Decidability via Filtration of Neighbourhood Models for Multi-Agent Systems

Clara Smith[1] and **Leandro Mendoza**[2] and **Agustín Ambrossio**[3]

Abstract. Lately, many multi-agent systems (MAS) are designed as multi-modal systems [9, 15, 23, 22, 26, 28, 18]. Moreover, there are different techniques for combining logics, such as products, fibring, fusion, and modalisation, among others [1, 14, 16]. In this paper we focus on the combination of special-purpose logics for building "on demand" MAS. From these engineering point of view, among the most used normal logics for modeling agents' cognitive states are logics for beliefs, goals, and intentions, while, perhaps, the most well-known non-normal logics for MAS is the logic of agency (and, possibly, ability). We explore combinations of these normal and non-normal logics. This lead us to handle Scott-Montague structures, (neighbourhood models, in particular) which can be seen as a generalization of Kripke structures [20].

Interested in the decidability of such structures, which is a guarantee of correct systems and their eventual implementations, we give a new presentation for existing theorems that generalize the well-known results regarding decidability through the finite model property via filtrations for Kripke structures. We understand that the presentation we give, based on neighbourhood models, better fits the most accepted and extended logic notation actually used within the MAS community.

1 Motivation and Aims

In [32] Smith and Rotolo adopted [13]s cognitive model of individual trust in terms of necessary mental ingredients which settle under what circumstances an agent x trusts another agent y with regard to an action or state-of-affairs, i.e. under which beliefs and goals an agent delegates a task to another agent. Using this characterization of individual trust, the authors provided a logical reconstruction of different types of collective trust, which for example emerge in groups with multi-lateral agreement, or which are the glue for grounding *in solidum* obligations raising from a "common front" of agents (where each member of the front can behave, in principle, as creditor or debtor of the whole). These collective cognitive states were characterized in [32] within a multi-modal logic based on [9]s axiomatisation for collective beliefs and intentions combined with a non-normal modal logic for the operator Does for agency.

In a subsequent work, the multi-relational model in [32] was reorganized as a fibring, a particular combination of logics which amounts to place one special-purpose normal logics on top of another [31]. In this case, the normal logic was put on top of the non-normal one. For doing this, authors first obtained two restrictions of the original logics. By exploiting results in regard to some techniques for combining logics, it was proved that [32]s system is complete and decidable. Hence, the sketch for an appropriate model checker is there outlined.

One motivation regarding a further combination of those special purpose logics for MAS is the aim to have an expressive enough system for modelling interactions between a behavioural dimension and a cognitive dimension of agents, and testing satisfiability of the corresponding formulas. For example, for modelling expressions such as $\mathrm{Does}_i\,(\mathrm{Bel}_j\,\mathscr{A})$ which can be seen as a form of persuasion or influence: agent i makes agent j have $\mathscr{A}$ as belief. This formula cannot be written in the fibred language in [31] neither in the original language in [32] because such languages have a restriction over the form of the *wffs*: no modal operator can appear in the scope of a Does. In [31], authors outlined a combination of the normal and the non-normal counterparts of the base logics. That combination lead to an ontology of pairs of situations allowing a structural basis for more expressiveness of the system. That combination is the result of (again) splitting of the original structure, which is a multi-relational frame of the form [32, 17]:

$$\mathfrak{F} = \langle A, W, \{B_i\}_{i\in A}, \{G_i\}_{i\in A}, \{I_i\}_{i\in A}, \{D_i\}_{i\in A}\rangle$$

where: A is a set of agents, W is a set of posible worlds, and $\{B_i\}, \{G_i\}, \{I_i\}, \{D_i\}$ are the accessibility relations for beliefs, goals, intentions, and agency respectively. The underlying set of worlds of the combination is an ontology of pairs of worlds (w_N, w_D). There are two structures where to respectively test the validity of the normal modalities and the non-normal modalities. The former is a Kripke model; the latter a neighbourhood model. The definition of a formula being satisfied in the combined model at a state (w_N, w_D) amounts to a scan through the combined structure, done according to which operator is being tested. Normal operators move along the first component w_N, and non-normal operators move along the second component of the current world w_D.

Regarding the application to agents, it is also common that the cognitive modalities are extended with temporal logics. For example, Schild [29] provides a mapping from Rao and Georgeff's BDI logic [27] to μ-calculus [24]. The model of Rao and Georgeff is based on a combination of the branching time logic CTL^* [8] and modal operators for beliefs, desires, and intentions. Schild collapses the (original) two dimensions of time and modalities onto a one dimensional structure. J. Broersen [5] presents an epistemic logic that incorporates interactions between time and action, and between knowledge and action.

Correspondingly, H. Wansing in [2] points out that (i) agents act in time, (ii) obligations change over time as a result of our actions and the actions of others, and (iii) obligations may depend on the

[1] FACET, UCALP, Argentina and Facultad de Informática, UNLP, Argentina
[2] Facultad de Informática, UNLP, Argentina and CONICET
[3] FACET, UCALP, Argentina

future course of events. In ([2], Section 10.3) he adopts a semantics reflecting the non-determinism of agency: models are based on trees of moments of time branching to the future. Agentive sentences are history dependent, formulas are not evaluated at points in time but rather at pairs (*moment*, *history*), where *history* is a linearly ordered set of moments.

Cohen and Levesque [7, 21] embed, using function mappings, a modal logic of beliefs and goals with a temporal logic with non-deterministic and parallel features.

