseteq \rho(b)|\rho(d)^{-1}.$$

The incestuality condition can be reformulated as follows:

If  $(w, w') \in \rho(a)$  and  $(w, w'') \in \rho(c)$  then there exists  $w'''$  such that  $(w', w''') \in \rho(b)$  and  $(w'', w''') \in \rho(d)$ .

**Theorem 3.** [2] *Let  $L_{BDI}$  be a normal multi-modal system built from a finite set of axioms  $\mathbf{G}^{a,b,c,d}$ . Then  $L_{BDI}$  is determined by the class of dovetailed models satisfying the  $a, b, c, d$ -incestuality properties.*

Catach originally proved the above theorem for what he calls multi-frames. Trivially multi-frames correspond to dovetailed models. In particular this result provides the characterization of a wide class of interaction axioms such as the relationships among mental attitudes of rational agents in terms of dovetailing.

## 5 Conditions on the Fibring Function

Section 3 establishes how BDI-like systems (without interaction axioms) can be reconstructed using dovetailing and section 4 introduces a general axiom schema through which we can generate a range of BDI-like interaction axioms. In this section we demonstrate with the help of an example what conditions would be required on the fibring functions in order to cope with the  $a, b, c, d$ -incestuality schema. As noted earlier we assume that the combination of two complete logics need not be complete when we include interaction axioms. We want to identify conditions under which completeness

can be preserved. But before identifying the specific conditions on the fibring functions we need to introduce certain notions and constructions related to completeness preservation in terms of canonical models and canonical logics.

In the canonical model construction a *world* is a maximal consistent sets of wff. Thus for any normal propositional modal system  $S$ , its canonical model  $\langle W, R, V \rangle$  is defined as follows:

- $W = \{w : w \text{ is a maximal } S\text{-consistent set of wff}\}$ ;
- For any pair of worlds  $w$  and any  $w' \in W$ ,  $wRw'$  iff  $\{\alpha : \Box\alpha \in w\} \subseteq w'$ ;
- For any variable  $p$  and any  $w \in W$ ,  $V(p, w) = 1$  iff  $p \in w$ .

But in the case of a fibred model the above construction needs to be modified accordingly as follows: Let  $L_i, i \in I$  be monomodal normal logic with languages  $\mathcal{L}_i$ . Let  $M_L$  be the set of all  $L$ -maximal consistent sets of formula. Given a set  $S$  of formulas,  $L^{\Box_i}(S) = \{A : \Box_i A \in S\}$  and  $L^{\Diamond_i}(S) = \{A : A = \Box_i B \text{ or } A = \Diamond_i B, A \in S\}$ . The canonical model for  $L_i^F, C_i^F$  is the structure  $\langle W, S, R, \mathbf{F}, a, \tau, h \rangle$ , where

- $W = M_L \times I$ .
- $S$  is a function  $W \mapsto \wp W$  such that  $S^w = \{(x, i) \in W : \tau(w) = i\}$ . In other words the set of worlds of the same type as  $w$ .
- $R^w \subseteq S^w \times S^w$  such that  $xR^w y$  iff  $L^{\Box_{\tau(w)}}(x) \subseteq y$ .
- $\mathbf{F}$  is the set of functions  $\mathcal{F} : I \times W \mapsto W$  (fibring functions) such that

$$\mathcal{F}_i(x, j) = \begin{cases} (x, j) & i = j \\ (x, i) & x = a^w \\ (y, i) & \text{otherwise} \end{cases}$$

- where  $L^{\mathcal{L}_i(x)} \subseteq y$ , and if  $x \neq y$ , then  $\mathcal{F}_i(x) \neq \mathcal{F}_i(y)$ .
- $a^w = w$ .
- $\tau(x, i) = i$
- $h(p, w) = 1$  iff  $p \in w$ , for  $p$  atomic.

