Indexed Semantics and Its Application in Modelling Interactive Unawareness

MSc Thesis (*Afstudeerscriptie*)

written by

Yanjing Wang
(born March 3rd, 1982 in Beijing, China)

under the supervision of **Prof. Dr. Frank Veltman** and **Dr. Maricarmen Martinez**, and submitted to the Board of Examiners in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:
March 8th, 2006

Members of the Thesis Committee:
Prof. Dr Frank Veltman
Prof. Dr. Johan van Benthem
Lector. Dr. Peter van Emde Boas

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Acknowledgments

First of all, I gratefully acknowledge the Beth Foundation for the scholarship I received, without which I would not have been able to come to ILLC.

I would like to thank my supervisors Frank Veltman and Maricarmen Martinez who encouraged me to find, sharpen and solve my own questions. I appreciate their patience for listening to my "gases in the air" and the efforts to get my feet back on the ground at the last moment.

Special thanks to Floris Roelofsen and Fenrong Liu, the friendship with them teaches me research, love and help me to find my way of life. I am also grateful to my mentor Khalil Sima'an who was willing to help me with all kinds of practical problems. Many thanks to those people in the Logic & Game reading group and the Stanford student visitors especially Tomasz Sadzik, I was inspired a lot by them.

Last but not least, I'd like to thank my parents for their unconditional support as always.

Contents

Chapter 1

Introduction

Chuangtse and Hueitse had strolled on to the bridge over the Hao, when the former observed, "See how the small fish are darting about! That is the happiness of the fish."

"You not being a fish yourself," said Huei, "how can you know the happiness of the fish?"

"And you not being I," retorted Chuangtse, "how can you know that I do not know?"

"If I, not being you, cannot know what you know," urged Huei, "it follows that you, not being a fish, cannot know the happiness of the fish."

"Let us go back to your original question," said Chuangtse. "You asked me how I knew the happiness of the fish. Your very question shows that you knew that I knew. I knew it from my own feelings on this bridge."

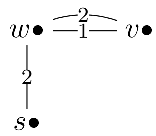
–Chuangtse, 300 B.C

1.1 Motivation

Possible-worlds models and the corresponding Kripke semantics have been used extensively in the literature about epistemic/doxastic logic in the last few decades. An example of a typical possible-worlds model for multi-agents[1]

[1] $S5_m$ model, we omitted the arrows.

is as follows:



Here the points stand for the possible worlds agents can be in; the labelled relations among them represent the possibility relations agents may consider. The status of a possible world is determined by the primitive propositions that hold on it and the structures reachable from it. The combination of possible-worlds models and the associated Kripke semantics enjoys the great advantages of both simplicity and expressive power. Numerous logic systems based on them have been studied. However, there are some fundamental problems which deserve more discussion.

1. Problem 1 on model building.

   - Problem: Given a consistent set of formulas, there are tons of classical possible-worlds models which satisfy those formulas on some points. Clearly, not all of them are desired. One reasonable criterion for a "right model" might be to require the model has as few "side effects" as possible(in terms of the extra formulas that hold but are not the logical consequences of the given set). Another criterion might be about complexity of the model–the simpler the better. Then the question is: do we have a way to build models step by step according to these two criterions? It is easy to see that to build the classical possible-world models with the fewest worlds usually won't do the job. We normally get lots of unwanted formulas which are satisfiable at the same point.
   - Example. Consider such a situation in which agent 1 thinks both p and $\neg p$ possible, no matter what the real world is. And so does agent 2. The most intuitive model for agent 1 is as follows[2]:

     $$p \text{ —}_1\text{— } \neg p$$

     Now we add agent 2 in the simplest way:

[2]The reflexive arrows are omitted.

p —1,2– $\neg p$

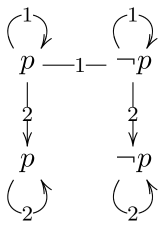
It is clear that $\Box_1 \neg \Box_2 \neg p$ holds on every world. However, this is undesired, since there is no constraint about how agent 1 thinks agent 2's information state. The above observation shows that it is not "safe" to add the relation 2 into the first model in the simplest way.

- Cause: In the classical possible-worlds model, the status of a world is determined also by all the structures reachable from that world. If we add relations to a world then actually the world changes. Suppose a world is considered possible by agent 1 and there is a 2-relation from that world, then agent 1 automatically considers that 2-relation.
- Possible Solutions:
  1. Use information structures instead, which constructs agents' high order information recursively(see [FHV91]). The price is to lose the beauty of simplicity of possible-worlds models.
  2. Try to avoid the undesired compositional side-effects of the relations. The most ideal way to build models is that we fix the worlds first and add relations or worlds whenever needed.

2. Problem 2 on Generalization and Uniform Substitution rules.

   - Problem: In the classical set-up, Generalization(from ϕ, prove $\Box\phi$) and Uniform Substitution (from $\phi(p)$, prove $\phi(\psi)$ where $\phi(\psi)$ is obtained by uniformly substituting ψ for p) preserve validity on any class of frames. It follows that they should be included in every complete normal modal logic. However, we may not be happy with these rules all the time.
   - Example: It is reasonable to have a logic with positive introspection $\Box_i p \to \Box_i \Box_i p$ as an axiom for each agent i, but at the same time, not want $\Box_j(\Box_i p \to \Box_i \Box_i p)$ to be a theorem for any agent $j \neq i$, since agent j may not believe agent i is positive introspective. Similarly, we may want a logic with the axiom $\Box_i p \to p$ but without $\Box_i \phi \to \phi$ for an arbitrary formula ϕ as a theorem. It is possible for agents only have true beliefs about propositions concerning only primitive facts in any case, but hardly know anything about more complex facts concerning other agents, due to the lack of information for others.

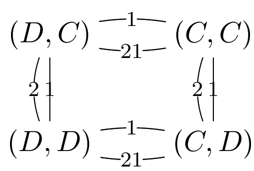
- Cause: Classical logic can only talk about the properties for all the worlds while we may care only about part of the worlds.
- Possible Solution:
  If we can restrict validity in some way on part of the whole worlds. Consider the following model:

  Clearly, the above model doesn't validate the axiom $\neg\Box_1\bot$ which is normally regarded important in modelling beliefs. However, if we only care about the above two worlds, then the serial axiom holds on both, while $\Box_2\neg\Box_1\bot$ does not. This is reasonable since the bottom two worlds only exist in the imaginations of agent 2, when he is at the top two worlds. And it could be the case that in agent 2's imaginations there is no such an agent 1.

In this thesis we will give an alternative semantics for multi agent doxastic/epistemic logic and try to solve the above problems. The basic idea of our approach is based on an simple intuition: when thinking about the others' information state, an agent is actually considering the "imaginary agents" in his mind, who may differ from the real agents in the possibility relations they have. Our trick is to include explicitly all the possibility relations for each imaginary agent in the so called "indexed models". Let's look at an example:

Example 1.1.1 Consider a variant of 2-person Muddy Children scenario. There are two children who may have mud on their foreheads. Normally, the children can only see the other one's forehead. However, child 2 is actually blind, therefore he can not get any information by looking at child 1. Unfortunately, child 1 has no ideas about that. Let's build a model according to this scenario as follows:



where there are 4 primitive possibilities: $(D, C), (D, D), (C, D)$ and (C, C).[3] Child 1 thinks all of the them are possible since in any case he can not see anything, while child 2 can not distinguish (D, C) from (D, D) and (C, D) from (C, C). These are represented by 1-relations and 2-relations respectively. Moreover, child 2 falsely believes that child 1 is normal, so he would think child 1 can not distinguish (D, C) from (C, C) and (C, D) from (D, D). This is represented by the 21-relation in the model[4].

It is easy to see that we have got rid of the compositional side-effects by setting relations for different agents independently. For example, the relation 1 will not be considered by child 2, since child 2's imagination about child 1(or say, the imaginary agent 21) is represented by relation 21.

We will develop an indexed semantics based on the indexed models which interprets the nested-modality formulas context-dependently, in order to capture the intuition about the imaginary agents. For example, the modal operator $\Box_i$ in the formula $\Box_j\Box_i\phi$ corresponds to the relations for the imaginary agent ji, and $\Box_i$ in formula $\Box_k\Box_i\phi$ corresponds to the relations for imaginary agent ki. As we will see in the next chapter, our approach will solve problem 2.

1.2 Overview of the thesis

The rest of this thesis is organized as follows. In Chapter 2, we first give the formal definition of indexed models and develop the intended indexed semantics. Then we discuss the corresponding results between this approach and the classical set-up. Finally we give several complete logics. Chapter 3 discusses the non-redundancy criterions of indexed models and give a simple application of it in modelling interactive unawareness. The final chapter concludes our results and discusses the possible further developments.

[3]For example, (D,C) means child 1 is dirty while child 2 is clean.

[4]We omit the higher order imaginations in the model.

Chapter 2

Indexed Models and Indexed Semantics

2.1 Indexed Model

According to the convention in doxastic logic and epistemic logic, we use the following language to talk about the factual and higher-order information that the agents in a non-empty group I have:

Definition 2.1.1 ***(Language*** $\mathcal{L}^I(\Phi)$***)*** *The formulas of* $\mathcal{L}^I(\Phi)$ *are formed based on a set of proposition letters* Φ *as follows:*

$$\top \quad | \quad p \quad | \quad \phi \wedge \psi \quad | \quad \neg\phi \quad | \quad \Box_j \phi$$

where $p \in \Phi$ *and* $j \in I$*.*

Notation As usual, we define $\bot$, $\phi \vee \psi$, $\phi \to \psi$ and $\Diamond_j \phi$ as the abbreviations of $\neg\top$, $\neg(\neg\phi \wedge \neg\psi)$, $\neg\phi \vee \psi$ and $\neg\Box_j\neg\phi$ respectively. Let $S(I)$ be the set of all non-empty finite sequences of elements in I. We use $\Box_c \phi$ as the abbreviation for $\Box_{j_1}\Box_{j_2}...\Box_{j_n}\phi$, where $c = \langle j_1, ..., j_n\rangle \in S(I)$. Similarly $\Diamond_c \phi$ is the abbreviation for $\Diamond_{j_1}\Diamond_{j_2}...\Diamond_{j_n}$. Especially, we denote $\Diamond_\epsilon \phi$ and $\Box_\epsilon \phi$ as two alternative forms of ϕ where ϵ is the empty sequence. In the following we assume that Φ is finite.

Remark 2.1.2 *Depending on the purpose,* $\Box_i \phi$ *will sometimes be read as "agent i believes* ϕ*" and at other times as "agent i knows* ϕ*."*

Definition 2.1.3 ***(Classical Relational Frame and Model)*** *A classical relational frame for language $\mathcal{L}^I(\Phi)$ is a pair:*

$$\mathfrak{F} = (W, \{R_i\}_{i \in I})$$

- *W is a non-empty set of possible worlds.*
- *$R_i \subseteq W \times W$ for each $i \in I$.*

A classical relational model M is a pair $(\mathfrak{F}, V)$, where $V : W \to \mathcal{P}(\Phi)$ is called a valuation on W.

Remark 2.1.4 *Intuitively, $(w, v) \in R_i$ means that when agent i is at world w, he actually thinks world v could be one possible candidate for the real world w. Moreover, v is actually part of w, in the sense that whenever any agent thinks w possible, he would also think agent i considers v possible at w. In such case, the relation $(w, v) \in R_i$ also means in any agents' imagination agent i would think v possible when at w.*

Definition 2.1.5 ***(Indexed Relational Frame and Model)*** *An indexed relational frame for language $\mathcal{L}^I(\Phi)$ is a pair:*

$$\mathfrak{F} = (W, \{R_c\}_{c \in S(I)})$$

- *W is a non-empty set of possible worlds.*
- *$S(I)$ is the set of all the non-empty finite sequences of agents in I.*
- *$R_c \subseteq W \times W$ for each $c \in S(I)$.*

An indexed relational model M is a pair $(\mathfrak{F}, V)$ where $V : W \to \mathcal{P}(\Phi)$ is a valuation on W.

Remark 2.1.6

- *Slightly different from the intuition behind the possibility relations in classical models, $(w, v) \in R_{di}$ here means that agent d(or imaginary agent d when $d \notin I$) thinks agent i would consider world v possible when agent i is at w. Especially, when d is ϵ, $(w, v) \in R_{\epsilon i} = R_i$ means that at world w agent i actually thinks world v possible, which is the same as in classical models.Notice that, R_{di} in indexed models has only one interpretation, it belongs to agent d's imagination about agent i, not in any other agent's imagination as the relations in classical models.*

- *It is easy to see that any indexed relational model* $M = (W, \{R_c\}_{c \in S(I)}, V)$ *can be looked as a classical relational model* $M' = (W, \{R_c\}_{c \in S(I)}, V)$ *for the language* $\mathcal{L}^{S(I)}(\Phi)$.

Definition 2.1.7 ***(Pointed model)*** *A pointed classical/indexed relational model is a classical/indexed relational model* M *with a specific point* $w \in W$. *We denote it as the pair* (M, w). *Usually the specific world* w *plays the role of the real world if* (M, w) *is a model for a situation.*

Notation In the following, we will call the classical relational models for language $\mathcal{L}^I(\Phi)$ "classical models" for short. Similarly, we call indexed relational models for language $\mathcal{L}^I(\Phi)$ "indexed models".

Definition 2.1.8 ***(c-path)***

- *Given a classical model* M, *suppose* $c = \langle j_1, ..., j_n \rangle \in S(I)$, *a* $c-path$ *in* M *from* w *to* v *is a tuple* $\langle w_0, w_1, ..., w_n, c \rangle$ *where* $\langle w_0, w_1, ..., w_{n-1}, w_n \rangle$ *is a sequence of possible worlds in* M*,such that* $w_0 = w$, $w_n = v$ *and* $w_{i-1} R_{j_i} w_i$ *for* $i \in \{1, .., n\}$.

- *Given an indexed model* M, *suppose* $c = \langle j_1, ..., j_n \rangle \in S(I)$, *a* $c - path$ *in* M *from* w *to* v *is a tuple* $\langle w_0, w_1, ..., w_n, c \rangle$ *where* $\langle w_0, w_1, ..., w_{n-1}, w_n \rangle$ *is a sequence of possible worlds in* W *such that* $w_0 = w$, $w_n = v$ *and* $w_{i-1} R_{j_1 \dots j_i} w_i$ *for* $i \in \{1, .., n\}$.