In this paper we define a combination of logics for MAS as a special case of neighbourhood structures. Previously, we give a new presentation of decidability results which apply to a particular kind of models: neighbourhood models. In the literature, the analysis of transfer of logical properties from special purpose logics to combined ones is usually based on properties of normal logics. It is claimed that the proof strategies in the demonstration of transference of properties of normal logics could in principle be applied to non-normal modal logics [12]. In a mono-modal logic with a *box* modality, normality implies that the following formulas are valid: $\Box(p \to q) \to (\Box p \to \Box q)$ and $\Box(p \wedge q) \leftrightarrow (\Box p \wedge \Box q)$, as well as the admission of the rule from $\vdash \mathcal{A}$ infer $\vdash \Box\mathcal{A}$ [3, 12]. None of this is assumed to hold for a non-normal logics. We indeed use a non normal modal logic for agency, as developed by Elgesem [11, 17]; and aim to put it to work with normal logics for, e.g, beliefs and goals. The logic of agency extends classical propositional logic with the unary symbol Does satisfying the following axioms: $\neg(\mathrm{Does}\,\top)$, $(\mathrm{Does}\,\mathscr{A}) \wedge (\mathrm{Does}\,\mathscr{B}) \Rightarrow \mathrm{Does}(\mathscr{A} \wedge \mathscr{B})$ and $\mathrm{Does}\,\mathscr{A} \Rightarrow \mathscr{A}$ together with the rule of Modus Ponens and the rule saying that from $\mathscr{A} \Leftrightarrow \mathscr{B}$ you can conclude $\mathrm{Does}\,\mathscr{A} \Leftrightarrow \mathrm{Does}\,\mathscr{B}$. The intended reading of $\mathrm{Does}\,\mathscr{A}$ is that 'the agent brings it about that $\mathscr{A}$'. (See Section 2.1 in [11].) A detailed philosophical justification for this logic is given in [11] and neighborhood and selection function semantics are discussed in [11, 17].

One advantage regarding the choice of a logic of agency such as Does relies on the issue of action negation. For Does, and for other related logics of action such as the one in [5], action negation is well-understood: given that the logic for Does is Boolean, it is easy to determine what $\neg\mathrm{Does}\,\mathscr{A}$ means. This allows providing accurate definitions for concepts such as e.g. "refrain", especially useful in normative MAS: I have the opportunity and ability to do something, but I do not perform it as I have the intention not to. Up to now, although addressed, there are no outstanding nor homogeneous solutions for the issue on action negation in other relevant logics for MAS such as dynamic logics (see e.g. [4, 5, 25]).

We organize the work as follows. In Section 2 we directly adapt for neighbourhood models the strategy in [3] regarding the *finite model property (FMP) via filtration*. This includes: (i) establishing conditions for finding a filtration of a neighbourhood model, (ii) the demonstration of a filtration theorem for the neighbourhood case, (iii) guaranteeing the existence of a filtration, and (iv) the proof of the *FMP Theorem* for a mono-modal neighbourhood model. In Section 3 we show how the results in Section 2 can be applied for proving decidability of a neighbourhood model with more than one modality. We also devise examples for a uni-agent mono-modal non-normal system, a uni-agent multi-modal system and a multi-modal multi-agent system. In Section 4 we concentrate on a combined *MAS*, with an underlying neighbourhood structure. Conclusions end the paper.

2 Decidability for the neighbourhood case through the extension of the FMP strategy for the Kripke case.

We mentioned that normal logics can be seen as a platform for the study of transference of decidability results for non-normal logics and combination of logics. We rely on well-studied results and existing techniques for Kripke structures, which are usual support of normal logics, to provide a new presentation of existing decidability results for a more general class of structures supporting non-normal logics.

We start from the definitions given by P. Blackburn *et. al.* [3]. In [3](Defs. 2.36, 2.38 and 2.40), the construction of a finite model for a Kripke structure is supported in: (i) the definition of a filtration, (ii) the Filtration Theorem, (iii) the existence of a filtration for a model and a subformula closed set of formulas, and (iv) the Finite Model Property Theorem via Filtrations.

B. Chellas, in its turn, defined filtrations for minimal models in [6] (Section 7.5). Minimal models are a generalization of Kripke ones. A minimal model is a structure $\langle W, N, P \rangle$ in which W is a set of possible worlds and P gives a truth value to each atomic sentence at each world. N, is a function that associates with each world a collection of sets of worlds. The notation used throughout is one based on truth sets ($\|\mathscr{A}\|$ is the set of points in a model where the *wff* $\mathscr{A}$ is true). Truth sets are a basic ingredient of selection function semantics.

In what follows we give a definition of filtration for Scott-Montague models using a neighbourhood approach and notation. Neighbourhood semantics is the most important (as far as we consider) generalization of Kripke style (relational) semantics. The set of possible worlds is replaced by a Boolean algebra, then the concept of validity is generalized to the set of true formulas in an arbitrary subset of the Boolean algebra, but (generally for every *quasi-classical* logics) the subset must be a filter. This 'neighbourhood approach' focuses on worlds, which directly leads us to the underlying net of situations that ultimately support the system: relative to a world w we are able to test whether agents believe in something or carry out an action. The neighbourhood semantics better adapts to the specification of most prevailing modal multi-agent systems, which lately tend to adopt the Kripke semantics with a notation given as in [3]. This because, probably, that notation is more intuitive for dealing with situations and agents acting and thinking according to situations, rather than considering formulas as 'first class' objects. This is crucial in current practical approaches to agents; in a world an agent realises its posibilities of succesful agency of $\mathscr{A}$, its beliefs, it goals, all relative to the actual world w, In this perspective, situations are a sort of "environmental support" for agent's *internal configuration* and *visible actions*. Worlds are, therefore, in a MAS context, predominantly, abstract descriptions of external circumstances of an agent's community that allow or disallow actions, activate or nullify goals.