**Lemma 1.** *For every formula  $\alpha$  and every world  $w$  in the canonical model*

$$h(w, \alpha) = 1 \text{ iff } \alpha \in w.$$

*Proof.* The proof is by induction on the complexity of  $\alpha$ . The only difference with the proof of the monomodal case is when  $\alpha = \Box_i \beta$  and  $\tau(w) \neq i$ . If  $h(w, \Box_i \beta) = 1$ , then for every  $\mathcal{F} \in \mathbf{F}$   $h(\mathcal{F}_i(w), \Box_i \beta) = 1$ , and we can apply the standard construction for modalities and we obtain that  $\Box_i \beta \in \mathcal{F}_i(w)$ . Let us now suppose that  $\Box_i \beta$  is not in  $w$ . Since  $w$  is maximal  $\neg \Box_i \beta \in w$ ; thus  $\Diamond_i \neg \beta \in w$ .  $L^{\mathcal{L}_i} \subseteq \mathcal{F}_i(w)$ , hence  $\Diamond_i \neg \beta \in \mathcal{F}_i(w)$ , from which we derive a contradiction. Thus  $\Box_i \beta \in w$ . The other direction is similar.

As an immediate consequence of the Lemma we have the following theorem.

**Theorem 4.**  $L_i^F \vdash \alpha$  iff  $C_i^F \models \alpha$ .

**Definition 5** Let  $F_L$  be the frame of the canonical model for  $L$ .  $L$  is canonical iff for every valuation  $V$ ,  $(F_L, V)$  is a model for  $L$ .

Clearly the above definition is equivalent to the usual condition for a modal logic to be canonical (i.e., that the frame of the canonical model is a frame for  $L$ ). However the fibring construction inherits the valuation functions from the underlying models, and we can obtain different logics imposing conditions on the fibring functions based on the assignments of the variables. The fibred frame for  $\mathcal{L}_{1,2}$  is obtained in the same way as the fibred model, replacing the occurrences of models with frames.

**Lemma 2.** *Let  $M_i^F = (W, S, R, \mathbf{F}, a, \tau, h)$  be the canonical model for  $L_i^F$ . Then for each  $w \in W$  ( $S^w, R^w, h^w$ ) is the canonical model for  $\tau(w)$ .*

*Proof.* By inspection on the construction of the canonical model for  $L_i^F$ .

From the above Lemma we obtain:

**Theorem 5.** *Let  $L_i, i \in I$  be canonical monomodal logics. Then  $L_i^F$  is canonical.*

For instance the inclusion axiom  $\Box_1 A \Rightarrow \Box_2 A$  is characterized by the dovetailed models where  $R_2 \subseteq R_1$ . However, such a constraint would be meaningless for fibred models where each modality has its own set of possible worlds. So, what is the corresponding condition on fibred models? A fibring function is defined as

$$\mathcal{F} : I \times W \rightarrow W$$

where  $I$  is the set of modalities involved and  $W$  is a set of possible worlds. It is worth noting that given a world we can identify the model it belongs to, and that there is a bijection  $M$  between the actual worlds and their models. So to deal with the inclusion axiom the following constraint must be satisfied:

$$\forall w \in W \forall \mathcal{F} \in \mathbf{F} : M(\mathcal{F}_2(w)) \sqsubseteq_N M(\mathcal{F}_1(w)) \quad (5)$$

where  $\sqsubseteq_N$  is the inclusion morphism thus defined:

**Definition 6** *Let  $\mathbf{m}_1$  and  $\mathbf{m}_2$  be two models. Then  $\mathbf{m}_2 \sqsubseteq_N \mathbf{m}_1$  iff there is a morphism  $\mathbf{w} : W_2 \mapsto W_1$ , such that*

- for each atom  $p$ ,  $h_2(w, p) = h_1(\mathbf{w}(w), p)$ ;
- if  $xR_2y$  then  $\mathbf{w}(x)R_1\mathbf{w}(y)$ .

The constraint on the fibring functions to support the *inclusion axiom*, is in alliance with the incestuality axiom  $G^{a,b,c,d}$  as stated in the previous section, that is,  $R_2 = \rho(c)$  and  $R_1 = \rho(b)$ . The incestuality axiom can be characterised by giving appropriate conditions that identify the (fibred) models  $\mathbf{m}_1$  and  $\mathbf{m}_2$  involved in the inclusion morphism.