- *For any* $c, d \in S(I)$, *we say a* $cd-path$ $\langle w_0, ..., w_n, cd \rangle$ *is an d-extension of a* $c - path$ $\langle v_0, ..., v_m, c \rangle$, *if* $\langle v_0, ..., v_m \rangle$ *is the initial segment of the* $\langle w_0, ..., w_n \rangle$.

Notation $\epsilon - path$ is the empty path, denoted as $\langle w, \epsilon \rangle$. In the following, for any $c \in S(I)$, we consider the sequences ϵc or $c \epsilon$ as c itself. We normally use c, d as sequences in $S(I)$, and i, j as agents in I.

2.2 Indexed Semantics

We now define two satisfiability relations $\Vdash, \vDash$ for the language $\mathcal{L}^I(\Phi)$ based on classical models and indexed models respectively.

Definition 2.2.1 ***(Truth Condition for*** $\Vdash$***)*** *The truth conditions* $\Vdash$ *for the* $\mathcal{L}^I(\Phi)$ *formulas are defined recursively as below:*

$M, s \Vdash \top$	$\iff$	*always*
$M, s \Vdash p$	$\iff$	$p \in V(s)$
$M, s \Vdash \neg\phi$	$\iff$	$M, s \nVdash \phi$
$M, s \Vdash \phi \wedge \psi$	$\iff$	$M, s \Vdash \phi$ *and* $M, s \Vdash \psi$
$M, s \Vdash \Box_i\phi$	$\iff$	*for all* t, *if* sR_it *then* $M, t \Vdash \phi$

We call the above semantics Classical Semantics.

Definition 2.2.2 ***(Truth Condition for*** $\vDash$***)*** *Let* $c \in S(I) \cup \{\epsilon\}$, *the truth conditions for* $\mathcal{L}^I(\Phi)$ *formulas are defined recursively by* $\vDash_c$ *as below:*

$M, s \vDash \phi$	$\iff$	$M, s \vDash_\epsilon \phi$
$M, s \vDash_c \top$	$\iff$	*always*
$M, s \vDash_c p$	$\iff$	$p \in V(s)$
$M, s \vDash_c \neg\phi$	$\iff$	$M, s \nvDash_c \phi$
$M, s \vDash_c \phi \wedge \psi$	$\iff$	$M, s \vDash_c \phi$ *and* $M, s \vDash_c \psi$
$M, s \vDash_c \Box_i\phi$	$\iff$	*for all* t, *if* $sR_{ci}t$ *then* $M, t \vDash_{ci} \phi$

We call the above semantics as Indexed Semantics.

Remark 2.2.3

- $\vDash_c$ *is used to encode the context and thus to define* $\vDash$ *recursively. For example:*
  $M, w \vDash \Diamond_i(p \wedge \Diamond_j q)$
  $\iff$ *there is a* $v \in W, wR_iv$ *and* $M, v \vDash_i (p \wedge \Diamond_j q)$
  $\iff$ *there is a* $v \in W, wR_iv$ *and* $M, v \vDash_i p$ *and* $M, v \vDash_i \Diamond_j q$
  $\iff$ *there is a* $v \in W$ *such that* wR_iv, $M, v \vDash p$, *and there is a* $t \in W$: $vR_{ij}t$ *and* $M, t \vDash q$.

- *The above truth conditions represent an explicit context-dependent feature of the indexed semantics, in the sense that the meaning of a modal operator is explicitly determined by its position in the formula. For example, the "John" in the sentence "I believe that John believes in God" is my certain imaginary agent who may differ from the real John. Then the meaning of "John believes in God" in that sentence is different from the one in "Mary believes that John believes in God".*

- *The context only matters for modal operators. We assume that all the agents have the same understanding towards the factual propositions, namely the boolean combinations of proposition letters.*

Definition 2.2.4 ***(Indexed Semantic Consequence)*** *Let $\Gamma \cup \{\phi\} \in \mathcal{L}^I$, K be a class of indexed frames then we say that ϕ is a semantic consequence of Γ over K(notation: $\Gamma \vDash \phi$) if for all models M based on the frames in K, and all points in M, if $M, w \vDash \Gamma$ then $M, w \vDash \phi$.*

Here are some straightforward propositions from the above truth conditions:

Proposition 2.2.5

- *Given any classical model M', $w' \in W'$, $M', w' \Vdash \Box_c \psi \iff$ for all $v' \in W$ if there is a $c-path$ from w' to v' in M', then $M', v' \Vdash \psi$.*
- *Given any indexed model M, $w \in W$, $M, w \vDash \Box_c \psi \iff$ for all $v \in W$ if there is a $c-path$ from w to v in M, then $M, v \vDash_c \psi$.*

Proof. Trivial.

QED

Similarly we have the following propositions about $\Diamond_c$:

Proposition 2.2.6

- *Given any classical model M', a $w' \in M'$: $M', w' \Vdash \Diamond_c \psi \iff$ there is a $c-path$ from w to v where $v \in W'$ and $M', v \Vdash \psi$.*
- *Given any indexed model M, a $w \in M$: $M, w \vDash \Diamond_c \psi \iff$ there is a $c-path$ from w to v where $v \in W$ and $M, v \vDash_c \psi$.*

It is clear that:

Proposition 2.2.7

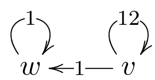
- *Given any classical model M', a $w' \in M'$, $M', w' \vDash \Diamond_c \psi \leftrightarrow \neg \Box_c \neg \psi$.*

- *Given any indexed model M, a $w \in M$, $M, w \vDash \Diamond_c \psi \leftrightarrow \neg \Box_c \neg \psi$.*[1]

Based on the above results, it is safe to denote $\neg \Box_c \neg$ as $\Diamond_c$.

2.3 Pruned Model

Notice that, in an arbitrary indexed model/frame, some relations may not be reachable by any path from any point. For example, consider the following frame:



It is clear that the above 12−relation is unnecessary at all in verifying any $\mathcal{L}^I(\Phi)$ formula ϕ, since there is no 1-path to v. Such models are also not reasonable intuitively. Consider the above one, how can agent 1 think about agent 2's possible uncertainties at v, if agent 1 never thinks v possible? In general, If there is a relation ci from w to v then there should be a path c from some world to w otherwise that ci-relation is nonsense. We hereby define the model without such unnecessary relations.

Definition 2.3.1 ***(Pruned indexed frame/model)*** *A pruned indexed frame/model is an indexed frame/model which satisfies the following condition:*

For any $w, v \in W, c \in S(I), j \in I$ if $wR_{cj}v$ then there is a $t \in W$ such that $tR_c w$.

Definition 2.3.2 ***(Pruned pointed indexed model)*** *A pruned pointed model is a pointed indexed model (M, w) which satisfies the following condition:*

For any $v, s \in W, c \in S(I), j \in I$ if $vR_{cj}s$ then there is a c−path from w to v.

Notation In the following, we say that $c \in S(I) \cup \{\epsilon\}$ *fits* $w \in W$ if $c = \epsilon$ or there is a c−path from some world to w in M.

[1] Actually we can prove the stronger version $M, w \vDash_d \Diamond_c \psi \leftrightarrow \neg \Box_c \neg \psi$ for any $c \in S(I), d \in S(I) \cup \{\epsilon\}$.

Notation Given two indexed models M, M' and $w \in M$, $w' \in M'$, we use $M, w \leftrightsquigarrow_c M', w'$(or $w \leftrightsquigarrow_c w'$ if the models are clear) to denote that for any $\mathcal{L}^I(\Phi)$ formula ϕ :

$$M, w \vDash_c \phi \iff M', w' \vDash_c \phi.$$

Especially let $w \leftrightsquigarrow w'$ be $w \leftrightsquigarrow_\epsilon w'$.

From the definition of pruned models and the truth conditions of indexed semantics, we have some immediate propositions.

Proposition 2.3.3 *Given an arbitrary indexed model $M = (W, \{R_c\}_{c \in S(I)}, V)$, there is a sub-model of it:$M' = (W', \{R'_c\}_{c \in S(I)}, V')$ which is a pruned model such that $W = W'$ and for any $w \in W : (M, w) \leftrightsquigarrow (M', w)$.*

Proof. We build the sub-model M' by throwing away unnecessary relations. Formally, $M' = (W, \{R'_c\}_{c \in S(I)}, V)$ where $R'_{di} = \{(w, v) | (w, v) \in R_{di}$ and d fits w$\}$ for any $d \in S(I) \cup \{\epsilon\}$. It is easy to check that $(M, w) \leftrightsquigarrow (M', w)$. QED

Proposition 2.3.4 *Given an arbitrary pointed indexed model $(M, w) = (W, \{R_c\}_{c \in S(I)}, V), w$, there is a sub-model of it: $M' = (W', \{R'_c\}_{c \in S(I)}, V')$ which is a pruned model such that $W = W'$ and $(M, w) \leftrightsquigarrow (M', w)$.*

Proof. Similar to the above proof; we cut off all the relations ci from v if there is no $c - path$ from w to v. QED

Notice Based on the above observation and results, without any special notice, we will only work with pruned frames and models in the following.

2.4 Translation From Language $\mathcal{L}^I(\Phi)$ to $\mathcal{L}^{S(I)}(\Phi)$

As we mentioned before, indexed models can also be looked as classical models for the language $\mathcal{L}^{S(I)}(\Phi)$. The following questions arise consequently: can we just use language $\mathcal{L}^{S(I)}(\Phi)$ to talk about those indexed models by adapting the classical semantics? And what is the difference between the two approaches? This section will answer these questions.

Definition 2.4.1 *Given a $\mathcal{L}^I(\Phi)$ formula ϕ, a translation T transforms ϕ into a $\mathcal{L}^{S(I)}(\Phi)$ formula $T(\phi)$ as follows:*

$T(\phi)$	$=$	$T_\epsilon(\phi)$
$T_c(p)$	$=$	p
$T_c(\top)$	$=$	$\top$
$T_c(\neg\psi)$	$=$	$\neg T_c(\psi)$
$T_c(\psi \wedge \psi')$	$=$	$T_c(\psi) \wedge T_c(\psi')$
$T_c(\Box_i\psi)$	$=$	$\Box_{ci}T_{ci}(\psi)$

For example: $T(\Box_i(p \wedge \neg\Box_j q) = \Box_i T_i(p \wedge \neg\Box_j q) = \Box_i(p \wedge \neg T_i(\Box_j q) = \Box_i(p \wedge \neg\Box_{ij}q)$.

Remark 2.4.2 *It is obvious that not every $\mathcal{L}^{S(I)}(\Phi)$ formula is a translation of a formula in $\mathcal{L}^I(\Phi)$. For example, $\Box_{12}\Box_3 p$ is not a translation for any $\mathcal{L}^I(\Phi)$ formula. Intuitively $\Box_{12}\Box_3 p$ is nonsense, since an imaginary agent 12 can not think about the actual agent 3, but only about the imaginary one* 123 *in his mind.*

Based on the insight mentioned in Remark 2.1.6, considering the two roles of an indexed model, we have the following result:

Theorem 2.4.3 *For any $\mathcal{L}^I(\Phi)$ formula ϕ, any indexed model M, any $w \in M$:*

$$M, w \Vdash T_c(\phi) \iff M, w \vDash_c \phi.$$

Proof. Induction on the structure of ϕ. According to the definition of T_c, we have the following:

$$\begin{array}{llllll}
M,s \vDash_c \top & \iff & \text{always} & \iff & M,s \Vdash T_c(\top) \\
M,s \vDash_c p & \iff & p \in V(s) & \iff & M,s \Vdash T_c(p) \\
M,s \vDash_c \neg\phi & \iff & M,s \nvDash_c \phi & \iff & M,s \Vdash T_c(\neg\phi) \\
M,s \vDash_c \phi \wedge \psi & \iff & M,s \vDash_c \phi \text{ and } M,s \vDash_c \psi & \iff & M,s \Vdash T_c(\phi \wedge \psi) \\
M,s \vDash_c \Box_i\phi & \iff & \text{for all } t, \text{ if } sR_{ci}t \text{ then } M,t \vDash_{ci} \phi & \iff & \text{for all } t, \text{ if } sR_{ci}t \text{ then } M,t \Vdash T_{ci}(\phi) \\
 & \iff & M,s \Vdash \Box_{ci}T_{ci}(\phi) & \iff & M,s \Vdash T_c(\Box_i\phi)
\end{array}$$

QED

As an immediate corollary, we have:

Corollary 2.4.4 *For any $\mathcal{L}^I(\Phi)$ formula ϕ, any pointed indexed model M, w, $M, w \Vdash T(\phi) \iff M, w \vDash \phi$.*

Remark 2.4.5 *The above result shows that instead of $\mathcal{L}^I(\Phi)$, we can also use the translatable part of the language $\mathcal{L}^{S(I)}(\Phi)$ along with the Kripke semantics to talk about the indexed models. The reason we choose the indexed semantics with $\mathcal{L}^I(\Phi)$ is that it is more intuitive and there is no restriction on the language. The translation, along with the later correspondence results, will help us to study the indexed models by making use of old results for classical models.*

2.5 Correspondence Between Indexed Model and Classical Model

In this section, we discuss the correspondence results between indexed models and classical models for the language $\mathcal{L}^I(\Phi)$.

We first introduce the most important technique in this section:

Definition 2.5.1 ***(Unravelling)***

Given a pointed classical model $(M, w) = ((W, \{R_i\}_{i\in S(I)}, V), w)$, the pointed classical model $(M^r, w^r) = ((W^r, \{R_i^r\}_{i\in I}, V^r), w^r)$ is called the unravelling of (M, w)(or say the unravelling of M around w) where:

- $W^r = \{s | s \text{ is a } c-\text{path from } w \text{ in } M \text{ for some } c \in S(I)\}$.
- $R_i^r = \{(s, s') | s' \text{ is an } i-\text{extension of } s \text{ in } M\}$.
- $V^r(\langle w, ..., w_n, c\rangle) = V(w_n)$.
- $w^r = (w, \epsilon)$

Similarly we can define the unravelling for indexed models.

Definition 2.5.2 ***(Indexed Unravelling)***

Given a pointed indexed model $(M, w) = ((W, \{R_c\}_{c\in S(I)}, V), w)$, the pointed classical model $(M^r, w^r) = ((W^r, \{R_i^r\}_{i\in I}, V^r), w^r)$ is called the unravelling of (M, w)(or say the unravelling of M around w) where:

- $W^r = \{s | s \text{ is a } c-\text{path from } w \text{ in } M \text{ for some } c \in S(I)\}$.
- $R_i^r = \{(s, s') | s' \text{ is an } i-\text{extension of } s \text{ in } M\}$.
- $V^r(\langle w, ..., w_n, c\rangle) = V(w_n)$.