That is why we prefer to work with neighbourhood models as models for MAS, keeping in mind that, while it is possible to devise selection function models for MAS, this is not nowadays usual practice. Also, as it is well-known, the difference between selection function semantics and neighbourhood semantics is merely at the intuitive level (their semantics are equivalent, and both known as ScottMontague semantics [17]).

P. Schotch has already addressed the issue of paradigmatic notation and dominating semantics for modalities. In his work [30] he points out that the necessity truth condition together with Kripkean structures twistedly "represent" the model-theoretic view of the area, given that -among other reasons- many "nice" logics can be devised with those tools. Moreover, due to this trend, he notes that previous complex and important logics (due to Lewis, or to the "Pennsylvania School") have become obsolete or curiosities just because their semantics is less elegant.

We adopt an eclectic position in this paper: we choose a structure that allows non-normal semantics and we go through it with the notation as given in [3], which is currently well-accepted and well-understood for modal MAS.

Next we outline some tools for finding a filtration of a neighbourhood model. We generalize the theorems for Kripke structures given in [3].

Definition 1 (Neighbourhood Frame)**.** A neighbourhood frame [20, 6] is a tuple $\langle W, \{N_w\}_{w\in W}\rangle$ where:

1. W is a set of worlds, and
2. $\{N_w\}_{w\in W}$ is a function assigning to each element w in W a class of subsets of W, the neighbourhoods of w.

We will be working with a basic modal language with a single unary modality, let us say '#'. We asume that this modality has a neighbourhood semantics. For example, '#' may be read as the Does operator, or an ability operator, as proposed by Elgesem [11]; or represent a "refrain" operator based on Does and other modalities such as ability, opportunity and intentions.

Definition 2 ((Recall Def. 2.35 in [3]) Closure)**.** A set of formulas Σ is closed under subformulas if for all formulas φ, if $\varphi \vee \varphi' \in \Sigma$ then so are φ and φ'; if $\neg\varphi \in \Sigma$ then so is φ; and if $\#\varphi \in \Sigma$ then so is φ. (For the Does modality, for example, if $\mathrm{Does}\,\varphi \in \Sigma$ so is φ).

Definition 3 (Neighbourhood Model)**.** We define $\mathfrak{M} = \langle W, \{N_w\}, V\rangle$ to be a model, where $\langle W, \{N_w\}\rangle$ is a neighbourhood frame, and V is a valuation function assigning to each proposition letter p in Σ a subset $V(p)$ of W (i.e. for every propositional letter we know in which worlds it is true).

Given Σ a subformula closed set of formulas and given a neighbourhood model $\mathfrak{M}$, let $\equiv_\Sigma$ be a relation on the states of $\mathfrak{M}$ defined by $w \equiv_\Sigma v$ iff $\forall\,\varphi \in \Sigma$ ($\mathfrak{M}, w \models \varphi$ iff $\mathfrak{M}, v \models \varphi$). That is, for all *wff* φ, φ is true in w iff is also true in v. Clearly $\equiv_\Sigma$ is an equivalence relation. We denote the equivalence class of a state w of $\mathfrak{M}$ with respect to $\equiv_\Sigma$ by $[w]_\Sigma$ (or simply $[w]$ when no confusion arises).

Let $W_\Sigma = \{[w]_\Sigma \,/\, w \in W\}$.

Next we generalize for neighbourhood models the concept of filtration given in [3].

Definition 4 (Filtrations for the neighbourhood case)**.** Suppose $\mathfrak{M}^f$ is any model $\langle W^f, \{N_w\}^f, V^f\rangle$ such that $W^f = W_\Sigma$ and:

1. If $U \in N_w$ then $\{[u]/u \in U\} \in N^f_{[w]}$,
2. For every formula $\#\varphi \in \Sigma$, if $\mathcal{U} \in N^f_{[w]}$ and $(\forall [u] \in \mathcal{U})(\mathfrak{M}, u \models \varphi)$, then $\mathfrak{M}, w \models \#\varphi$,
3. $V^f(p) = \{[w] \,/\, \mathfrak{M}, w \models p\}$, for all proposition letter p in Σ.

Condition (1) requires that for every neighbourhood of w there is a corresponding neighbourhood of classes of equivalences for the class of equivalence of w (i.e. $[w]$) in the filtration. Condition (2) settles, among classes of equivalences, the satisfaction definition regarding a world and its neighbourhoods.

We use U for the neighbourhoods in the original model $\mathfrak{M}$, and $\mathcal{U}$ for the neighbourhoods of $[w]$ in the filtration $\mathfrak{M}^f$.