It is now possible to provide a characterization of the fibring/dovetailing of normal modal logics  $L_1$  and  $L_2$  with inclusion axiom (i.e.,  $\Box_1 \alpha \Rightarrow \Box_2 \alpha$ ).

**Theorem 6.** *Let  $L_1$  and  $L_2$  be two canonical normal modal logics and let  $L_{1,2}$  be the logics obtained by fibring/dovetailing  $L_1$  and  $L_2$ . Then  $L_{1,2} \oplus \Box_1 \alpha \Rightarrow \Box_2 \alpha$  is characterized by the class of fibred/dovetailed models satisfying (5).*

*Proof.* For the proof we have to note that, thanks to the fact that  $L_1$  and  $L_2$  are canonical, for any pair of world  $w$  and  $v$ , the sets of maximal consistent associated with them are the same, i.e.,  $S^w = S^v$ , they are the set of all the maximal consistent sets. Thus no matter of the fibring function we chose, we have that the structure of the the models obtained from  $\mathcal{F}_1$  and  $\mathcal{F}_2$  are the same. Therefore we can use the identity over  $M_L$  as the morphism  $\mathbf{w}$  in the inclusion morphism. Let  $\mathbf{m}_1 = \mathcal{F}_1(w)$  and  $\mathbf{m}_2 = \mathcal{F}_2(w)$ . Clearly for every world  $v$  and every atom  $p$ ,  $h^2(v, p) = h^1(\mathbf{w}(v), p)$ . Since  $\Box_1 A \Rightarrow \Box_2 A$  is an axiom, then  $\forall w \in W, \Box_1 A \Rightarrow \Box_2 A \in w$ . This implies  $L^{\Box_1}(w) \subseteq L^{\Box_2}(w)$ . Hence if  $wR^2y$ , then  $L^{\Box_2}(w) \subseteq y$ . Since the morphism  $\mathbf{w}$  is the identity over  $M_L$ , we have that  $L^{\Box_1}(\mathbf{w}(w)) \subseteq L^{\Box_2}(\mathbf{w}(w)) \subseteq \mathbf{w}(y)$ . Therefore  $\mathbf{w}(w)R^1\mathbf{w}(y)$ .

It is well known that normal modal operators have certain properties that are occasionally considered undesirable for the common sense notions that they are intended to formalise. For instance the property of *Logical Omniscience* though could hold for the beliefs of an agent is certainly undesirable for the knowledge part. For example to say that an agent knows all the logical consequences of its knowledge ( $\Box\varphi \wedge \Box(\varphi \Rightarrow \psi) \Rightarrow \Box\psi$ ) is to live in an idealized world. The fibring methodology can be used to combine single modal logics that are not normal. However, in general, simple adjustments are required to deal with classes of non-normal modal logics. In what follows we show the modifications required for quasi-normal modal logics (i.e. modal logics containing  $K$  and closed under  $\text{RM } \vdash \alpha \Rightarrow \beta / \vdash \Box\alpha \Rightarrow \Box\beta$ ). The first thing we have to consider is that the structure of models appropriate for such a class is  $(W, N, R, a, h)$  where  $W$  and  $a$  are as usual,  $N \subseteq W$  (representing the set of normal worlds),  $R \subseteq N \times W$ , and we have the following two additional clauses on the valuation function  $h$ :

$$\text{if } w \notin N, h(w, \Box\alpha) = 0; \quad \text{if } w \notin N, h(w, \Diamond\alpha) = 1.$$

Fibred, dovetailed, and canonical models can be obtained accordingly with the appropriate trivial modifications (cf. [10]).<sup>3</sup> We are now ready to give the completeness theorem for the fibring of monotonic modal logics.