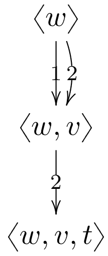
- $w^r = (w, \epsilon)$

Remark 2.5.3

- *The above definition is slightly different from the traditional definition of unravelling. The elements in a traditional unravelling are simply sequences of worlds. We use $c-$paths here to get rid of the unnecessary compositional side-effects. For example, if we unravel the following indexed model traditionally around w:*

  we get:

  which is not exactly what we want, since in the original model there is no 22-path from w to t.

- *Unravelling an indexed model transforms it into a classical model. This is the key to the correspondence result as we will see in the following.*

- *In general, the unravelling of an pruned indexed model around some point only contains partial information of the original model. Only paths starting from the selected points matter for the unravelling around that point.*

Notation It is easy to see that for each pointed classical/indexed model (M, w), there is an unique unravelling $(M^r, (w, \epsilon))$. In the following we denote it as $(Unr(M, w), (w, \epsilon))$.

Proposition 2.5.4 *Given a classical model $M' = (W', \{R'_i\}_{i \in S(I)}, V')$, an indexed model $M = (W, \{R_c\}_{c \in S(I)}, V)$, $w' \in W'$ and $w \in W$. Suppose*

$V(w) = V'(w')$ then for any $c \in S(I) \cup \{\epsilon\}$, any $\mathcal{L}^I(\Phi)$ formula ϕ which doesn't contain modalities, we have:

$$M, w \vDash_c \phi \iff M', w' \Vdash \phi.$$

Proof. Trivial. QED

Lemma 2.5.5 ***(Invariance for Unravelling)***

(a)If a classical model $M', (w, \epsilon)$ is the unravelling of a pointed classical model (M, w), then for any $\mathcal{L}^I(\Phi)$ formulas ϕ, and any $\langle w, ..., w_n, c\rangle \in M'$ we have:

$$M, w_n \Vdash \phi \iff M', \langle w, ..., w_n, c\rangle \Vdash \phi.$$

(b)If a classical model $M', (w, \epsilon)$ is the unravelling of a pointed indexed model (M, w), then for any $\mathcal{L}^I(\Phi)$ formulas ϕ, and any $\langle w, ..., w_n, c\rangle \in M'$ we have:

$$M, w_n \vDash_c \phi \iff M', \langle w, ..., w_n, c\rangle \Vdash \phi.$$

Proof. For part (a):

It is easy to see that the w-generated sub-model of M is a bounded morphism image of M' w.r.t the mapping $f : f\langle w, ..., w_n, c\rangle = w_n$. Then we have $M, w_n \Vdash \phi \iff M', \langle w, ..., w_n, c\rangle \Vdash \phi$ cf. [BRV].

For part (b):

Induction on the structure of ϕ.

- When ϕ is formed by boolean combinations of $p \in \Phi$ and $\top$ then we have $M, w_n \vDash_c \phi \iff M', \langle w, ..., w_n, c\rangle \Vdash \phi$ from Proposition 2.5.4, since $V'(\langle w, ..., w_n, c\rangle) = V(w_n)$.
- Let ψ, ψ' be $\mathcal{L}^I(\Phi)$ formulas. Suppose for any $c - path$ $\langle w, ..., w_n, c\rangle$ in M, we have $M, w_n \vDash_c \psi \iff M', \langle w, ..., w_n, c\rangle \Vdash \psi$ and $M, w_n \vDash_c \psi' \iff M', \langle w, ..., w_n, c\rangle \Vdash \psi'$. It is easy to see that $M, w_n \vDash_c \neg\psi \iff M', \langle w, ..., w_n, c\rangle \Vdash \neg\psi$ and $M, w_n \vDash_c \psi \wedge \psi' \iff M', \langle w, ..., w_n, c\rangle \Vdash \psi \wedge \psi'$.

- Suppose $\phi = \Box_i \psi$. For any $\langle w, ..w_n, c\rangle \in W'$:
$M', \langle w, ..., w_n, c\rangle \Vdash \phi$
$\iff$ for all $s \in W'$, if $\langle w, ..., w_n, c\rangle R_i s$, then $M', s \Vdash \psi$
$\iff$ for all s, if $s = \langle w, ..., w_n, v, ci\rangle \in W'$ for some $v \in W$, $M', s \Vdash \psi$
$\iff$ for all $v \in W$, if $\langle w, ..., w_n, v, ci\rangle \in W'$, then $M', \langle w, ..., w_n, v, ci\rangle \Vdash \psi$
$\iff$ for all $v \in W$, if $\langle w, ..., w_n, v, ci\rangle \in W'$ then $M, v \vDash_{ci} \psi$ (from Induction Hypothesis)
$\iff$ for all $v \in W$, if $w_n R_{ci} v$, then $M, v \vDash_{ci} \psi$ (since $\langle w, ..., w_n, c\rangle \in W'$)
$\iff$ $M, w_n \vDash_c \Box_i \psi$
$\iff$ $M, w_n \vDash_c \phi$.

QED

Since $(w, \epsilon) \in Unr(M, w)$, as an immediate corollary of Lemma 2.5.5, we have:

Corollary 2.5.6 *Given any pointed indexed model* (M, w)*, for all* $\mathcal{L}^I(\Phi)$ *formulas* $\phi : M, w \vDash \phi \iff Unr(M, w), (w, \epsilon) \Vdash \phi$.

Moreover, we now prove a proposition which will be useful in the later sections.

Proposition 2.5.7 *Given an indexed pointed model* (M, w)*, if* vR_iw *and for every sequence* $\langle t_1...t_n\rangle, n \in \mathbb{N}, c \in S(I)$ *the following holds:*

$\langle v, w, t_1, ..., t_n, ic\rangle$ *is an* $ic-$*path* $\iff$ $\langle w, t_1, ..., t_n, c\rangle$ *is a* $c-$*path*

then for any $\phi \in From(\mathcal{L}^I(\Phi))$:

$$M, w \vDash_i \phi \iff M, w \vDash \phi$$

Proof. Consider the unravellings $Unr(M, v)$ of M around v and $Unr(M, w)$ of M around w. From part (b) of Lemma 2.5.5, we have for every $\mathcal{L}^I(\Phi)$ formulas ϕ: $M, w \vDash_i \phi \iff Unr(M, v), \langle v, w, i\rangle \Vdash \phi$ and $M, w \vDash \phi \iff Unr(M, w), \langle w, \epsilon\rangle \Vdash \phi$. However since for every sequence of $\langle t_1...t_n\rangle, n \in \mathbb{N}, c \in S(I)$:

$\langle v, w, t_1, ..., t_n, ic\rangle$ is an $ic-$path $\iff$ $\langle w, t_1, ..., t_n, c\rangle$ is a $c-$path

Then it is obvious that the the sub-model of $Unr(M, v)$ generated by the point $\langle v, w, i\rangle$ is isomorphic to $Unr(M, w)$. It follows that for all ϕ, $Unr(M, w), \langle w, \epsilon\rangle \Vdash \phi \iff Unr(M, v), \langle v, w, i\rangle \Vdash \phi$. Namely, $M, w \vDash \phi \iff M, w \vDash_i \phi$.

QED

Now we are ready to build the relationship between indexed models and classical models.

Theorem 2.5.8 ***(Correspondence)***

(a)Given an indexed model $M = (W, \{R_c\}_{c\in S(I)}, V)$ there is a classical model $M' = (W', \{R'_i\}_{i\in I}, V')$ such that for any $w \in W$, there is a $w' \in M'$: for any $\mathcal{L}(\Phi)$ formula ϕ, $M, w \vDash \phi \iff M', w' \Vdash \phi$.

(b)Given a classical model $M' = (W', \{R'_i\}_{i\in I}, V')$ there is an indexed model $M = (W, \{R_c\}_{c\in S(I)}, V)$ such that $W = W'$ and for any $w' \in W'$, any $\mathcal{L}(\Phi)$ formula ϕ: $M, w' \vDash \phi \iff M', w' \Vdash \phi$.

Proof.

- For part (a): given an indexed model $M = (W, \{R_c\}_{c\in S(I)}, V)$. We first define a function f on W: $f(w) = Unr(M, w)$. In other words, f gives the unravelling of each pointed model based on M. Consider the disjoint union of those unravelling: $\biguplus_{w\in W} f(w)$. Since classical modal satisfaction is invariant under disjoint unions, we have:

  $$\biguplus_{w\in W} f(w), (w, \epsilon) \Vdash \phi \iff f(w), (w, \epsilon) \Vdash \phi.$$

  From the Corollary 2.5.6, we have that

  $$M, w \vDash \phi \iff f(w), (w, \epsilon) \Vdash \phi.$$

  Then for each $w \in W$, we have $(w, \epsilon) \in \biguplus_{w\in W} f(w)$ such that for any $\mathcal{L}^I(\Phi)$ formula ϕ:

  $$M, w \vDash \phi \iff \biguplus_{w\in W} f(w), (w, \epsilon) \Vdash \phi.$$

  That is to say, $\biguplus_{w\in W} f(w)$ is the classical model we want.

- For part (b), given a classical model $M' = (W', \{R'_i\}_{i \in I}, V')$ we can build an indexed model $M = (W, \{R_c\}_{c \in S(I)}, V)$ where:

  - $W = W', V = V'$.
  - $R_{ci} = R'_i$ where $c \in S(I) \cup \{\epsilon\}$.

  We claim that:

  Claim : For any $w' \in W'$, $Unr(M', w'), (w', \epsilon)$ is isomorphic to $Unr(M, w'), (w', \epsilon)$.

  If this is true, then from Lemma 2.5.5, we have for any $w \in W', \mathcal{L}(\Phi)$ formula ϕ:

  $$M, w' \models \phi \iff Unr(M, w'), (w', \epsilon) \Vdash \phi$$

  and

  $$Unr(M', w'), (w', \epsilon) \Vdash \phi \iff M', w' \Vdash \phi.$$

  It follows from the above claim that:
  For all $w' \in W'$, $M, w' \models \phi \iff M', w' \Vdash \phi$.

  Now we move on to the proof of the claim. For simplicity, we now call w' in the claim "w_0". From the definition of c-path, we have that if $c = j_1...j_n$, then $\langle w_0, ..., w_n, c\rangle$ is a c-path in $M \iff$ for all $k \in [1, n] : w_{k-1} R_{j_1..j_k} w_k$. Since $R_{ci} = R'_i$ for all $c \in S(I) \cup \{\epsilon\}$, it is clear that for all $k \in [1, n]$:

  $$w_{k-1} R_{j_1..j_k} w_k, \iff w_{k-1} R'_{j_k} w_k$$

  Then we have:

  $\langle w_0, ..., w_n, c\rangle$ is a c-path in $M \iff \langle w_0, ..., w_n, c\rangle$ is a c-path in M'.

  That is to say, the worlds in $Unr(M', w'), (w', \epsilon)$ and the worlds in $Unr(M, w'), (w', \epsilon)$ are the same. Since the worlds actually determine the relations in the unravellings, clearly we have

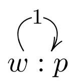
  $\langle w', .., w_n, c\rangle R^r_j \langle w', ..., w_n, v, cj\rangle \iff \langle w', .., w_n, c\rangle R'^r_j \langle w', ..., w_n, v, cj\rangle$.

  Moreover, since $V = V'$, then $Unr(M', w')$ and $Unr(M, w')$ are isomorphic, which completes the proof.

QED

Remark 2.5.9

- *In the proof of part (a), we actually built a "forest" of the unravelled trees. However, only the roots of those trees in* $\biguplus_{w \in W} f(w)$ *are the counterpoints for the worlds in the original indexed models.*
- *Given an indexed pointed model* (M, w)*, it is not always possible, generally, to find a classical pointed model* M', w' *such that* $|M| = |M'|$ *and* $M, w \vDash \phi \iff M, w \Vdash \phi$ *for any* $\mathcal{L}^I(\Phi)$ *formula* ϕ*. See the following simplest example of an indexed model* M*:*



It is impossible to find a singleton classical model M' *such that* $M', w \vDash \neg\Diamond_i\Diamond_i p$ *but* $M', w \vDash \Diamond_i p$ *Normally the classical model* M' *contains many more worlds than* M*, if the above invariance is to be guaranteed.*

Now we want to know under what conditions, two indexed pointed models satisfy the same set of formulas in $\mathcal{L}^I$. The Classical bisimulation seems to be the best starting point.

Definition 2.5.10 ***(Bisimulation)*** *Let* $M = (W, \{R_\Delta\}_{\Delta \in \mathcal{T}}, V)$ *and* $M' = (W', \{R'_\Delta\}_{\Delta \in \mathcal{T}}, V')$ *be two relational models.*

A non-empty binary relation $Z \subseteq W \times W'$ *is called a* bisimulation *between* M *and* M' $(M \leftrightarrows M')$ *if the following conditions are satisfied:*

1. *If* wZw' *then* $V(w) = V(w')$.
2. *If* wZw' *and* $wR_\Delta v$*, then there exists* $v' \in M'$ *such that* vZv' *and* $w'R'_\Delta v'$.
3. *If* wZw' *and* $w'R'_\Delta v'$*, then there exists* $v \in M'$ *such that* vZv' *and* $wR_\Delta v$.

If there is a bisimulation linking $w \in M$ *and* $w' \in M'$*, we say that* w *and* w' *are bisimilar* $(M, w \leftrightarrows M', w')$*.*

It is commonly known that modal formulas are invariant under bisimulation w.r.t to the classical semantics. Now let's see the counter part of this results for indexed model and semantics.

