Decidability via Filtration of Neighbourhood Models for Multi-Agent Systems

Clara Smith$^{12}$ Agustín Ambrossio$^1$ Leandro Mendoza$^{23}$

(1) FACET, UCALP, Argentina

(2) Facultad de Informática, UNLP, Argentina

(3) CONICET, Argentina

SNAMAS 2012

July 2, 2012
1 General Overview
  - General Overview
  - Motivation and Previous Work

2 Background
  - Normality vs. Non-Normality
  - Combination of Logics

3 Neighbourhood Semantics
  - Decidability and FMP
  - Neighbourhood Approach

4 Our Combination
  - Definition

5 Summary
  - Summary
We focus on the combination of special-purpose logics for building “on demand” MAS.

Most used normal logics for modeling agents’ cognitive states are: logics for beliefs, goals and intentions. Perhaps, the most well-known non-normal modal logic for MAS is the logic of agency (and, possibly, ability).
Motivation and Previous Work

- We organized a given multi-modal MAS for modeling Collective Trust as a fibring [SAMR2012].
- A fibring is a particular combination of logics which intuitively amounts to place one logics on top of another.
- In this case, the normal logic was put on top of the non-normal one. Before, we obtained two restrictions of the original logics.
- We proved that the system is complete and decidable.
- We outlined a sketch for an appropriate model checker.

![Diagram showing models for L and M]
Normality vs. Non-Normality

In a mono-modal logic with a box modality, normality implies that the following formulas are valid:

\[ \Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q), \text{ and} \]

\[ \Box(p \land q) \rightarrow (\Box p \land \Box q) \]

as well as the admission of the rule:

\[ \text{from } \vdash \mathcal{A} \text{ infer } \vdash \Box \mathcal{A}. \]

Some or None of these is assumed to hold for a non-normal logics [Chellas80].

We use a non-normal modal logic for agency, and aim to put it to work with normal logics for beliefs and goals.
Advantages regarding the choice of a logic of agency such as Does [Elgesem 1997]

- Action negation: for Does, and for other related logics of action (e.g. ability), action negation is well understood.
- The logic for Does is Boolean; it is easy to determine what $\neg \text{Does } A$ means.
- “refrain” in terms of Does: “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, there are no outstanding non homogeneous solutions for the issue on action negation in other relevant logics for MAS, such as dynamic logics.
Neighbourhood Semantics are a generalization of Kripke Semantics. [Hansson and Gärdenfors]

A *neighbourhood frame* is a tuple $\langle W, \{N_w\}_{w \in W}\rangle$ where:
- $W$ is a set of worlds, and
- $\{N_w\}_{w \in W}$ is a function assigning to each element $w$ in $W$ a class of subsets of $W$, the neighbourhoods of $w$.

In its turn, $\mathcal{M} = \langle W, \{N_w\}_{w \in W}, V\rangle$ is a neighbourhood model, where:
- $\langle W, \{N_w\}_{w \in W}\rangle$ is a neighbourhood frame, and
- $V$ is a valuation function assigning to each proposition letter $p$ a subset $V(p)$ of $W$ (i.e, for every propositional letter, we know in which worlds it is true).
Combinations of Logics for MAS

Cognitive modalities are usually extended with *time*:

- Schild maps Rao and Georgeff’s BDI to $\mu$-calculus, collapsing the (original) two dimensions of time and modalities onto a one dimensional structure. [Schild 2000]
- J. Broersen presents an epistemic logic that incorporates interactions between time and action, and between knowledge and action. [Broersen 2008]
- H. Wansing 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 future course of events.
- L F. del Cerro and A. Herzig combine classical and intuitionistic logics, A. Herzig also combines beliefs and databases...and much more.
Combinations of *Bel, Int, Goal, Does*

- We aim to have an expressive enough system for modeling interactions between a behavioral dimension and a cognitive dimension of agents.
- For example: $\text{Does}_i (\text{Bel}_j \ A)$ is a form of persuasion. This formula cannot be written in the fibred language in [SAMR2012] because *no modal operator can appear in the scope of a Does*.
- We devised a join of logics that lead to an ontology of pairs of situations
  
  $$(w_N, w_D)$$

  where to test behavior and visible actions, respectively.
Extension of the *FMP* to Neighbourhoods.