Theorem 1 (Filtration Theorem for the neighbourhood case.)**.** *Consider a unary modality '#'. Let $\mathfrak{M}^f$ be a filtration of $\mathfrak{M}$ through a subformula closed set Σ. Then for all φ in Σ and all w in $\mathfrak{M}$, $\mathfrak{M}, w \models \varphi$ iff $\mathfrak{M}^f, [w] \models \varphi$. That is, filtration preserves satisfiability.*

Proof. We show that $\mathfrak{M}, w \models \varphi$ iff $\mathfrak{M}^f, [w] \models \varphi$. As Σ is subformula closed, we use induction on the structure of φ. We focus on the case $\varphi = \#\gamma$. Assume that $\#\gamma \in \Sigma$, and that $\mathfrak{M}, w \models \#\gamma$. If $\mathfrak{M}, w \models \#\gamma$ then there is a neighbourhood U such that $U \in N_w$ and $(\forall u \in U)(\mathfrak{M}, u \models \gamma)$, that is, for every world in that neighbourhood, γ holds. Thus, by application of the induction hypothesis, for each of those u we have that $\mathfrak{M}^f, [u] \models \gamma$. By condition (1) above, $\{[u]/u \in U\} \in N^f_{[w]}$. Hence $\mathfrak{M}^f, [w] \models \#\gamma$.

Conversely we have to prove that if $\mathfrak{M}^f, [w] \models \varphi$ then $\mathfrak{M}, w \models \varphi$.

Assume that $\varphi = \#\gamma$ and $\mathfrak{M}^f, [w] \models \#\gamma$. By truth definition, there exists $\mathcal{U}$ neighbourhood of $[w]$ such that $(\forall [u] \in \mathcal{U})(\mathfrak{M}^f, [u] \models \gamma)$. Then by inductive hypothesis $(\forall [u] \in \mathcal{U})(\mathfrak{M}, u \models \gamma)$. Then by condition (2) $\mathfrak{M}, w \models \#\gamma$. ☐

Note that clauses (1) and (2) above are devised to make the neighbourhood case of the induction step straightforward.

Existence of a filtration.

Notation. $[U] = \{[u]/u \in U\}$ i.e. $[U]$ is a set of classes of equivalences. Define $N^s_{[w]}$ as follows: $[U] \in N^s_{[w]}$ iff $(\exists\, w \equiv_\Sigma w'/U \in N_{w'})$. That is, $[U]$ is a neighbourhood of $[w]$ if there exists a neighbourhood U in the original model reachable through a world w' which is equivalent to w (under $\equiv_\Sigma$). This definition leads us to the smallest filtration.

Lemma 1 (See Lemma 2.40 in [3])**.** *Let $\mathfrak{M}$ be any model, Σ any subformula closed set of formulas, W_Σ the set of equivalence classes of W induced by $\equiv_\Sigma$, V^f the standard valuation on W_Σ. Then $\langle W_\Sigma, N^s_{[w]}, V^f\rangle$ is a filtration of $\mathfrak{M}$ through Σ.*

Proof. It suffices to show that $N^s_{[w]}$ fulfills clauses (1) and (2) in Definition 4. Note that it satisfies (1) by definition. It remains to check that $N^s_{[w]}$ fulfills (2).

Let $\#\varphi \in \Sigma$, we have to prove that $(\forall \mathcal{U} \in N^s_{[w]})$ $(\forall [u] \in \mathcal{U})(\mathfrak{M}, u \models \varphi) \to (\mathfrak{M}, w \models \#\varphi)$. We know that $\mathcal{U} = [U]$ for some $U \in N_{w'}$ such that $w \equiv_\Sigma w'$. Recall that $(\forall [u] \in \mathcal{U})(\mathfrak{M}, u \models \varphi)$ means that $(\forall u \in U)(\mathfrak{M}, u \models \varphi)$. By truth definition $\mathfrak{M}, w' \models \#\varphi$, then because $w \equiv_\Sigma w'$ we get $\mathfrak{M}, w \models \#\varphi$. ☐

Theorem 2 (Finite Model Property via Filtrations)**.** *Assume that φ is satisfiable in a model $\mathfrak{M}$ as in Definition 3; take any filtration $\mathfrak{M}^f$ through the set of subformulas of φ. That φ is satisfiable in $\mathfrak{M}^f$ is immediate from the Filtration Theorem for the neighbourhood case.*

Being $\equiv_\Sigma$ an equivalence relation, and using Theorem 1 it's easy to check that, a model $\mathfrak{M}$ and any filtration $\mathfrak{M}^f$ are equivalent modulus φ. This result is useful to understand why the original properties of the frames in the models are preserved. This results are provided in [Chellas] for the preservation of frames clases through filtrations.

Example 1 (uni-agent mono-modal system)**.** A simple system can be defined with structure as in Definition 3, where we can write and test situations like the one following:

Bus stop scenario ([13], revisited). Suppose that agent y is at the bus stop. We can test whether y raises his hand and stops the bus by testing the validity of the formula: $\mathrm{Does}_y(\mathit{StopBus})$. This simple kind of systems are proved decidable via FMP through Definition 4, Theorem 1 and Lemma 1 in this Section. They are powerful enough to monitor a single agent's behaviour.

Note that $\mathrm{Does}_y(StopBus)$ holds in a world w in a model $\mathfrak{M}$, that is, $\mathfrak{M}, w \models \mathrm{Does}_y(StopBus)$ iff $(\exists U \in N_{y_w})$ such that $(\forall u \in U)\,(\mathfrak{M}, u \models StopBus)$.