Proposition 2.5.11 *Suppose M, M' are indexed models, for any $w \in M, w' \in M'$, we have $M, w \underline{\leftrightarrow} M', w'$ implies $w \leftrightsquigarrow_c w'$.*

Proof. Consider the indexed models M and M' as classical models for language $\mathcal{L}^{S(\mathcal{I})}(\Phi)$. $M, w \underline{\leftrightarrow} M', w'$ implies for any $\mathcal{L}^{S(I)}(\Phi)$ formula ψ, $M, w \Vdash \psi \iff M', w' \Vdash \psi$. By Proposition 2.4.3, we have for any $\phi \in \mathcal{L}^I(\Phi) : M, w \Vdash T_c(\phi) \iff M, w \vDash_c \phi$. Since for any $\mathcal{L}^I(\Phi)$ formula ϕ, $T_c(\phi)$ is a $\mathcal{L}^{S(I)}(\Phi)$ formula, we have for all $\mathcal{L}^I(\Phi)$ formula ϕ, $M, w \Vdash T_c(\phi) \iff M', w' \Vdash T_c(\phi)$. Namely for all $\mathcal{L}^I(\Phi)$ formula ϕ, $M, w \vDash_c \phi \iff M', w' \vDash_c \phi$. QED

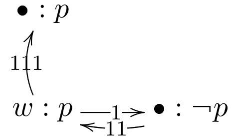
An immediate corollary is as follows:

Corollary 2.5.12 *For any two pointed indexed models $(M, w) and (M', w')$, if $M, w \underline{\leftrightarrow} M', w'$ then $w \leftrightsquigarrow w'$.*

Remark 2.5.13

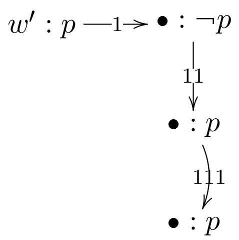
- *Similar invariance results could be proved for generated sub-models and bounded morphisms if we look indexed models as classical models for $\mathcal{L}^{S(\mathcal{I})}(\Phi)$.*
- *The above results shows that bisimulation is too strong to obtain solely $w \leftrightsquigarrow w'$. So we should not expect $w \leftrightsquigarrow w'$ to imply that w and w' are bisimilar, even in the image-finite cases[2]. For example, take two pruned pointed indexed I-models as follows:*

  M:



[2]A model is called finite-image, if no world of it has infinitely many successors.

M' :



It is easy to see that M, w and M', w' are pruned pointed models which are finite but not bisimilar. However, $w \leftrightsquigarrow w'$ and it is not the case that $w \leftrightsquigarrow_c w'$[3].

Notice that in the above counter-example, although M, w and M', w' are not bisimilar, we still have the bisimulation *somewhere else*:

$$Unr(M, w), (w, \epsilon) \,\underline{\leftrightarrow}\, Unr(M', w'), (w', \epsilon).$$

In general, we have the following result:

Proposition 2.5.14 *Let $(M, w), (M', w')$ be two pointed indexed models. $Unr(M, w), (w, \epsilon) \,\underline{\leftrightarrow}\, Unr(M', w'), (w', \epsilon)$ implies that $(M, w) \leftrightsquigarrow (M', w')$.*

Proof. Suppose $Unr(M, w), (w, \epsilon) \,\underline{\leftrightarrow}\, Unr(M', w'), (w', \epsilon)$, then we have $Unr(M, w), (w, \epsilon) \leftrightsquigarrow Unr(M', w'), (w', \epsilon)$. Since $Unr(M, w), (w, \epsilon) \leftrightsquigarrow M, w$ and $Unr(M', w'), (w', \epsilon) \leftrightsquigarrow M', w'$ then we have $(M, w) \leftrightsquigarrow (M', w')$.
QED

2.6 Validity and Indexed Frames

Definition 2.6.1 ***(C-Validity)***

A formula ϕ is c-valid at a world w in an indexed frame $\mathfrak{F}$(notation : $\mathfrak{F}, w \vDash_c \phi$) if $(\mathfrak{F}, V), w \vDash_c \phi$ in every model $(\mathfrak{F}, V)$ based on $\mathfrak{F}$; ϕ is c-valid on a frame $\mathfrak{F}$(notation : $\mathfrak{F} \vDash_c \phi$) if it is $c-$valid at every world in $\mathfrak{F}$. A formula ϕ is $c-$valid on a class of frames $\mathsf{K}(\mathsf{K} \vDash_c \phi)$ if it is c-valid on every frame $\mathfrak{F}$ in K. When $c = \epsilon$ we obtain the normal validity. In the following we call ϵ-validity "validity".

Proposition 2.6.2 $\Box_i(p \to q) \to (\Box_i p \to \Box_i q)$ *is c-valid on the class of all indexed frames for any $c \in S(I) \cup \{\epsilon\}$.*

[3]For example, when $c = 11$.

Proof. Trivial. QED

Proposition 2.6.3 *Any instance of any tautology is c-valid on the class of all indexed frames for any $c \in S(I) \cup \{\epsilon\}$.*

Proof. Trivial. QED

It follows from the above propositions that any instance of tautologies and $\Box_i(p \to q) \to (\Box_i p \to \Box_i q)$ are valid($\epsilon - valid$) on the class of all indexed frames.

Now we prove a proposition which will turn out to be useful in the next section.

Proposition 2.6.4 *A $\mathcal{L}^I(\Phi)$ formula ϕ is classically valid on the class of all classical frames*
$\iff$ ϕ is $c-$valid on the class of all indexed frames for every $c \in S(I)$
$\iff$ ϕ is valid on the class of all indexed frames.

Proof. For the first $\iff$:

$\Rightarrow$: Suppose towards contradiction that ϕ is classically valid on the class of all classical frames, but there is a pointed indexed model $(M, w) = (W, \{R_d\}_{d \in S(I)}, V)$ such that $M, w \not\models_c \phi$ for some $c \in S(I) \cup \{\epsilon\}$. Then we can build an indexed pointed model $(M', w') = (W', \{R'_d\}_{d \in S(I)}, V')$ such that $W = W', V = V', w = w', R'_d = R_{cd}$. Then it is obvious that $M', w' \not\models \phi$. Then from Lemma 2.5.5 we have $Unr(M', w'), (w', \epsilon) \not\models \phi$. Since $Unr(M', w')$ is a classical model then ϕ is not valid w.r.t. the class of all the classical frames. Contradiction.
$\Leftarrow$: Suppose towards contradiction that ϕ is c-valid on the class of all indexed frames for every $c \in S(I) \cup \{\epsilon\}$, but there is a pointed classical model (M, w), such that $M, w \not\models \phi$. Then from the correspondence result, there is an pointed indexed model (M', w') such that $M', w' \not\models \phi$ then ϕ is not $\epsilon-$valid on the class of all indexed frames. Contradiction.

For the second $\iff$:
$\Rightarrow$: Trivial, follows from the above.
$\Leftarrow$: Suppose ϕ is valid on the class of all indexed frames then it is easy t see that ϕ is valid on the class of all the classical frames, thus it is c-valid on the class of all indexed frames. Suppose not, there is a classical pointed

model (M, w) such that $M, w \not\Vdash \phi$ then from the correspondence result we have there is an indexed pointed model such that $M, w \not\models \phi$. Contradiction. QED

We have already showed in Proposition 2.4.4 that every $\mathcal{L}^I(\Phi)$ formula ϕ can be safely translated to a $\mathcal{L}^{S(I)}(\Phi)$ formula $T(\phi)$ such that $M, v \models \phi \iff M, v \Vdash T(\phi)$. Clearly this invariance result also holds for validity, namely, $\mathfrak{F} \models \phi \iff \mathfrak{F} \Vdash T(\phi)$ for any frame $\mathfrak{F}$. Thus, by using the standard translation `ST` from modal formulas to first-order formulas we can translate a modal frame property[4] into a second-order frame property:

Proposition 2.6.5 *Let ϕ be a $\mathcal{L}^I(\Phi)$ formula. Then for any indexed frame $\mathfrak{F}$, we have:*

$$\mathfrak{F} \models \phi \iff \mathfrak{F} \Vdash T(\phi) \iff \forall P_1 ... \forall P_n \forall x ST_x(T(\phi)).$$

where the second order quantifiers bind second-order variables P_i corresponding to the proposition letters p_i appearing in ϕ.

Proof. The first $\iff$ result is directly from Proposition 2.4.4, and the second follows from the standard proof. QED

Remark 2.6.6 *The above result shows that the validity in indexed semantics is actually a second-order concept just like the validity in classical semantics. The tools for studying definability also work here, the only thing is that we are now only concerned with part of the formulas in $\mathcal{L}^I(\Phi)$. We will not go into details on this issue. In the following, we will focus on some important modal formulas which have first-order correspondents.*

Let us use an infinitary first-order language to talk about indexed frames. It has binary predicates R_c for each $c \in S(I)$ describing the relations in the frames.

Proposition 2.6.7 $\mathfrak{F} \models \Box_i p \to p \iff \mathfrak{F} \Vdash \forall w(wR_iw)$.

Proof. $\Leftarrow$: Suppose $\mathfrak{F} \Vdash \forall w(wR_iw)$ then for all (M, w) based on $\mathfrak{F}$, if $M, w \models \Box_i p$ we have $M, w \models p$.

$\Rightarrow$: Assume towards contradiction that $\mathfrak{F} \not\Vdash \forall w(wR_iw)$ then there is a M based on $\mathfrak{F}$ in which there is a non-reflexive point w. We can revise the valuation V such that $w \notin V(p)$ and for all v, $v \in V(p)$ if wR_iv. Clearly, $M, w \not\models \Box_i p \to p$. Contradiction. QED

[4] If we regard the valid modal formulas are modal frame properties.

Proposition 2.6.8 $\mathfrak{F} \models \Box_i p \to \Box_i \Box_i p \iff \mathfrak{F} \Vdash \forall w \forall v \forall s((wR_i v \wedge vR_{ii} s) \to wR_i s)$.

Proof. $\Leftarrow$: Suppose $\mathfrak{F} \Vdash \forall w \forall v \forall s(wR_i v \wedge vR_{ii} s \to wR_i s)$ and for a (M, w) based on $\mathfrak{F}$: $M, w \models \Box_i p$. Then we have that for all v: $wR_i v$ implies $v \in V(p)$. From the property of $\mathfrak{F}$ we have for all v if there is an $ii - path$ from w to v then $wR_i v$. That is to say for all v if there is an $ii - path$ from w to v then $v \in V(p)$ namely $M, w \models \Box_i \Box_i p$.

$\Rightarrow$: Assume towards contradiction that $\mathfrak{F} \nVdash \forall w \forall v \forall s(wR_i v \wedge vR_{ii} s \to wR_i s)$ then there is a point w in a model M based on $\mathfrak{F}$ such that there is an $ii - path$ from w to s but it is not the case $wR_i s$. We can revise the valuation V to let $s \notin V(p)$ and for all v, if $wR_i v$ then $v \in V(p)$. Clearly, $M, w \not\models \Box_i p \to \Box_i \Box_i p$. Contradiction.

QED

Proposition 2.6.9 $\mathfrak{F} \models \neg \Box_i p \to \Box_i \neg \Box_i p \iff \mathfrak{F} \Vdash \forall w \forall v \forall s(wR_i v \wedge wR_i s \to vR_{ii} s)$.

Proof. Similar to the classical case, trivial. QED

Proposition 2.6.10 $\mathfrak{F} \models \neg \Box_i \bot \iff \mathfrak{F} \Vdash \forall w \exists v(wR_i v)$.

Proof. Trivial. QED

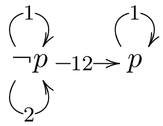
Remark 2.6.11

- *Following the traditional notation, we give the names for the above formulas and the correspondent first-order frame properties:*

$T_i^p : \Box_i p \to p$	$r_i^p : \forall w(wR_i w)$
$4_i^p : \Box_i p \to \Box_i \Box_i p,$	$t_i^p : \forall w \forall v \forall s(wR_i v \wedge vR_{ii} s \to wR_i s)$
$5_i^p : \neg \Box_i p \to \Box_i \neg \Box_i p$	$e_i^p : \forall w \forall v \forall s(wR_i v \wedge wR_i s \to vR_{ii} s)$
$d_i : \neg \Box_i \bot.$	$s_i : \forall w \exists v(wR_i v)$

- *The corresponding frame properties also have intuitive interpretations. For example, the transitive property $\forall w \forall v \forall t(wR_i v \wedge vR_{ii} t \to wR_i t)$ according to 4_i^p actually says: at world w, if agent i indeed thinks v possible and he has the reflection on himself that he would think t possible when at v, then for consistency he must also think t possible since he thinks v could be real world right now at w.*

- *In the above notations, i stands for the agent and p denotes a formula which is only about the factual propositions. We can not substitute arbitrary formula for p. For example, $\mathfrak{F} \models \Box_i\phi \to \phi \iff \mathfrak{F} \Vdash \forall w(wR_iw)$ doesn't hold for arbitrary ϕ:*



Although 1-relations are reflexive everywhere, $\Box_1\Diamond_2p \to \Diamond_2p$ doesn't hold at the $\neg p$ world. This is reasonable since we didn't say anything about how one thinks about others in terms of constraints on relations.

Proposition 2.6.12 *Let T_i be the set of formulas $\{\phi \to \Diamond_i\phi | \phi \in Form(\mathcal{L}^I(\Phi))\}$, r_i be the first-order formula $\bigwedge_{dj \in S(I)} \forall w\forall v((wR_{dj}v \to wR_{idj}v) \wedge ((wR_{idj}v \wedge \exists w'(w'R_dw)) \to wR_{dj}v)) \wedge \forall w(wR_iw)$, then:*

$$\mathfrak{F} \models T_i \iff \mathfrak{F} \Vdash r_i.$$

Proof.

$\Leftarrow$: Suppose $\mathfrak{F} \Vdash r_i$. Given an arbitrary $\phi \in Form(\mathcal{L}^I(\Phi))$, let (M, w) be a pointed indexed model based on $\mathfrak{F}$, such that $M, w \models \phi$. We need to prove that $M, w \models \Diamond_i\phi$. Since $\mathfrak{F} \Vdash \forall w(wR_iw)$, then every world is reflexive. Then we only need to show that $M, w \models_i \phi$. To prove this, we have the following claim:

Claim: For any $n \in \mathbb{N}$, any sequence $t_1, ..., t_n, t$: $\langle w, w, t_1, ..., t_n, t, idj\rangle$ is an *idj*-path $\iff$ $\langle w, t_1, ..., t_n, t, dj\rangle$ is a *dj*-path.

Induction on the length of d:

- Suppose the length of d is 0, namely $d = \epsilon$. If $\langle w, t, j\rangle$ is an j-path then since $\mathfrak{F} \Vdash \forall w\forall v(wR_{dj}v \to wR_{idj}v) \wedge \forall w(wR_iw)$, we have $\langle w, w, t, ij\rangle$ is an $ij-$path. Suppose $\langle w, w, t, ij\rangle$ is an $ij-$path. Then from $wR_\epsilon w$ and $\mathfrak{F} \Vdash \bigwedge_{dj \in S(I)} \forall w\forall v((wR_{idj}v \wedge \exists w(w'R_dw) \to wR_{dj}v))$, we have wR_jt, then $\langle w, t, j\rangle$ is an j-path.
- Suppose when the length of d is k, we have $\langle w, w, t_1, ..., t_k, t, idj\rangle$ is an *idj*-path $\iff$ $\langle w, t_1, ..., t_k, t, dj\rangle$ is a *dj*-path.