**Theorem 7.** *Let  $L_i, i \in I$  be quasi-normal modal logics classes of structures  $\mathcal{K}_i$  and set of theorems  $T_i$ . Let  $T_i^F$  be the following set of wffs of  $L_i^F$ .*

1.  $T_i \subseteq T_i^F$ , for every  $i \in I$ ;
2. If  $A(x_m) \in T_i$  then  $A(x_m/\Box_j\alpha_j) \in T_i^F$ , for any  $\Box_i\alpha_m, i \in I$ ;
3. **Monotonic Modal Fibring Rule:** If  $\Box_i$  is the modality of  $L_i$  and  $\Box_j$  that of  $L_j$ , where  $i, j$  are arbitrary, with  $i \neq j$ .

$$\frac{\bigwedge_{k=1}^n \Box_i A_k \Rightarrow \bigvee_{k=1}^m \Box_i B_k \in T_i^F}{\Box_j^n \bigwedge_{k=1}^n \Box_i A_k \Rightarrow \Box_j^n \bigvee_{k=1}^m \Box_i B_k \in T_i^F} \text{ for all } n;$$

4.  $T_i^F$  is the smallest set closed under 1, 2, 3, modus ponens and substitution.

Then  $T_i^F$  is the set of all wffs of  $L_i^F$  valid in all the fibred monotonic structures of  $L_i^F$ .

*Proof.* The proof is a trivial modification of that of Theorem 3.10 of [5].

<sup>3</sup> For normal modal logics  $N = W$ , thus any normal modal logic is also quasi-normal.

Intuitively the meaning of the Monotonic Modal Fibring Rule has to do with the substitutions of formulas of one language into a formula of the other language. If the substituted formulas are proof theoretically related we want to propagate this relation to the other language. Moreover there are formal similarities between the Monotonic Fibred Rule and RM. Consider an implication of the form  $A \Rightarrow B$  where  $A$  and  $B$  are built from atoms of the form  $\Box_i C$ . There our special RM says that if  $\vdash A \Rightarrow B$  then we can derive  $\vdash \Box_j A \Rightarrow \Box_j B$  for any modality  $\Box_j$  other than  $\Box_i$ .

A similar theorem can be proved for the dovetailing of quasi-normal modal logics with the appropriate modifications on the Dovetail Modal Rule given by Gabbay [5].

At this stage we have to revise our definition of inclusion morphism.

**Definition 7** Let  $\mathbf{m}_1$  and  $\mathbf{m}_2$  be two quasi-normal models.  $\mathbf{m}_1 \sqsubseteq_M \mathbf{m}_2$  iff

1.  $\mathbf{m}_1 \sqsubseteq_N \mathbf{m}_2$ ; and
2. if  $w \notin N_1$  then  $\mathbf{w}(w) \notin N_2$ .

**Theorem 8.** Let  $L_1$  and  $L_2$  be the logic obtained by the canonical quasi-normal modal logic fibring/dovetailing of  $L_1$  and  $L_2$ . Then  $L_{1,2}^M \oplus \Box_1 \alpha \Rightarrow \Box_2 \alpha$  is characterized by the class of fibred/dovetailed models satisfying

$$\forall w \in W \forall \mathcal{F} \in \mathbf{F} : M(\mathcal{F}_2(w)) \sqsubseteq_M M(\mathcal{F}_1(w)).$$

*Proof.* The proof is analogous to that of Theorem 6.

The main consequence of the above theorem is that it shows how to extend the full power of fibring to non-normal modal logics with interaction axioms, including combinations of a range of modalities required to model complex BDI systems.

**Corollary 1.** Let  $L_1$  and  $L_2$  be the logic obtained by the canonical {quasi-}normal modal logic fibring/dovetailing of  $L_1$  and  $L_2$ . Then  $L_{1,2}^M \oplus \Box_1 \alpha \Rightarrow \Box_2 \alpha$  is canonical.

## 6 Discussion

We have investigated the relationships between BDI logics and Gabbay's fibring semantics. In particular we have shown how to reconstruct the logical account of BDI in terms of dovetailing and Catach's approach. Roughly fibring (dovetailing) examines and develops each single operator on its own, within its own semantics, and then the operators and models are glued together to form a comprehensive system.