3 Extension to the multi-agent multi-modal case

Recall that the original base structure discussed in [32] is a multi-relational frame of the form:

$$\mathfrak{F} = \langle A, W, \{B_i\}_{i\in A}, \{G_i\}_{i\in A}, \{I_i\}_{i\in A}, \{D_i\}_{i\in A}\rangle$$

where:

- A is a finite set of agents;
- W is a set of situations, or points, or possible worlds;
- $\{B_i\}_{i\in A}$ is a set of accessibility relations wrt Bel, which are transitive, euclidean and serial;
- $\{G_i\}_{i\in A}$ is a set of accessibility relations wrt Goal, (standard K_n semantics);
- $\{I_i\}_{i\in A}$ is a set of accessibility relations wrt Int, which are serial; and
- $\{D_i\}_{i\in A}$ is a family of sets of accessibility relations D_i wrt Does, which are pointwise closed under intersection, reflexive and serial [17].

This original structure contains the well-known normal operators Bel, Goal, and Int. They have a necessity semantics, plus characterizing axioms (see for example [19, 9]). These operators are the ones we aim to arbitrarily combine with the non-normal Does.

Note that the necessity semantics for the Kripke case can be written using neighbourhood semantics in the following way (see [6] Theorem 7.9 for more detail):

$\mathfrak{M}^K, w \models \varphi$ iff $(\forall v \,/\, wRv)(\mathfrak{M}^K, v \models \varphi) \Longleftrightarrow \mathfrak{M}^N, w \models \varphi$ iff $(\forall v_k \in N_w)\,(\forall u \in v_k)(\mathfrak{M}^N, u \models \varphi)$

where $\mathfrak{M}^K$ is a Kripke model, and $\mathfrak{M}^N$ is a neighbourhood model.

The intuition behind this definition is that each world v accessible from w in $\mathfrak{M}^K$ is a neighbourhood of w in $\mathfrak{M}^N$. Standard models can be paired one-to-one with neighbourhood models in such a way that paired models are pointwise equivalent [6].

So we can think of having a $\{N_{i_w}\}$ for each normal modality, as we do for the Does modality.

Now let us consider a multi-modal system with structure $\langle W, \{N_{1_w}\}, ..., \{N_{m_w}\}\rangle$ and let us assume that we have one agent. It is straightfoward to extend the application of Theorem 1 (Section 2) to this structure. Asume a basic modal language with modalities $\#_1, ..., \#_m$, each with a neighbourhood semantics. Also, consider a set Σ closed for subformulas that satisfies: (i) if $\varphi \vee \varphi' \in \Sigma$ then $\varphi \in \Sigma$ and $\varphi' \in \Sigma$; (ii) if $\neg\varphi \in \Sigma$, then $\varphi \in \Sigma$; and (iii) if $\#_i\,\varphi \in \Sigma$, then $\varphi \in \Sigma$ for every $\#_i$.

Definition 5 (Extends Definition 4)**.** Let $\mathfrak{M} = \langle W, \{N_{1_w}\}, ..., \{N_{m_w}\}, V\rangle$ be a model, Σ a subformula closed set, $\equiv_\Sigma$ an equivalence relation. Let $\mathfrak{M}^f = \langle W^f, \{N_{1_w}\}^f, ..., \{N_{m_w}\}^f, V^f\rangle$ such that $W^f = W_\Sigma$ and:

1. If $U \in N_{i_w}$ then $\{[u]/u \in U\} \in N^f_{i_{[w]}}$; and
2. For every formula $\#_i\,\varphi \in \Sigma$, if $\mathcal{U} \in N^f_{i_{[w]}}$ and $(\forall [u] \in \mathcal{U})(\mathfrak{M}, u \models \varphi)$, then $\mathfrak{M}, w \models \#_i\,\varphi$.
3. $V^f(p) = \{[w] \,/\, \mathfrak{M}, w \models p\}$, for all proposition letter p in Σ.

It is easy to check that if Σ is a subformula closed set of formulas, then $\mathfrak{M}^f$ is a filtration of $\mathfrak{M}$ through Σ. That is, for all φ in Σ and all w in $\mathfrak{M}$, $\mathfrak{M}, w \models \varphi$ iff $\mathfrak{M}^f, [w] \models \varphi$. Proof is done by repeated application of Theorem 1 (Section 2). Clearly, it suffices to prove the result for a single '$\#_i$' as all modalities have a neighbourhood semantics. It is worth mentioning that authors in [10], for example, proceed with the direct repeated application of the notion of filtration for proving the FMP of their (normal) multi-modal system.

Example 2 (uni-agent multi-modal system)**.** A simple system can be defined according to Definition 5, where we can depict scenarios and test situations like the one following:

Bus stop example (revisited). Agent x is at the bus stop having the goal to stop the bus: $\mathrm{Goal}_x(\mathrm{Does}_x(StopBus))$.

Note that $\mathrm{Goal}_x(\mathrm{Does}_x(StopBus))$ holds in a world w in a model $\mathfrak{M}$, that is, $\mathfrak{M}, w \models (\mathrm{Goal}_x\,\mathrm{Does}_x(StopBus))$ iff $(\exists U \in N_{x_w})$ such that $(\forall u \in U)(\mathfrak{M}, u \models \mathrm{Does}_x(StopBus))$, and $(\forall u \in U)(\mathfrak{M}, u \models \mathrm{Does}_x(StopBus))$ iff $(\exists U' \in N_{y_u})$ such that $(\forall u' \in U')(\mathfrak{M}, u' \models StopBus)$.

Further extension: multi-agent case

Extending the system to many agents will not add anything substantially new to Definition 5. A multi-agent system is a special case of the multi-modal case; the structure is merely extended with the inclusion of new modalities. For example, include Bel_i, Goal_i, and Int_i, for each agent i and a Does_i for each agent i. Thus, for every agent, include its corresponding modalities, each of which brings in its own semantics.