- Suppose the length of d is $n+1$. If $\langle w, t_1, ..., t_{k+1}, t, dj\rangle$ is an dj-path, then since $\mathfrak{F} \Vdash \forall w \forall v(wR_{dj}v \rightarrow wR_{idj}v)$, we have $\langle w, w, ..., v, t, idj\rangle$ is an $idj-$path. Suppose $\langle w, w, t_1, ..., t_{k+1}, t, idj\rangle$ is an $idj-$path. Then $\langle w, w, t_1, ..., t_{k+1}, id\rangle$ is an $id-$path. From Induction Hypothesis, we have $\langle w, t_1, ..., t_{k+1}, d\rangle$ is a $d-$path. Then there is a t_k such that $t_k R_d t_{k+1}$. From $\mathfrak{F} \Vdash \bigwedge_{dj\in S(I)} \forall w \forall v((wR_{idj}v \wedge \exists w(w'R_d w) \rightarrow wR_{dj}v))$, we have $t_{k+1}R_{dj}t$. Since $\langle w, t_1, ..., t_{k+1}, d\rangle$ is a $d-$path then $\langle w, t_1, ..., t_{k+1}, t, dj\rangle$ is a $dj-$path.

From proposition 2.5.7, $M, w \vDash_i \phi \iff M, w \vDash \phi$. Then we have $M, w \vDash_i \phi$.

$\Rightarrow$: Suppose that $\mathfrak{F} \nVdash r_i$. If $\mathfrak{F}$ is not i-reflexive then from Proposition 2.6.12, we can find a formula ϕ in the shape of $p \rightarrow \Diamond_i p$ such that $\mathfrak{F} \nvDash \phi$. There are two other cases to be considered:

1. There are $w, v \in \mathfrak{F}$ such that $(w, v) \in R_{dj}$ but $(w, v) \notin R_{idj}$.
2. There are $w, v \in \mathfrak{F}$ such that $wR_{idj}v \wedge \exists w'(w'R_d w)$ but $(w, v) \notin R_{dj}$.

For case 1: We define a valuation V as follows: $V(p) = \{w\}; V(q) = \{v\}$[5]. Thus we can talk about w and v by their unique valuations. Since $\mathfrak{F}$ is a pruned frame and $(w, v) \in R_{dj}$, there is a $dj-$path from some world w_0 to v extending a d-path from w_0 to w. Let's consider the formula $\phi = \Diamond_d(p \wedge \Diamond_j q)$. It is clear that $(\mathfrak{F}, V, w_0) \vDash \phi$. However, $(\mathfrak{F}, V, w_0) \nvDash \Diamond_i \phi$ since there is no R_{idj} relation from w to v.

For case 2: We define a valuation V as above: $V(p) = \{w\}; V(q) = \{v\}$. Since $\mathfrak{F}$ is a pruned frame and $(w', w) \in R_d$, there is an $d-$path from some world w_0 to w. Let's consider the formula $\phi = \Diamond_d(p \wedge \neg\Diamond_j q)$. Since there is no R_{dj} relation from w to v, then it is clear that $(\mathfrak{F}, V, w_0) \vDash \phi$. However, $(\mathfrak{F}, V, w_0) \nvDash \Diamond_i \phi$ since $(w, v) \in R_{idj}$.

QED

Proposition 2.6.13 *Let 4_i be the set of formulas $\{\Diamond_i\Diamond_i\phi \rightarrow \Diamond_i\phi | \phi \in Form(\mathcal{L}^I(\Phi))\}$, t_i be the first-order formula $\bigwedge_{dj\in S(I)} \forall w \forall v((wR_{iidj}v \rightarrow wR_{idj}v) \wedge ((wR_{idj}v \wedge \exists w'(w'R_{iid}w)) \rightarrow wR_{iidj}v)) \wedge \forall w \forall v \forall s(wR_i v \wedge vR_{ii}s \rightarrow wR_i s)$, then:*

$$\mathfrak{F} \vDash 4_i \iff \mathfrak{F} \Vdash t_i.$$

[5]For other proposition letters in Φ, arbitrary.

Proof.

$\Leftarrow$: Suppose $\mathfrak{F} \Vdash t_i$. Given an arbitrary $\phi \in Form(\mathcal{L}^I(\Phi))$, let (M, w) be a pointed indexed model based on $\mathfrak{F}$ such that $M, w \vDash \Diamond_i \Diamond_i \phi$. We need to show that $M, w \vDash \Diamond_i \phi$. Since $\mathfrak{F} \Vdash \forall w \forall v \forall s (w R_i v \wedge v R_{ii} s \to w R_i s)$, then if there is an ii-path from w to v then there is an i-path from w to v. Since $M, w \vDash \Diamond_i \Diamond_i \phi$, there is a *ii*-path from w to v extending an *i*-path $\langle w, w', i \rangle$ such that $M, v \vDash_{ii} \phi$. We only need to show that $M, v \vDash_i \phi$. To prove this, we claim that:

Claim: For any $j \in I$, $n \in \mathbb{N}$ and any sequence $\langle t_1 ... t_{n-1}, t \rangle$: $\langle w, w', v, t_1, ..., t, iidj \rangle$ is an $iidj-$path $\iff$ $\langle w, v, t_1, ..., t, idj \rangle$ is an $idj-$path.

Do induction on the length of d:

- Suppose the length of d is 0, namely $d = \epsilon$. If $\langle w, w', v, t, iij \rangle$ is an iij-path then since $\mathfrak{F} \Vdash \bigwedge_{dj \in S(I)} \forall w \forall v (w R_{iidj} v \to w R_{idj} v)$, we have $vRijt$. Then $\langle w, v, t, ij \rangle$ is an $ij-$path. Suppose $\langle w, v, t, ij \rangle$ is an $ij-$path. Then from $w' R_{ii} v$ and $\mathfrak{F} \Vdash \bigwedge_{dj \in S(I)} \forall w \forall v (w R_{idj} v \wedge \exists w' (w' R_{iid} w) \to w R_{iidj} v)$, we have $v R_{iij} t$, then $\langle w, w', v, t, iij \rangle$ is an iij-path.
- Suppose when the length of d is k, we have $\langle w, w', v, t_1, ... t_k, t, iidj \rangle$ is an $iidj$-path $\iff$ $\langle w, v, t_1, ... t_k, t, idj \rangle$ is an $idj-$path.
- Suppose the length of d is $k + 1$. If $\langle w, w', v, t_1, ..., t_{k+1}, t, iidj \rangle$ is an $iidj$-path, then from $\mathfrak{F} \Vdash \bigwedge_{dj \in S(I)} \forall w \forall v (w R_{iidj} v \to w R_{idj} v)$, we have $\langle w, v, t_1, ..., t_{k+1}, t, idj \rangle$ is an $idj-$path. Suppose $\langle w, v, t_1, ..., t_{k+1}, t, idj \rangle$ is an $idj-$path. Then $\langle w, v, t_1, ..., t_k, t_{k+1}, id \rangle$ is an $id-$path. From the Induction Hypothesis, $\langle w, w', v, t_1, ..., t_k, t_{k+1}, iid \rangle$ is an $iid-$path. Then $t_k R_{iid} t_{k+1}$. From $\mathfrak{F} \Vdash \bigwedge_{dj \in S(I)} \forall w \forall v (\exists w' (w' R_{iid} w) \to w R_{iidj} v)$, then $t_{k+1} R_{iidj} t$. Namely, we have $\langle w, w', v, t_1, ..., t_{k+1}, t, iidj \rangle$ is an $iidj-$path.

Similar to the proof of Proposition 2.5.7, it is easy to see that $M, v \vDash_{ii} \phi \iff M, v \vDash_i \phi$. Then it follows that $M, v \vDash_i \phi$.

$\Rightarrow$: Suppose that $\mathfrak{F} \nVdash t_i$. If $\mathfrak{F} \nVdash \forall w \forall v \forall s (w R_i v \wedge v R_{ii} s \to w R_i s)$ then from Proposition 2.6.8, we can find a formula ϕ in the shape of $\Diamond_i \Diamond_i p \to \Diamond_i p$ such that $\mathfrak{F} \nvDash \phi$. There are two cases to be considered:

1. There are $w, v \in \mathfrak{F}$ such that $w R_{idj} v \wedge \exists w' (w' R_{iid} w)$ but $(w, v) \notin R_{iidj}$.
2. There are $w, v \in \mathfrak{F}$ such that $(w, v) \in R_{iidj}$ but $(w, v) \notin R_{idj}$.

For case 1: We define a valuation V as follows: $V(p) = \{w\}; V(q) = \{v\}$. Since $\mathfrak{F}$ is a pruned frame and there is a w' such that $(w', w) \in R_{iid}$, then there is an $iid-$path from some world w_0 to w. Let's consider the formula $\phi = \Diamond_d(p \wedge \neg\Diamond_j q)$. Since $(w, v) \notin R_{iidj}$ then $(\mathfrak{F}, V, w_0) \vDash \Diamond_i\Diamond_i\phi$. However, $(\mathfrak{F}, V, w_0) \nvDash \Diamond_i\phi$ since $(w, v) \in R_{idj}$[6].

For case 2: We define a valuation V as above: $V(p) = \{w\}; V(q) = \{v\}$. Since $\mathfrak{F}$ is a pruned frame and $(w, v) \in R_{iidj}$, there is a $iidj-$path from some world w_0 to v extending a iid-path from w_0 to w. Let's consider the formula $\phi = \Diamond_d(p \wedge \Diamond_j q)$. It is clear that $(\mathfrak{F}, V, w_0) \vDash \Diamond_i\Diamond_i\phi$. However, $(\mathfrak{F}, V, w_0) \nvDash \Diamond_i\phi$ since there is no R_{idj} relation from w to v.

QED

Proposition 2.6.14 *Let* 5_i *be the set of formulas* $\{\neg\Box_i\phi \to \Box_i\neg\Box_i\phi | \phi \in Form(\mathcal{L}^I(\Phi))\}$. *Let* e_i *be the first-order formula* $\bigwedge_{dj \in S(I)} \forall w\forall v((wR_{idj}v \to wR_{iidj}v) \wedge ((wR_{iidj}v \wedge \exists w'(w'R_{id}w)) \to wR_{idj}v)) \wedge \forall w\forall v\forall s(wR_i v \wedge wR_i s \to vR_{ii}s)$, *then:*

$$\mathfrak{F} \vDash 5_i \iff \mathfrak{F} \Vdash e_i.$$

Proof. Similar to the above proof. QED

Remark 2.6.15 *The corresponding first-order formula for* T_i *roughly requires the i-reflexivity and* $R_c = R_{ic}$ *for every* $c \in S(I)$, *which coincides with our intuition about true beliefs. The first-order correspondents for* 4_I *and* 5_i *roughly says* $R_{iic} = R_{ic}$. *Here we said "roughly", since there are extra constraints in those first-order correspondents. For example, in*

$$r_i : \bigwedge_{dj \in S(I)} \forall w\forall v(wR_{dj}v \to wR_{idj}v \wedge (wR_{idj}v \wedge \exists w'(w'R_d w) \to wR_{dj}v)) \wedge \forall w(wR_i w),$$

we don't have the exact $R_c = R_{ic}$. *Instead, we have* $R_{dj} \subseteq R_{idj}$ *and if* $\exists w'(w'R_d w)$ *then* $wR_{idj}v$ *implies* $wR_{dj}v$. *Unfortunately, the intuition of this extra constraint is not very clear so far.*

If we want to model knowledge, then following the traditions in epistemic logic, we should have T_i, 4_i and 5_i all as axioms in the intended logic. It is easy to check that:

Proposition 2.6.16 $\mathfrak{F} \vDash T_i \cup 4_i \cup 5_i \iff \mathfrak{F} \Vdash r_i \wedge t_i^p \wedge e_i^p \wedge \forall w\forall v \bigwedge_{c \in S(I)}(wR_{iic}v \leftrightarrow wR_{ic}v)$.

[6]No matter whether there is an $id-$path to w.

If we want to model the situations in which agents only have true beliefs about the primitive facts, but no enough information about others, then we should require T_i^p, 4_i and 5_i. The corresponding first-order frame property is $r_i^p \wedge t_i \wedge e_i$.

Now we conclude the above corresponding results here:

Modal formula	First-order correspondent	Class of frames
K_i	$\top$	all frames
T_i^p	r_i^p	i-reflexive frames
T_i	r_i	restricted i-reflexive frames
4_i^p	t_i^p	i-transitive frames
4_i	t_i	restricted i-transitive frames
5_i^p	e_i^p	i-euclidean frames
5_i	e_i	restricted i-euclidean frames
d_i	s_i	i-serial frames

As we have shown above, the classical uniform substitution doesn't preserve validity w.r.t to arbitrary class of frames. Instead of uniform substitution, we can have a weaker version. To define it, we will use the following concept.

Definition 2.6.17 ***(I-modal Depth)*** *The I-modal depth of an occurrence of a proposition letter p in a $\mathcal{L}^I(\Phi)$ formula ϕ (notation: $(ID(o(p), \phi))$ is a sequence in $S(I) \cup \{\epsilon\}$ which is defined recursively as follows:*

- $ID(o(p), p) = ID(\top) = \epsilon$
- $ID(o(p), \neg\phi) = ID(o(p), \phi)$
- $ID(o(p), \phi \wedge \psi) = \begin{cases} ID(\phi) & if \quad o(p) \text{ appears in } \phi \\ ID(\psi) & if \quad o(p) \text{ appears in } \psi \end{cases}$
- $ID(o(p), \Box_i\phi) = iID(o(p), \phi)$.

where $o(p)$ is the occurrence of p.

If all the occurrences of a proposition letter p in a formula have the same I-modal depth c, then we say p in ϕ has the uniform $I-$modal depth c.