The proposed methodology provides a general framework for BDI in so far as it is not restricted to particular operators and, at the same time, it offers easy tools to study properties (e.g., soundness, completeness, ...) of the resulting systems. Transfer results, even if limited to very small cases of interaction axioms, would be of extreme importance to BDI theorists as this will help shift their attention from the single case analysis to the general problem of combination of mental states. The problem of interaction axioms between logic fragments in a combined logic might become more central among the BDI theorists. Moreover the proposed approach is not restricted to normal modal operators —the use of non-normal epistemic operators is one of the common strategies to (partially) obviate the problem of logical omniscience.

As we have seen dovetailing is a particular form of fibring, and the latter offers a more fine grained analysis of the concepts at hand. In other words we can say that fibring is more selective than dovetailing. Indeed, there are formulas which are valid under dovetailing but false under some interpretations using fibring; thus some unwanted consequences could be discarded by fibring. Remember that the condition for dovetailing states that sources and targets of the fibring functions agree on the evaluation of atoms. However, this is not required for fibring, so, we can say that fibring can be thought of as a change of perspective in passing from a model (modal operator) to another model (modal operator). This has some important consequences: for example, the interpretation of a piece of evidence may depend on the mental state used to look at it.

## References

1. M.E. Bratman, D.J. Israel, and M.E Pollack. Plans and resource-bounded practical reasoning. *Computational Intelligence*, 4:349–355, 1988.
2. L. Catach. Normal multimodal logics. In *In Proc. National Conference on AI(AAAI-88)*, pages 491–495, 1988.
3. P. R. Cohen and H. J. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42(3), 1990.
4. K. Fischer and Muller J. P. & Pischel M. A pragmatic BDI architecture. In *Intelligent Agents II, LNAI-1037*, volume 1037, pages 203–218. Springer-Verlag, 1996.
5. D. M. Gabbay. *Fibring Logics*. Oxford University Press, Oxford, 1998.
6. D. M. Gabbay and G. Governatori. Dealing with Label Dependent Deontic Modalities. In *Norms, Logics And Information Systems*. IOS Press, Amsterdam, 1999.
7. M. Georgeff and F. Ingrand. Decision making in an embedded reasoning system. In *In proc. of the International Joint Conference on Artificial Intelligence- IJCAI*, pages 972–978, 1989.
8. M. Kracht. *Tools and Techniques for Modal Logics*. Elsevier, 1999.
9. M. Kracht and F. Wolter. Properties of independently axiomatizable bimodal logics. *The Journal of Symbolic Logic*, 56(4):1469–1485, 1991.
10. E.J. Lemmon. Algebraic semantics for modal logic II. *Journal of Symbolic Logic*, 31:191–218, June 1966.
11. Franz Baader & Carsten Lutz, Holger Sturn, and Frank Wolter. Fusions of description logics and abstract description systems. *Journal of Artificial Intelligence Research*, 16:1–58, 2002.
12. Lin Padgham and Patrick Lambrix. Agent capabilities: Extending bdi theory. In *Proceedings of AAAI-2000, Austin, Texas, USA*, pages 68–73. AAAI/MIT Press, 2000.
13. Vineet Padmanabhan, Guido Governatori, and Abdul Sattar. Actions Made Explicit in BDI. In *AI 2001: Advances in Artificial Intelligence, LNAI-2256*. Springer Verlag, 2001.
14. A. S. Rao. Decision procedures for propositional linear-time belief- desire-intention logics. In *Intelligent Agents Volume II, LNAI-1037*, pages 33–48. Springer-Verlag, 1996.
15. A.S. Rao and Georgeff M.P. Modelling rational agents within a bdi-architecture. In *Principles of Knowledge Representation and Reasoning (KR'91)*. Morgan Kaufmann, 1991.
16. Y. Shoham. Agent oriented programming. *Artificial Intelligence*, 60:51–92, 1993.
17. Munindar P. Singh. Semantical considerations on intention dynamics for bdi agents. *Journal of Experimental and Theoretical Artificial Intelligence*, 1998.
18. B. Van Linder. *Modal Logic for Rational Agents*. PhD thesis, Department of Computer Science, Utrecht University, 19th June 1996.
19. Frank Wolter. Fusions of modal logics revisited. In *Advances in Modal Logic*, volume 1. CSLI Lecture notes 87, 1997.