Example 3 (multi-agent multi-modal system)**.** A multi-agent multi-modal system for the bus stop scenario is, for example:

Bus stop example (re-revisited). The formula $\mathrm{Bel}_x(\mathrm{Does}_y(StopBus)))$ stands for 'agent x believes that agent y will stop the bus', meaning that he thinks he will not have to raise the hand himself. This formula holds in a world w in a model $\mathfrak{M}$, that is, $\mathfrak{M}, w \models \mathrm{Bel}_x\,\mathrm{Does}_y(StopBus)$ iff $(\exists U \in N_{x_w})$ such that $(\forall u \in U)(\mathfrak{M}, u \models \mathrm{Does}_y(StopBus))$, and $(\forall u \in U)(\mathfrak{M}, u \models \mathrm{Does}_y(StopBus))$ iff $(\exists U' \in N_{y_u})$ such that $(\forall u' \in U')(\mathfrak{M}, u' \models StopBus)$.

Another example.

Bus stop example (persuasion). $\mathrm{Does}_x(\mathrm{Goal}_y(StopBus))$ can be seen as a form of persuasion, meaning that 'agent x makes agent y stop the bus'. $\mathrm{Does}_x(\mathrm{Goal}_y(StopBus)))$ holds in a world w in a model $\mathfrak{M}$, that is, $\mathfrak{M}, w \models \mathrm{Does}_x\,\mathrm{Goal}_y(StopBus)$ iff $(\exists U \in N_{x_w})$ such that $(\forall u \in U)(\mathfrak{M}, u \models \mathrm{Goal}_y(StopBus))$, and $(\forall u \in U)(\mathfrak{M}, u \models \mathrm{Goal}_y(StopBus))$ iff $(\exists U' \in N_{y_u})$ such that $(\forall u' \in U')(\mathfrak{M}, u' \models StopBus)$.

Recall that we could not write and test *wff* with modalities within the scope of a Does in [32] and [31]. $\mathrm{Does}_i(\mathrm{Goal}_j\,\mathscr{A})$ is a formula in which the normal modality appears within the scope of a (non-normal) Does.

4 Combination of Mental States and Actions

Up to now, we described MAS under a single point of view: in this situation an agent believes this way, and acts that way. We are now interested in describing systems in which two points of view coexist: a cognitive one, and a behavioural one. These differ from the former ones on the ontology adopted.

We already referred in the Introduction that it is common to combine agent's behaviour with time. As a further example, a combination between a basic temporal and a simple deontic logic for MAS has been recently depicted in [33]. That combination puts together two normal modal logics: a temporal one and a deontic one. In the resultant system it is possible to write and test the validity of formulas with arbitrarily interleaved deontic and tense modalities. There are two structures (W, R) and $(T, <)$ which are respectively the underlying ontologies where a deontic point of view and a temporal point of view are interpreted (both are Kripke models). (W, R) represents a multigraph over situations, $(T, <)$ represents a valid time line. Next, it is built an ontology $W \times T$ of pairs (*situation*, *point in time*) representing the intuition "this situation, at this time". We note that such combination can be seen as a special case of the structure that we outline next. This outline (which is more general) allows combinations of non-normal operators having neighbourhood semantics.

For simplifying our presentation, we work again with the less possible number of modalities (say just two). We choose a normal, cognitive modality (let us say Bel, for beliefs), and a non-normal behavioural one (let us say Does, for agency).

Proposition 1. *If* $\langle W_B, \{N_B\}_{b \in W_B} \rangle$ *and* $\langle W_D, \{N_D\}_{d \in W_D} \rangle$ *are neighbourhood frames, then:*

$\mathfrak{C} = \langle W_B \times W_D, \{N_B\}_{(b,d) \in W_B \times W_D}, \{N_D\}_{(b,d) \in W_B \times W_D} \rangle$ *is a combined frame, where:*

- $W_B \times W_D$ *is a set of pairs of situations;*
- $S \in N_{B_{(b,d)}}$ *iff* $S = m \times \{d\}$, $m \in N_{B_b}$*; and*
- $T \in N_{D_{(b,d)}}$ *iff* $T = \{b\} \times n$, $n \in N_{D_d}$.

At a point (w_B, w_D) we have a pair of situations which are, respectively, environmental support for an internal configuration and for an external one. According to both dimensions, we test the validity of *wffs*: beliefs are tested on w_B and throughout the neigbourhoods of w_B provided by dimension S. The S dimension keeps untouched the behavioral dimension bound to w_B i.e. w_D is the second component on the neighbourhood S of w_B. (respectively for w_d and T).

In its turn, a combined model is a strucure $\langle \mathfrak{C}, V \rangle$ where V is a valuation function defined as expected. It is plain to see that this structure is an instance of Definition 5. That means there exists a flitration for a model based on this structure.

A MAS with structure as in Proposition 1 is said to be two-dimensional in the sense given by Finger and Gabbay in [14]: the alphabet of the system's language contains two disjoint sets of operators, and formulas are evaluated at a two-dimensional assignment of points that come from the prime frames' sets of situations. Moreover, in this "Beliefs $\times$ Actions" outline, there is no strong interaction among the logic of beliefs and the logic of agency as we define no interaction axioms among both special purpose logics. Our Proposition 1 much resembles the definition of full join given in [14] (Def 6.1) (two-dimensional plane).