Definition 2.6.18 ***(I-Uniform Substitution)*** *A transformation from a $\mathcal{L}^I(\Phi)$ formula ϕ to another $\mathcal{L}^I(\Phi)$ formula θ is called an I-uniform substitution, if θ is obtained from ϕ in one of the following ways:*

1. *uniformly replacing proposition letters of the uniform I-modal depth in ϕ with arbitrary $\mathcal{L}^I(\Phi)$ formulas.*

2. *uniformly replacing proposition letters in ϕ with arbitrary $\mathcal{L}^I(\Phi)$ formulas without any modalities.*

Proposition 2.6.19 *I-uniform substitution preserves validity on any class* K *of indexed frames.*

Proof. Suppose towards contradiction that there is a frame class K such that K $\models \phi(p)$ but K $\not\models \phi(\psi)$ where $\phi(\psi)$ is obtained by I-uniformly substituting ϕ for p in $\phi(p)$. Then there is a pointed indexed model $(\mathfrak{F}, V, w)$ based on a frame $\mathfrak{F} \in$ K such that $(\mathfrak{F}, V, w) \not\models \phi(\psi)$.

- Suppose $\phi(\psi)$ is obtained in the first way of the definition of I-substitution. Then let $V'(q) = V(q)$ for any proposition letter $q \in \Phi$ except p, but let $V'(p) = \{w | w \models_c \psi\}$ where c is the uniform I-modal depth of p in ϕ. then it is easy to see that $(\mathfrak{F}, V', w) \not\models \phi(p)$. Contradiction.

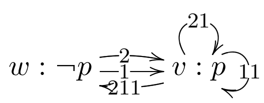
- Suppose $\phi(\psi)$ is obtained by in second way of the definition of I-substitution. Then let $V'(q) = V(q)$ for any proposition letter $q \in \Phi$ except p and let $V'(p) = \{w | w \models \psi\}$. It is easy to see that $(\mathfrak{F}, V', w) \not\models \phi(p)$. Contradiction.

That is to say I-substitution preserves validity[7]. QED

Definition 2.6.20 ***(I-Uniform Substitution Rule)*** *Given a $\mathcal{L}^I(\Phi)$ formula ϕ, prove θ where θ is obtained by I-uniform substitution from ϕ.*

Classical generalization(that is, given ϕ prove $\Box_i\phi$) doesn't preserve validity on arbitrary class of frames either. For example, $\Box_1 p \to \Box_1\Box_1 p$ is valid w.r.t to the class of 1-transitive frames but $\Box_2(\Box_1 p \to \Box_1\Box_1 p)$ is not. Consider the following model:

M:



[7] Informally, when a formula is valid on some frame, this can not depend on the particular value its propositional letters have. Thus it should be safe to uniformly replace these letters with any other formulas as long as every occurrence of the uniform replacement has the same meaning. The 3 conditions of I-substitution guarantee this.

It is easy to see that M is based on a 1-transitive frame, but clearly, $M, w \not\models \Box_2(\Box_1 p \to \Box_1\Box_1 p)$ which means $\Box_2(\Box_1 p \to \Box_1\Box_1 p)$ is not valid on the class on 1−transitive frames.

Remark 2.6.21 *It is reasonable not to have the classical generalization all the time as we discussed in the first chapter. In the classical set-up, if we assume an agent is positive introspective, then other agents must "know/believe" this by classical generalization.*

2.7 I-modal Logics

Definition 2.7.1 ***(I-modal Logics)*** *An I-modal logic Λ is a set of modal formulas that contains all instances of propositional tautologies and is closed under modus ponens(MP: $\phi \in \Lambda$ and $\phi \to \psi \in \Lambda$ then $\psi \in \Lambda$.) and I-uniform substitution. If $\phi \in \Lambda$, we say that ϕ is a theorem of Λ(notation: $\vdash_\Lambda \phi$).*

Soundness and completeness are defined as in the classical set-up.

Definition 2.7.2 ***(Soundness)***
Let S be a class of frames(or models). A I-modal logic Λ is sound with respect to S if for all formula ϕ, and all structures $\Theta \in S$, $\vdash_\Lambda$ implies $\Theta \models \phi$.

Definition 2.7.3 ***(Completeness)*** *Let S be a class of frames(or models). An I-modal logic Λ is strongly complete with respect to S if for any set of formulas $\Gamma \cup \{\phi\}$, $\Gamma \models_S \phi$ implies $\Gamma \vdash_\Lambda \phi$. An I-modal logic Λ is weakly complete with respect to S if for any formula ϕ, $S \models \phi$ implies $\vdash_\Lambda \phi$.*

Proposition 2.7.4 *A logic Λ is strongly complete with respect to a class of structures $\mathcal{S}$ iff every Λ-consistent set of formulas is satisfiable on some $\Theta \in \mathcal{S}$.*

Proof. Standard, Cf [BRV]. QED

Let K_I be the classical modal logic that contains all instances of tautologies, all the formulas:

$$K_i : \qquad \Box_i(p \to q) \to (\Box_i p \to \Box_i q)$$

for $i \in I$, and closed under MP, uniform substitution(SUB) and generalization(GEN). It is commonly known that this logic is sound and strongly complete with respect to the class of all classical frames. Then we have $\vdash_{K_I} \phi \iff \Vdash \phi$. From Proposition 2.6.4, we have the following:

Proposition 2.7.5 $\vdash_{K_I} \phi \iff \mathtt{K} \vDash \phi$, *where* $\mathtt{K}$ *is the class of all indexed frames.*

Actually we can prove the following stronger result:

Proposition 2.7.6 K_I *is sound and strongly complete with respect to the class of all indexed frames.*

Proof. Soundness follows directly from the above proposition. For the completeness part:
Fix any K_I-consistent set of formulas Δ. From the completeness of K_I w.r.t the class of all classical frames, we have there is a classical pointed model (M, w) such that $M, w \Vdash \Delta$. From the correspondence result, there is an indexed pointed model (M', w') such that $M', w' \vDash \Delta$. QED

Since K_I is clearly an I-modal logic according to the definition, then we can define the normal I-modal logics based on it.

Definition 2.7.7 ***(Normal I-modal Logic)*** *An I-modal logic* Λ *is normal, if it contains the logic* K_I *and is closed under MP and I-uniform substitution(I-SUB). Equivalently, an normal I-modal logic can also be represented as an axiom system, which contains at least all the formulas in* K_I *as its axioms and MP, I-SUB as its rules.*

Notation In the following, we will use the name of certain formulas to denote the normal I-modal logic containing those formulas as extra axioms. For example, $K_I T_I^p$ denotes the normal logic that contains axioms T_i^p for each $i \in I$. Similarly, we call a frame I-reflexive, if it is i-reflexive for each $i \in I$.

Since K_I is sound w.r.t to the class of all the indexed frames, it follows that:

Proposition 2.7.8 K_I *is sound with respect to any class of indexed frames.*

It is obvious that the rule MP preserves validity. From Proposition 2.6.19, we know that the rule $I - SUB$ also preserves validity w.r.t any class of indexed frames. Then to show the soundness for a normal I-modal logic system w.r.t to a frame class $\mathtt{K}$, we only need to check that the extra axioms are valid w.r.t to $\mathtt{K}$. From the correspondence results in the last section, we can easily obtain soundness for several normal logics. For example, logic $K_I T_I^p$ is sound w.r.t to the class of I-reflexive indexed frames; $K_I 4_I 5_I$ is sound w.r.t to the class of restricted transitive and restricted euclidean indexed frames.

We should notice that although classical generalization can not preserve validity on arbitrary class of indexed frames, it indeed could preserve validity on some certain class of frames. From the completeness result for K_I, we know that GEN preserves validity on the class of all indexed frames[8]. Moreover we have the following results.

Proposition 2.7.9 *If ϕ is valid on a class of restricted i-reflexive indexed frames, then $\Box_i\phi$ is also valid on this class of frames.*

Proof. Remember that $\mathfrak{F}$ is a restricted i-reflexive frame $\iff$ $\mathfrak{F} \Vdash T_i$. From the claim in Proposition 2.6.12, and a straightforward generalization of Proposition 2.5.7 to validity, we know that $\mathfrak{F}, w \vDash \phi \iff \mathfrak{F}, w \vDash_i \phi$. Thus if $\mathfrak{F} \vDash \phi$ then $\mathfrak{F} \vDash_i \phi$. It follows that $\mathfrak{F} \vDash \Box_i\phi$. QED

Now let's consider the normal I-logic $K_IT_I + GEN$ which has all the formulas in $\{\Box_i\phi \to \phi | i \in I, \phi \in Form(\mathcal{L}^I(\Phi))\}$ as its extra axioms and GEN as its extra rule of proof. We have the following result:

Proposition 2.7.10 *Logic $K_IT_I + GEN$ is sound and strongly complete with respect to the class of restricted I-reflexive frames.*

Proof. The soundness follows easily from the Proposition 2.7.9(for GEN rule) and Proposition 2.6.12 (for the validity of axioms). For the completeness part:
It is easy to see that $K_IT_I + GEN$ is equivalent to the classical multi-agent logic KT, since T_I actually contains all the uniform substitution instances of $\Box_i p \to p$ and both systems have the generalization rule. Since KT is strongly complete w.r.t the class of I-reflexive classical frames, then given a $K_IT_I + GEN$-consistent set Δ, it is satisfiable at some reflexive classical model, suppose it is $M = (W, \{R_i\}_{i\in I}, V)$. From the construction in Theorem 2.5.8, we know that there is an indexed model M' such that $M' = (W, \{R'_c\}_{c\in S(I)}, V)$ where $R'_{ci} = R_i$ and $M, w \Vdash \phi \iff M', w \vDash \phi$ for any $\phi \in Form(\mathcal{L}^I(\Phi)), w \in W$. Since M is I-reflexive and $R'_i = R_i$ for each $i \in I$, then M' is I-reflexive. Moreover, for any $i \in I$, $c = \langle j_1, ..., j_n\rangle \in S(I)$, $R'_{ic} = R_{j_n} = R'_{j_1,...,j_n} = R'_c$[9]. It follows M' is restricted I-reflexive. So any $K_IT_I + GEN$-consistent set Δ is satisfiable at some restricted I-reflexive indexed model. QED

[8] It follows also from Proposition 2.6.4.

[9] Which is stronger than the requirement for R'_{ic} and R'_c in r_i.

Similarly we have the following result:

Proposition 2.7.11 *$K_IT_I4_I5_I+GEN$ is sound and strongly complete w.r.t the class of restricted I-transitive, restricted I-euclidean and restricted I-reflexive frames.*

Proof. The soundness is straightforward. For completeness: It is easy to see that $K_IT_I4_I5_I+GEN$ is equivalent to the classical multi-agent logic $S5$. Then any $K_IT_I4_I5_I+GEN$-consistent set of formulas Δ is also S_5-consistent. Then according to the completeness result for $S5$, any $K_IT_I4_I5_I+GEN$-consistent set of formulas Δ can be satisfied on some model M, in which the relations are equivalence relations. We still use the construction in the above proof to obtain an indexed model M'. We only need to check if M' is I-euclidean and I-transitive. Suppose there are $w, v, s \in M'$ such that wR'_iv and wR'_is. Since $R'_i = R_i$ and R_i is an equivalence relation then R'_i is an equivalence relation too. It follows that vR'_is. Since $R'_{ii} = R'_i = R_i$ then we have $vR'_{ii}s$. That is to say M is i-euclidean. Since i is arbitrary then M is I-euclidean. Similarly, we can show that M is I-transitive. Since $wR'_iv = wR'_{ic}v$, then M is restricted I-euclidean and restricted I-transitive. QED

$K_IT_I4_I5_I+GEN$ and K_IT_I+GEN are rather special I-modal logics which are equivalent to some classical modal logics. Normally we don't have such equivalence relation, since the normal I-modal logics extending K_I but without T_I axioms, don't contain the classical GEN and SUB in general[10]. Moreover, for a normal I-modal logic, there is no obvious way of constructing the canonical model which is usually used to prove the completeness[11]. In such case, we'd better prove completeness indirectly. Here is a strategy to prove the completeness for arbitrary I-modal logic Λ by making use of the completeness results for some other classical logics:

Try to find a complete classical normal logic Λ_c, such that given any Λ-consistent set Δ, Δ is also Λ_c-consistent. Since Λ_c is complete w.r.t some class of classical frames, then any Λ-consistent set Δ is satisfiable on some classical model M in with certain frame properties. Then we just need to transform this M into an indexed model with the desired property according to the class of frames we want.

[10] The corresponding frame property of T_i guarantees that GEN preserves validity, as we showed in Proposition 2.7.9. However, this is a very special case.

[11] For example, considering how to build R_c relations for the canonical model when $c \notin I$.

Let's look at an example:

Theorem 2.7.12 $\mathrm{K_I T_I^p}$ *is strongly complete with respect to the class of I-reflexive frames.*

To prove this theorem we take the above strategy.

Lemma 2.7.13 *A set of $\mathcal{L}^I(\Phi)$ formulas Δ is $\mathrm{K_I T_I^p}$-consistent $\iff$ $\Delta \cup \{\Box_i \phi \to \phi | i \in I, \phi$ is a formula in $\mathcal{L}^I(\Phi)$ without modalities$\}$ is K_I-consistent.*

Proof. Let $A = \{\Box_i \phi \to \phi | i \in I, \phi$ is a formula in $\mathcal{L}^I(\Phi)$without modalities$\}$. Since $I - SUB$ only substitutes formulas without modalities for the p in $\Box_i p \to p$, then we can equally add all those instances of substitutions as premises in K_I. Namely, we have $\Delta \vdash_{\mathrm{K_I T_I^p}} \phi \iff \Delta \cup A \vdash_{\mathrm{K_I}} \phi$. That is to say $\Delta \nvdash_{\mathrm{K_I T_I^p}} \bot \iff \Delta \cup A \nvdash_{\mathrm{K_I}} \bot$. QED

Since K_I is strongly complete w.r.t the class of all indexed frames then we have:

Lemma 2.7.14 *For any $\mathrm{K_I T_I^p}$-consistent set of $\mathcal{L}^I(\Phi)$ formulas Δ, $\Delta \cup A$ is satisfiable in some indexed model.*

Now we only need to transform the model which satisfies Δ, into another one with the desired property[12].

Let us first define a term which is useful here.

Definition 2.7.15 ***(Descriptive List)*** *We call a $\mathcal{L}^I(\Phi)$ formula π a "descriptive list" if π is in the form of $\bigwedge_{p \in \Phi} \pm p$, which can be regarded as a full truth list of proposition letters in the finite set Φ. We denote the set of all the possible descriptive lists as $V(\Phi)$.*

Remark 2.7.16 *For example, if $\Phi = \{p, q\}$, then the formula $p \wedge \neg q$ is a descriptive list. Intuitively, a descriptive list is like a valuation on a world in a model.*

Lemma 2.7.17 *Suppose $\{\Box_i \phi \to \phi | i \in I, \phi$ is a formula in $\mathcal{L}^I(\Phi)$ without modalities$\}$ is satisfiable at a pointed indexed model (M, w), then there is an I-reflexive pointed indexed model (M^*, w^*) such that for all formula ϕ, $M^*, w^* \models \phi \iff M, w \models \phi$.*

[12]It is easy to see that we can not simply add reflexive relations at each point.