- Normal logics can be seen as a platform for the study of transference of decidability results for non-normal logics and combination of logics.
- There are well-studied results and existing techniques for Kripke structures, which are usual support of normal logics.
- We provide a new presentation of existing decidability results for a *more general class of structures* supporting non-normal logics.
- B. Chellas defines filtrations for minimal models (generalization of Kripke ones). The is based on *truth sets* ($\| A \|$ is the set of points in a model where $A$ is true). Truth sets are a basic ingredient of selection function semantics. **We will concentrate on neighbourhoods.**
P. Schotch has already addressed the issue of paradigmatic notation and dominating semantics for modalities: [Schotch 1992]

- He points out that the necessity truth condition (box) 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.

- 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: a structure that allows non-normal semantics go through it with the notation as given e.g. by Blackburn et al., which is currently well-accepted and well understood for modal MAS.
Choosing a Neighbourhood Approach

- It focuses on *worlds*, which are abstract descriptions of external circumstances of an agent’s community that allow or disallow actions, activate or nullify goals (no ‘special’ worlds).
- This is crucial in current practical approaches: in a world an agent realizes its possibilities of successful agency of $\mathcal{A}$, its beliefs, its goals, *all relative to the actual world* $w$. Situations are an “environmental support” for agent’s internal configuration and visible actions.
- Neighbourhoods better adapt to the specification of most prevailing modal MAS, which lately tend to adopt Kripke semantics.
- This because, probably, that notation is more intuitive for dealing with situations and agents acting and thinking according to situations, rather than considering the formulas as ‘first class’ objects.
Finding a filtration of a neighbourhood model

- A filtration for a neighbourhood model: is a finite model $M^f = \langle W^f, \{N_w\}^f, V^f \rangle$ which satisfies certain properties.
- We prove the Filtration Theorem for the neighbourhood case: given a model and a filtration, filtration preserves satisfiability of formulas,
- We prove that the (smallest) neighbourhood filtration exists,
- We prove the FMP via filtrations for the neighbourhood case: assume that $\varphi$ is satisfiable in a model $M$, take any filtration $M^f$ of $M$, that $\varphi$ is satisfiable in $M^f$ is true, given the Filtration Theorem.
A Simple, decidable system with neighbourhood semantics:

- Bus stop scenario (monitoring system). 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 $\text{Does}_y(\text{StopBus})$.

$\text{Does}_y(\text{StopBus})$ holds in a world $w$ in a model $\mathcal{M}$, that is: $\mathcal{M}, w \models \text{Does}_y(\text{StopBus})$ iff $(\exists U \in N_w)$ such that $(\forall u \in U) (\mathcal{M}, u \models \text{StopBus})$. 
MAS with behavioral and agency modalities

- Necessity Kripkean semantics fit within neighborhoods [REF]. Intuition: each accessible world \( v \) from \( w \) in \( M^K \) is a neighborhood of \( w \) in \( M^N \).
- So, we can think of a \( \{N_w\} \) for each normal modality (first, extend the previous theorems).
- Bus stop example (persuasion).

\[ M, w \models Does_x \text{Goal}_y(StopBus) \iff (\exists U \in N_{xw}) \text{ such that } \forall u \in U (M, u \models \text{Goal}_y(StopBus)) \iff (\exists U' \in N_{yw}) \text{ such that } \forall u' \in U' (M, u' \models \text{StopBus}). \]
An ontology for testing internal configuration and visible behaviour

Definition:

If \( \langle W_B, \{ N_B \}_{b \in N_B} \rangle \) and \( \langle W_D, \{ N_D \}_{d \in N_D} \rangle \) are neighborhood frames, then:

\[
\mathcal{C} = \langle W_B \times W_D, \{ N_B \}_{(b,d) \in N_D \times N_B}, \{ N_B \}_{(b,d) \in N_D \times N_B} \rangle
\]

is a combined frame, where:

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

\( S \) and \( T \) are respectively the behavioral and the agency dimensions; i.e. environmental support for testing wffs.
Agents’ combined beliefs and actions

Example:

\[ \text{Bel}_x(\text{Does}_x(\text{Bel}_x \ A)) \] can be seen as “positive introspection” regarding agency.

It may be interesting to test it 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 behavior).
- occasions where one performs an action which one does not believes in (e.g. obeying immoral orders).
Engineering approach: combining special purpose logics for building “on demand” MAS.

Combinations can be used as a basis for modeling systems with normal and non-normal operators.

Within the MAS community, the neighborhood notation is, possibly, better understood than the selection function notation for dealing with non-normal logics.

We give a “neighbourhood outline” to decidability via filtration for neighbourhood models. They are suitable for capturing the semantics of some non-normal operators found in the MAS literature (agency, or ability).
Thanks!!
Thanks!!

Questions?