Example 4 (Uni-agent combined system)**. Agent's beliefs and actions.** According to Proposition 1, we can define a system where to write and test formulas like e.g. $\mathrm{Bel}_x(\mathrm{Does}_x(\mathrm{Bel}_x \mathscr{A}))$. This formula is meant to stand for "agent x believes that s/he does what s/he believes" which can be seen as a kind of "positive introspection" regarding agency. This formula is not to be understood as an axiom bridging agency and beliefs; nonetheless it may be interesting to test its validity in certain circumstances: one may indeed believe that one is doing what meant to (expected correspondence between behaviour and belief), while one may believe one is doing something completely different to what one is effectively doing (e.g. poisoning a plant instead of watering it; or some other forms of erratic behaviour). Moreover, there are occasions where one performs an action which one does not believes in (e.g. obeying immoral orders).

For testing such formula, one possible movement along the multigraph is:

$\mathfrak{M}, (w_B, w_D) \models \mathrm{Bel}_x(\mathrm{Does}_x(\mathrm{Bel}_x \mathscr{A}))$ iff $(\exists\, U \in N_{B_{(w_B, w_D)}})$ such that $(\forall\, (u, w_D) \in U)$ $(\mathfrak{M}, (u, w_D) \models \mathrm{Does}_x(\mathrm{Bel}_x \mathscr{A})$. In its turn, $(\mathfrak{M}, (u, w_D) \models \mathrm{Does}_x(\mathrm{Bel}_x \mathscr{A})$ iff $(\exists\, \mathbb{V} \in N_{D_{(u, w_D)}})$ such that $(\forall\, (u, v) \in \mathbb{V})$ $(\mathfrak{M}, (u, v) \models \mathrm{Bel}_x \mathscr{A})$. Finally, $(\mathfrak{M}, (u, v) \models \mathrm{Bel}_x \mathscr{A})$ iff $(\exists\, U' \in N_{B_{(u,v)}})$ such that $(\forall\, (u', v) \in U')$ $(\mathfrak{M}, (u', v) \models \mathscr{A})$.

In connection with our Example 4, it is worth mentioning that J. Broersen defines and explains in [5] a particular logics for doing something (un)knowingly. In that work (Section 3) the author explicitly defines some constraints for the interaction between knowledge and action, namely (1) an axiom that reflects that agents can not knowingly do more than what is affected by the choices they have, and (2) an axiom establishing that if agents knowingly see to it that a condition holds in the next state, in that same state agents will recall that such condition holds. The frames used are two-dimensional, with a dimension of histories (linear timelines) and a dimension of states agents can be in. Behaviours of agents can be interpreted as trajectories going from the past to the future along the dimension of states, and jumping from sets of histories to subsets of histories (choices) along the dimension of histories.

5 Conclusions

The idea of combining special purpose logics for building "on demand" MAS is promising. This engineering approach is, in this paper, balanced with the aim to handle decidable logics, which is a basis for the implementation and launching of correct systems. We believe that the decidability issue should be a prerequisite to be taken into account during the design phase of MAS.

Within the MAS community the neighbourhood notation is, possibly, most widely used, well-understood, and well-recognized than the selection function notation. We gave a "neighbourhood outline" to decidability via filtration for a particular kind of models, namely neighbourhood models. These models are suitable for capturing the semantics of some non-normal operators found in the MAS literature (such as agency, or ability, among others) and, of course, also the semantics of normal modal operators as most MAS use.

We also offered technical details for combining logics which can be used as a basis for modeling multi-agent systems. The logics resulting from different possible combinations lead to interesting levels of expressiveness of the systems, by allowing different types of complex formulas. The combinations outlined in this paper are, given the logical tools presented in Section 2, decidable. There are for sure several other possible combinations that can be performed. For example, Proposition 1 can be extended to capture more cognitive aspects such as e.g. goals, or intentions. In that case, the cognitive dimension (In Proposition 1, characterized by S) is to be extended with the inclusion of normal operators. Moreover, within our neighbourhood outline and on top of the uni-agent modalities, collective modalities such as mutual intention, collective intention; also elaborated concepts such as trust or collective trust can also be defined.

We can push the combination strategy even further, by proposing the combination of modules which are in its turn combinations of special purpose logics, in a kind of multiple level combination. This strategy has to be carefully studied, and is matter of our future research.

REFERENCES

[1] C. Areces, C. Monz, H. de Nivelle, and M. de Rijke, 'The guarded fragment: Ins and outs', in *JFAK. Essays Dedicated to Johan van Benthem on the Occasion of his 50th Birthday*, eds., J. Gerbrandy, M. Marx, M. de Rijke, and Y. Venema, Vossiuspers, AUP, (1999).

[2] Patrick Blackburn, Johan F. A. K. van Benthem, and Frank Wolter, *Handbook of Modal Logic, Volume 3 (Studies in Logic and Practical Reasoning)*, Elsevier Science Inc., New York, NY, USA, 2006.

[3] Patrick Blackburn, Maarten de Rijke, and Yde Venema, *Modal Logic*, volume 53 of *Cambridge Tracts in Theoretical Computer Science*, Cambridge University Press, Cambridge, 2001.

[4] Jan Broersen, 'Relativized action complement for dynamic logics', in *Advances in Modal Logic*, pp. 51–70, (2002).

[5] Jan Broersen, 'A complete stit logic for knowledge and action, and some of its applications', in *DALT*, pp. 47–59, (2008).