Proof. Suppose $\{\Box_i\phi \to \phi | i \in I, \phi$ is a formula in $\mathcal{L}^I(\Phi)$ without modalities$\}$ is satisfiable at a pointed indexed model (M, w). Then for each $i \in I$, since $M, w \Vdash \Box_i\neg\pi \to \neg\pi$ then $M, w \Vdash \pi \to \Diamond_i\pi$ where π is the descriptive list according to $V(w)$[13]. Namely, for each $i \in I$, there is a world v such that wR_iv and $V(w) = V(v)$. It follows that, in the unravelling model $Unr(M, w) = (S, \{R^r_i\}_{i\in I}, V^r)$, for each $i \in I$ there is a (w, v, i) such that $(w, \epsilon)R_i(w, v, i)$ and $V((w, \epsilon)) = V(((w, v, i))$. We pick one such point for each i and let T be the set of them. We now construct a pointed indexed model $(M^*, w^*) = (S^*, \{R^*_c\}_{c\in S(I)}, V^*), w^*$ as follows:

- $S^* = S - T$.
- $V^*(s) = V^r(s)$.
- $w^* = (w, \epsilon)$.
- R^*_c is obtained by the following operations:
  1. First step: let $R'_{ci} = \{(s, t) | s = \langle w, ..., w_n, c\rangle, t = \langle w, ..., w_n, w_{n+1}ci\rangle \in S\}$ for each $c \in S(I)$ and $i \in I$.
  2. Second step: Let $R''_c = R'_{ci}|S^*$, and if $t = \langle w, v, i\rangle \in T$ and $tR^r_{ij}s$ for any $i, j \in I$ then add $(\langle w, \epsilon\rangle, s)$ into R''_{ij}.
  3. Third step: Let $R^*_c = R''_c$ for all the $c \in S(I) - I$, Let $R^*_i = R''_i \cup \{(s, s) | s \in S^*\}$ for each $i \in I$.

The intuition behind such construction is this: we first we rename the relations to encode the path information. Secondly, for each $i \in I$, we cut off all the certain worlds (w, v, i) in T. Then we engraft the subtrees rooted at $t \in T$ to (w, ϵ). Finally, we make the model I-reflexive closure, namely to add all the reflexive i-relations at each world for each $i \in I$.

It is obvious that $Unr(M^*, w^*)$ is isomorphic to $Unr(M, w)$. Intuitively $Unr(M^*, w^*)$ unravels the i-paths we hide at $\langle w, \epsilon\rangle$ in (M^*, w^*).

From Lemma 2.5.5, we have $M^*, w^* \vDash \phi \iff M, w \vDash \phi$. QED

Then from this Lemma 2.7.17 and Lemma 2.7.14, we have the following Lemma which implies the complete theorem directly.

Lemma 2.7.18 *Any* $\mathrm{K_IT^p_I}$*-consistent set* Δ *is satisfiable in some I-reflexive indexed model.*

[13] We suppose Φ is finite.

Remark 2.7.19 *In general, to prove completeness for arbitrary normal I-modal logic is difficult. We don't have an elegant way to deal with it uniformly. Besides the above results, we are more interested in the following conjectures about the normal I-modal logics indeed without SUB and GEN:*

- $K_I 4^p$ *is strongly complete w.r.t the class of I-transitive frames.*
- $K_I 4$ *is strongly complete w.r.t the class of restricted I-transitive frames.*
- $K_I 4_I 5_I D_I$ *is strongly complete w.r.t the class of restricted I-transitive, restricted I-euclidean and I-serial frames.*

Proving such results will definitely give us better understanding of indexed model and its semantics. We leave them here for further study.

Chapter 3

Non-redundant Models and Awareness

3.1 Non-redundant Models

In many cases, the agents are only interested in the primitive possibilities and what they care about others is also others' information about those primitive possibilities. However, in the models we often have multiple worlds sharing the same primitive facts. A natural question arises: When can we have the models with unique primitive possibilities? In this section, we will give the criteria for a set of formulas to have the desirable indexed/classical models containing only unique primitive possibilities. It will also be clear that how our indexed approach reduce the size of models than the classical setting in this specific case.

Definition 3.1.1 ***(Redundant Model)*** *A classical(indexed) model M is called a redundant model if there are $w, v \in W$ such that $V(w) = V(v)$. If a classical(indexed) model M is not a redundant model, then we say M is non-redundant.*

Let Δ and Δ^* be two sets of $\mathcal{L}^I(\Phi)$ formulas:

$$\Delta = \{\Diamond_c(\pi \wedge \Diamond_i \pi') \wedge \Diamond_d(\pi \wedge \neg\Diamond_i \pi') | i \in I; c, d \in S(I) \cup \{\epsilon\}, \pi, \pi' \in V(\Phi)\}^1,$$

$$\Delta^* = \{\Diamond_c(\pi \wedge \Diamond_i \pi') \wedge \Diamond_c(\pi \wedge \neg\Diamond_i \pi') | i \in I; c \in S(I) \cup \{\epsilon\}, \pi, \pi' \in V(\Phi)\}.$$

Moreover, let $\neg\Delta = \{\neg\phi | \phi \in \Delta\}, \neg\Delta^* = \{\neg\phi | \phi \in \Delta^*\}$.

[1]Remember that as we defined in last chapter, $V(\Phi)$ is the set of all descriptive lists and Φ is finite.

Lemma 3.1.2

- *(a) For any pointed classical model (M, w), $M, w \Vdash \neg\Delta \iff$ there is a non-redundant pointed classical model (M^n, w^n) such that $M^n, w^n \leftrightsquigarrow M, w$.*

- *(b) For any pointed indexed model (M, w), $M, w \vDash \neg\Delta^* \iff$ there is a non-redundant pointed indexed model (M^n, w^n) such that $M^n, w^n \leftrightsquigarrow M, w$.*

Proof. For (a):
$\Rightarrow$: Suppose $M, w \Vdash \neg\Delta$. We claim that:

Claim: If two worlds v, v' are connected to w[2] and $V(v) = V(v')$ then $M, v \leftrightsquigarrow M, v'$.

Now we prove the above claim by induction on the structure of a $\mathcal{L}^I(\Phi)$ formula ψ:

- Suppose ψ is a boolean combination of primitive propositions and $\top$. It is clear that $M, v \Vdash \psi \iff M, v' \Vdash \psi$, since $V(v) = V(v')$.

- Let ϕ, ϕ' be two $\mathcal{L}^I(\Phi)$ formulas. Suppose for any v, v' such that $V(v) = V(v')$ we have $M, v \Vdash \phi \iff M, v' \Vdash \phi$ and $M, v \Vdash \phi' \iff M, v' \Vdash \phi'$. It is easy to see that $M, v \Vdash \phi \wedge \phi' \iff M, v' \Vdash \phi \wedge \phi'$ and $M, v \Vdash \neg\phi \iff M, v' \Vdash \neg\phi$.

- When ψ is of the form $\Diamond_i \psi'$, suppose towards contradiction that $M, v \Vdash \psi$ and $M, v' \nVdash \psi$, namely $M, v' \Vdash \neg\psi$. Since $M, v \Vdash \psi$, there is a world s such that $vR_i s$ and $M, s \Vdash \psi'$. Let π' be the descriptive list according to $V(s)$, we claim that

  $$\textbf{Claim}' : M, v' \Vdash \neg\Diamond_i \pi'.$$

  Suppose not, then $M, v' \Vdash \Diamond_i \pi'$, namely there is a world s' such that $v' R_i s'$ and $V(s') = \pi' = V(s)$. It is obvious that s and s' are connected to w since v and v' are connected to w. Since $M, s \Vdash \psi'$, then from induction hypothesis we have $M, s' \Vdash \psi'$. It follows that $M, v' \Vdash \Diamond_i \psi'$

[2] In the sense that there are paths from w to v and v'. Especially, the path can be the ϵ-path namely v or v' could be w itself.

which contradicts $M, v' \Vdash \neg\psi$. Then **Claim'** is proved.
Moreover it is obvious that $M, v \Vdash \Diamond_i \pi'$ since $V(s) = \pi'$. Suppose v, v' are connected to w by a $c-path$ and a $d-path$ respectively. Let $\pi = V(v) = V(v')$ and formula $\phi = \Diamond_c(\pi \wedge \Diamond_i \pi') \wedge \Diamond_d(\pi \wedge \neg\Diamond_i \pi') \in \Delta$[3]. Then it is clear that $M, w \Vdash \phi$. Contradiction. Now **Claim** is proved.

Suppose for now (M, w) is a pointed model in which every world is connected to w[4]. Based on the **Claim**, we can build a non-redundant model (M^n, w^n) from (M, w) as follows:

- $W^n = \{|v| | v \in W\}$ where $|v|$ is the equivalence class w.r.t $\leftrightsquigarrow$ relation;
- $R_i^n |s||t| \iff \exists s' \in |s| \exists t' \in |t|$ such that $s' R_i t'$;
- $V^n(|v|) = V(v)$;
- $w^n = |w|$.

It is easy to check that M^n is a filtration of M' through $Form(\mathcal{L}^I(\Phi))$. Now from the Filtration Theorem cf.[BRV], we have $M, w \Vdash \phi \iff M^n, |w| \Vdash \phi$.

Suppose there are two worlds $|w|, |v|$ in M^n such that $V^n(|w|) = V^n(|v|)$. From the definition of V^n, we have $V(w) = V(v)$. But from the **Claim**, we have $M^n, w \leftrightsquigarrow M^n, v$. It follows that $v \in |w|$ which means $|v| = |w|$, namely M^n is non-redundant.

$\Leftarrow$: Suppose there is a non-redundant classical pointed model (M^n, w^n) such that $M^n, w^n \Vdash \phi \iff M', w' \Vdash \phi$ for any $\phi \in Form(\mathcal{L}^I(\Phi))$. Suppose towards contradiction that $M^n, w^n \nVdash \neg\Delta$ namely there is a formula $\neg\phi \in \neg\Delta$ such that $M^n, w^n \nVdash \neg\phi$. Then it follows $M^n, w^n \Vdash \phi$ where $\phi \in \Delta$. That is to say there are two worlds $v, v' \in M^n$ such that $V(v) = V(v') = \pi$ and $M^n, v \Vdash \Diamond_i \pi'$ and $M^n, v' \Vdash \neg\Diamond_i \pi'$ for some descriptive list π'. However, since M^n is non-redundant, $v = v'$. But it is not possible for any world $v \in M^n$ to satisfy $\Diamond_i \pi'$ and $\neg\Diamond_i \pi'$ at the same time. Contradiction.

For (b):
$\Rightarrow$: Suppose $M, w \models \neg\Delta^*$. We build the non-redundant model M^r as follows:

[3]Especially, if $c = \epsilon$ (or $d = \epsilon$), then $\phi = \Diamond_\epsilon(\pi \wedge \Diamond_i \pi') \wedge \Diamond_d(\pi \wedge \neg\Diamond_i \pi') = \pi \wedge \Diamond_i \pi' \wedge \Diamond_d(\pi \wedge \neg\Diamond_i \pi')$ according to the notation for $\Diamond_\epsilon$ we mentioned before.
[4]Otherwise we can simply "cut off" all the unreachable worlds.

- Let $W^n = \{|v| | v \in W'\}$ where $|v|$ is the equivalent class w.r.t to the relation $\leftrightsquigarrow_\Phi$[5];
- $R^n_c|s||t| \iff \exists s' \in |s| \exists t' \in |t|$ such that $s'R_c t'$;
- $V^n(|v|) = V(v)$;
- $w^n = |w|$.

Assume without generality that (M, w) is a pruned pointed indexed model[6]. We claim that:

Claim: $Unr(M, w), (w, \epsilon) \leftrightarrows Unr(M^n, |w|), (|w|, \epsilon)$.

If the claim is true, then from Proposition 2.5.14 we have $M, w \leftrightsquigarrow M^n, |w|$ which is what we want.

To prove the claim we have to find the bisimulation between those two models. Let's define a relation $Z \subseteq Unr(M, w) \times Unr(M^n, |w|) : (s, t) \in Z \iff s = \langle w, ..., w_n, c\rangle$ and $t = \langle |w|, ..., |w_n|, c\rangle$ for some $c \in S(I) \cup \{\epsilon\}$. We now show that Z is a bisimulation.

Suppose not, then there are $s = \langle w, ..., w_n, c\rangle \in Unr(M, w)$ and $t = \langle |w|, ..., |w_n|, c\rangle \in Unr(M^n, |w|)$ which violate at least one of the three conditions of bisimulation. It is clear that $V(s) = V(t)$. Suppose $sR^r_i s'$ for some $s' \in Unr(M, w)$, namely $s' = \langle w, ..., w_n, v, ci\rangle$[7]. Then from the definition of R^n_c, there must be a $t' = \langle |w|, ..., |w_n|, |v|, ci\rangle \in Unr(M^n, w^n)$ such that $tR^{nr}_i t'$. In such case, the back-condition of bisimulation must be violated. That is to say, there is a t' such that $tR^{nr}_i t'$ for some $i \in I$, but there is no $s' \in Unr(M, w)$ such that $sR^r_i s'$ and $s'Zt'$. Namely there is $t' = \langle |w|, ..., |w_n|, |v|, ci\rangle \in Unr(M^n, |w|)$ but there is no $s' = \langle w, ..., w_n, v^*, ci\rangle$ in $Unr(M, w)$ where $v^* \in |v|$. Since $t' = \langle |w|, ..., |w_n|, |v|, ci\rangle$, according to the definition of R^n_c, in (M, w) there is a ci-path from a $w' \in |w|$ to a $v' \in |v|$ which extends a c-path from w' to a $w'_n \in |w_n|$. Since (M, w) is a pruned pointed model, then all the paths in (M, w) are from w. It follows that there is a ci-path from w to a v' which extends a c-path from w to a w'_n. Evidently,

[5] $w \leftrightsquigarrow_\Phi v \iff V(w) = V(v)$.

[6] Otherwise we can make it pruned and the formulas satisfiable at w will not be changed.