[6] Brian Chellas, *Modal Logic: An Introduction*, Cambridge University Press, 1980.

[7] Philip R. Cohen and Hector J. Levesque, 'Intention is choice with commitment', *Artif. Intell.*, **42**(2-3), 213–261, (March 1990).

[8] J. W. de Bakker, Willem P. de Roever, and Grzegorz Rozenberg, eds. *Linear Time, Branching Time and Partial Order in Logics and Models for Concurrency, School/Workshop, Noordwijkerhout, The Netherlands, May 30 - June 3, 1988, Proceedings*, volume 354 of *Lecture Notes in Computer Science*. Springer, 1989.

[9] Barbara Dunin-Keplicz and Rineke Verbrugge, 'Collective intentions.', *Fundam. Inform.*, 271–295, (2002).

[10] Marcin Dziubiński, Rineke Verbrugge, and Barbara Dunin-Kęplicz, 'Complexity issues in multiagent logics', *Fundam. Inf.*, **75**(1-4), 239–262, (January 2007).

[11] Dag Elgesem, 'The modal logic of agency', *Nordic Journal of Philosophical Logic*, **2**, 1–46, (1997).

[12] Rogerio Fajardo and Marcelo Finger, 'Non-normal modalisation.', in *Advances in Modal Logic'02*, pp. 83–96, (2002).

[13] R. Falcone and C. Castelfranchi, *Trust and Deception in Virtual Societies*, chapter Social Trust: A Cognitive Approach, 55–90, Kluwer Academic Publishers, 2001.

[14] Marcelo Finger and Dov Gabbay, 'Combining temporal logic systems', *Notre Dame Journal of Formal Logic*, **37**, (1996).

[15] Massimo Franceschet, Angelo Montanari, and Maarten De Rijke, 'Model checking for combined logics with an application to mobile systems', *Automated Software Eng.*, **11**, 289–321, (June 2004).

[16] Dov Gabbay, *Fibring Logics*, volume 38 of *Oxford Logic Guides*, Oxford Univesity Press, 1998.

[17] Guido Governatori and Antonino Rotolo, 'On the Axiomatization of Elgesem's Logic of Agency and Ability', *Journal of Philosophical Logic*, **34**(4), 403–431, (2005).

[18] Guido Governatori and Antonino Rotolo, 'Norm compliance in business process modeling', in *Semantic Web Rules*, eds., Mike Dean, John Hall, Antonino Rotolo, and Said Tabet, volume 6403 of *Lecture Notes in Computer Science*, 194–209, Springer Berlin / Heidelberg, (2010).

[19] J. Halpern and Y. Moses, 'A guide to completeness and complexity for modal logics of knowledge and belief', *Artificial Intelligence*, **54**, 311–379, (1992).

[20] Bengt Hansson and Peter Gärdenfors, 'A guide to intensional semantics. Modality, Morality and Other Problems of Sense and Nonsense. Essays Dedicated to Sören Halldén', (1973).

[21] Philip R. Cohen Hector J. Levesque and Jose H. T. Nunes, 'On acting together', Technical Report 485, AI Center, SRI International, 333 Ravenswood Ave., Menlo Park, CA 94025, (May 1990).

[22] H. Herrestad and C. Krogh, *Deontic Logic Relativised to Bearers and Counterparties*, 453–522, J. Bing and O. Torvund, 1995.

[23] A.J.I. Jones and M. Sergot, 'A logical framework. in open agent societies: Normative specifications in multi-agent systems', (2007).

[24] Dexter Kozen, 'Results on the propositional mu-calculus', *Theor. Comput. Sci.*, **27**, 333–354, (1983).

[25] J. J. Ch Meyer, 'A different approach to deontic logic: Deontic logic viewed as a variant of dynamic logic', *Notre Dame Journal of Formal Logic*, **29**(1), 109–136, (1987).

[26] Jeremy Pitt, 'A www interface to a theorem prover for modal logic', in *Department of Computer Science, University of York*, pp. 83–90, (1996).

[27] ANAND S. RAO and MICHAEL P. GEORGEFF, 'Decision procedures for bdi logics', *Journal of Logic and Computation*, **8**(3), 293–343, (1998).

[28] Antonino Rotolo, Guido Boella, Guido Governatori, Joris Hulstijn, Regis Riveret, and Leendert van der Torre, 'Time and defeasibility in FIPA ACL semantics', in *Proceedings of WLIAMAS 2008*. IEEE, (2008).

[29] Klaus Schild, 'On the relationship between bdi logics and standard logics of concurrency', *Autonomous Agents and Multi-Agent Systems*, **3**(3), 259–283, (September 2000).

[30] Peter K. Schotch, 'Paraconsistent logic: The view from the right', in *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, volume Two: Symposia and Invited Papers, pp. 421–429. The University of Chicago Press, (1992).

[31] Clara Smith, Agustin Ambrossio, Leandro Mendoza, and Antonino Rotolo, 'Combinations of normal and non-normal modal logics for modeling collective trust in normative mas', *AICOL XXV IVR, Forthcoming LNAI*, (2011).

[32] Clara Smith and Antonino Rotolo, 'Collective trust and normative agents', *Logic Journal of IGPL*, **18**(1), 195–213, (2010).

[33] Clara Smith, Antonino Rotolo, and Giovanni Sartor, 'Representations of time within normative MAS', *Frontiers in Artificial Intelligence and Applications*, **223**, 107–116, (2010).