[7] Remember we use R^r to denote the relation in unravelling models. Here we use R^r_i as relations in $Unr(M, w)$ and R^{nr}_i as relations in $Unr(M^n, w^n)$ for each $i \in I$.

$M, w \vDash \Diamond_c(\pi \wedge \Diamond_i \pi')$ where π coincides $V(|w_n|) = V(w_n)$ and π' coincides $V(v')$. However, since there is no such $\langle w, ..., w_n, v^*, ci\rangle \in Unr(M, w)$ for any $v^* \in |v|$ while $s = \langle w, ..., w_n, c\rangle$ exists, then it is easy to see that $M, w \vDash \Diamond_c(\pi \wedge \neg\Diamond_i \pi')$. Therefore $M, w \vDash \Diamond_c(\pi \wedge \neg\Diamond_i \pi') \wedge \Diamond_c(\pi \wedge \Diamond_i \pi')$ which contradicts to the assumption that $M, w \vDash \neg\Delta^*$. That is to say, Z is indeed a bisimulation, which completes the proof.

$\Leftarrow$: Similar to the $\Leftarrow$ proof in part (a).

QED

Theorem 3.1.3 ***(Non-redundancy Criteria for finite Φ)***

- *(a) Given a set of $\mathcal{L}^I(\Phi)$ formulas Γ, Γ has a non-redundant classical model $\iff$ $\Gamma \cup \neg\Delta$ has a classical model.*
- *(b) Given a set of $\mathcal{L}^I(\Phi)$ formulas Γ, Γ has a non-redundant indexed model $\iff$ $\Gamma \cup \neg\Delta^*$ has an indexed model.*

Proof. For (a):

$\Rightarrow$: Given a set of $\mathcal{L}^I(\Phi)$ formulas Γ, suppose Γ has a non-redundant classical model (M, w). We claim that $\neg\Delta$ is also satisfiable on (M, w). Suppose not, then there is a formula $\phi = \Diamond_c(\pi \wedge \Diamond_i \pi') \wedge \Diamond_d(\pi \wedge \neg\Diamond_i \pi') \in \Delta$ such that $M, w \Vdash \phi$. However since (M, w) is non-redundant then there is at most one world which satisfies π, suppose it is v, but $M, v \Vdash \Diamond_i \pi' \iff M, v \nVdash \neg\Diamond_i \pi'$. Then $M, w \nVdash \Diamond_c(\pi \wedge \Diamond_i \psi) \wedge \Diamond_d(\pi \wedge \Diamond_i \psi)$. Contradiction.

$\Leftarrow$: Suppose there is a model for $\Gamma \cup \neg\Delta$, then from Lemma 3.1.2 we have Γ has an non-redundant model.

For (b):

$\Rightarrow$: Easy, similar to (a).

$\Leftarrow$: Suppose there is a model for $\Gamma \cup \neg\Delta^*$. From Lemma 3.1.2 we have Γ has an non-redundant model.

QED

Remark 3.1.4 *Let's take a look at the formulas in Δ and Δ^*. We now call a π as a* primitive state*, since it is the full description of the primitive factual propositions. Then the formula $\Diamond_c(\pi \wedge \Diamond_i \pi') \wedge \Diamond_c(\pi \wedge \neg\Diamond_i \pi')$ roughly says that: the (imaginary)agent c thinks it is possible that the primitive state is π but he is not sure whether agent i would consider π' possible when at* *primitive state* π. *And the formula* $\Diamond_c(\pi \wedge \Diamond_i\pi') \wedge \Diamond_d(\pi \wedge \neg\Diamond_i\pi')$ *says that: the imaginary agent c thinks it is possible that the primitive state is* π *and agent i considers* π' *possible at such primitive state but the imaginary agent d thinks it is possible that agent i doesn't consider* π' *possible at such primitive state* π.

Based on the above interpretations of formulas, the above theorem shows that:

- *We can have a non-redundant classical model for a situation if and only if in that situation, any two (imaginary)agents don't have different opinions on whether an agent j considers a primitive state* π' *possible when j is at primitive state* π.

- *We can have a non-redundant indexed model for a situation, if in that situation, any (imaginary)agent doesn't have uncertainties about whether an agent j considers a primitive state* π' *possible when at a possible primitive state* π. *In other words, if every primitive state can determine the (imaginary)agents' attitudes towards all the primitive states, then we can have a non-redundant indexed model for it. Although this determinacy criterion looks very strong, there are still lots of interesting situations like Muddy Children*[8] *and many card games satisfying it. For example, consider the muddy children example mentioned in Chapter 1. Both children are sure that at any primitive state how the other would think, although child 2 falsely believes child 1 is just as normal as he is. For instance, child 2 thinks that at primitive state* $< dirty, clean >$ *child 1 would thinks* $< clean, clean >$ *and* $< dirty, clean >$ *both possible. Though child 2 is wrong, he still doesn't have any doubts(uncertainties) about child 1's attitude towards the primitive states at any given primitive state.*

The criterions for having non-redundant indexed model is weaker and more reasonable than the one for having non-redundant classical model. This shows that the indexed model/semantics approach is more suitable if we are dealing with some specific situations or have strong preference to use non-redundant models. Modelling interactive unawareness is a good application for indexed non-redundant models. We will discuss it in the next section.

[8]Even with wrong information as we mentioned in the first chapter.

3.2 Indexed Model with Awareness

The "awareness" that we want to discuss in this section corresponds to the details of the primitive possibilities agents may consider. Suppose the primitive possibilities with full details are described by all the propositions in Φ. Then for any agent who is only aware of $\Phi' \subset \Phi$, his unawareness intuitively corresponds to filtering or collapsing certain full-detail possibilities into to one. According to such intuition, several semantic approaches based on the state-space have been proposed recently[MR99][HMS03][Sad05]. However, under some constraints, we think the projections and multiple levels in their models can be encoded in a simpler way via indexedmodels. We now define our own awareness models based on indexed models.

Definition 3.2.1 ***(Indexed Models with Awareness)*** *An indexed model with awareness is a tuple:*

$$M = (W, \{(R_c, A_c)\}_{c \in S(I)}, V)$$

where:

- *W is a non-empty set of possible worlds.*
- *A_c is a subset of Φ.*
- *R_c is a relation on $W \times W$.*
- *$V : W \to \mathcal{P}\Phi$ is a valuation on W.*

Such that for any $w, v \in W$, $V(w) = V(v) \iff w = v$.

Remark 3.2.2

- *It is clear that an indexed model with awareness can be seen as a pair $(M^*, \{A_c\}_{c \in S(I)})$ where M^* is a non-redundant indexed model. We can choose the preferable underlying indexed models depending on the purpose*[9]*.*
- *The intended interpretations of R_c relations are as before.*

[9]For example, to deal with belief and awareness, we'd better use the indexed models which are restricted I-euclidean, restricted I-transitive and I-serial.

- *A_c represents the awareness set of the (imaginary) agent c. We allow A_c to be $\varnothing$ in some cases. Notice that the awareness set in our model is world-independent. Namely, the awareness is uniform for an (imaginary) agent at any world. It follows that in our model, we don't allow (imaginary)agents to have uncertainties about others' awareness ability. For example, j's awareness set is certain in agent i's mind, although agent i might be wrong. As Remark 3.1.4 discussed, by using such non-redundant models, we are aiming at the situations in which agents only have uncertainties about primitive possibilities.*

Notation In the following, we call the indexed models with awareness "*IA-models*".

According to our intuition, If an agent is not aware of p, then he can not distinguish two primitive possibilities which only differ in the truth value of p. To represent this explicitly, we should add some constraints on the IA-models. First of all, we define the indistinguishable relation $\sim_c$ w.r.t to an awareness set A_c as following:

Definition 3.2.3 ***(Indistinguishable relations)*** *Given an IA-model $M = (W, \{(A_c, R_c) | c \in S(I)\}, V)$, for any $c \in S(I)$, $\sim_c$ is a relation on $W \times W$:*

$$\sim_c = \{(w, v) | (V(w) \cap A_c) = (V(v) \cap A_c)\}$$

Obviously, $\sim_c$ is an equivalence relation.

Let's define an useful concept here:

Definition 3.2.4 ***(A_c-bisimulation)*** *Let $M = (W, \{(R_c, A_c)\}_{c \in S(I)}, V)$ be an IA-model. A binary relation $Z \subseteq W \times W$ is said to be an A_c-bisimulation in M if the following conditions are satisfied:*

- *If wZv then $p \in V(w) \iff p \in V(v)$ for all $p \in A_c$.*
- *If wZv and $wR_{ci}w'$ for some $i \in I$, then there is $v' \in W$ such that $vR_{ci}v'$ and $w'Zv'$.*
- *If wZv and $vR_{ci}v'$ for some $i \in I$, then there is $w' \in W$ such that $wR_{ci}w'$ and $w'Zv'$.*

We say that two worlds v, w of M are A_c-bisimilar(notation $w \leftrightarrows_{A_c} v$), if vZw and Z is an A_c-bisimulation on M.

There are several intuitive basic constraints on IA-models:

1. **Limited Awareness:** $A_{ci} \subseteq A_c$ for any $c \in S(I)$.

   The idea behind this constraint is that one agent can not really imagine that others could be aware of what he is unaware of[10].

2. **Consistency for Indistinguishable States:** For any $c \in S(I)$, $w, v, t \in W$:
   1. if $w \sim_c v$ then $(t, w) \in R_c \iff (t, v) \in R_c$,
   2. $\sim_c$ is an A_c-bisimulation.

   The idea behind this is that if agent can not distinguish two primitive states w, v then:
   1. At any world, he thinks $w\ iff$ he thinks v possible.
   2 He has the equal imaginations for others on these two worlds.
   In sum we can identify the worlds with the same A_c for (imaginary)agent c.

Like many properties about indexed models that we mentioned in the last chapter, there are some stronger constraints about awareness sets. Those stronger constraints don't hold in general, but may be useful depending on certain purpose. We now list some of them:

- Stronger alternatives of **Limited Awareness**:

  1. **Positive Introspection of Awareness:** $A_{cii} = A_{ci}$ for any $c \in S(I)\{\epsilon\}, i \in I$.
     The idea behind this is that agents have the correct reflection about their own awareness sets, and they also assume everybody does so.
  2. **Weak Correctness of Imaginary Awareness:** $A_{ci} \subseteq A_c \cap A_i$ for any non-empty sequence $c \in S(I)$ and $i \in I$.
     The idea behind this is that we assume agents could only think others are less or equally aware than they actually are. For example, teachers always think students are aware of less things than

[10]Although he may doubt about this.

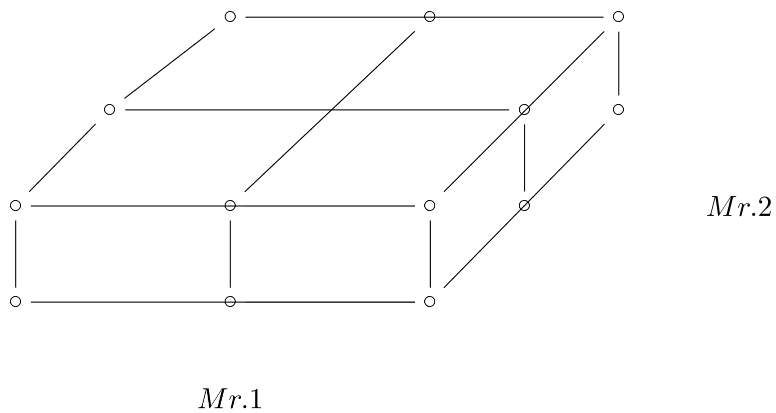
they actually are. On the other hand, this stronger constraint excludes the possibilities in which agents can be wrong about others' awareness abilities. For example, students always think teachers are aware of more things than they actually are.

3. **Strict Correctness of Imaginary Awareness:** $A_{ci} = A_c \cap A_i$ for any non-empty sequence $c \in S(I)$ and $i \in I$.

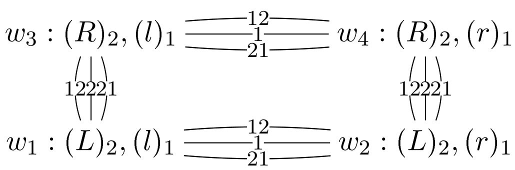
   The idea behind this is that we assume the agents can do their best in guessing correctly about others' awareness. Such constraint doesn't hold in general, but may be helpful when modelling many situations in which a rather small Φ is considered and agents know each other pretty well.

Now take a look at an example which often appears in the multi-context systems.

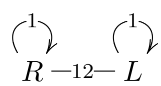
Example 3.2.5 *There are two agents Mr.1 and Mr.2 who are looking from different sides of the "magic box" as the following picture shows. The box is called "magic" since the agents only can see if there is a ball in some columns but can not tell the depth where the ball actually is. Now put a ball in the box, then both agents have asymmetric and partial information about the position of the ball. One thinks the other won't see anything useful at the other position(since they can not see the depth).*



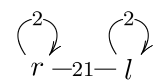
Let $\Phi = \{L, l\}$ where L means "there is a ball at the left-hand side from the view of Mr.1", while l means that "there is a ball at the left-hand side from the view of Mr.2". Since there is only one ball in the box, then if you don't see the ball at the left-hand side then you must see it at the right-hand side so we can denote proposition letter R as $\neg L$, and r as $\neg l$. According to the scenario, $A_1 = \{L\}$ while $A_2 = \{l\}$ and $A_{12} = A_{21} = \varnothing$. Then $A_{12c} = A_{21c}$ must be $\varnothing$ for any $c \in S(I)$ according to **Limited Awareness**. According to **Positive Introspection of Awareness** we have $A_{1..1} = A_1 = \{L\}$ and $A_{2..2} = A_2 = \{l\}$. Moreover, $A_{1..12} = A_{12} = A_{21} = A_{2..21} = \varnothing$. So far we have defined A_c for any non-empty $c \in S(I)$. For possibility relations: since $A_{1..12c} = A_{2..21c} = \varnothing$ for any $c \in S(I)$ from **Consistency for Indistinguishable States** we have $R_{12} = R_{21} = \{(w, v)|w, v \in W\}$. Let $R_{1..1} = R_1$ and $R_{2..2} = R_2$. Now we only need to define the relation R_1 and R_2. According to the scenario, we have the following model[11]:



where the proposition letters in the brackets labelled by i represent the unawareness set of agent i. It is easy to see that we can generate the sub-model from it for each agent c by identifying worlds according to the equivalence relation A_c. For example, for agent 1 we have:

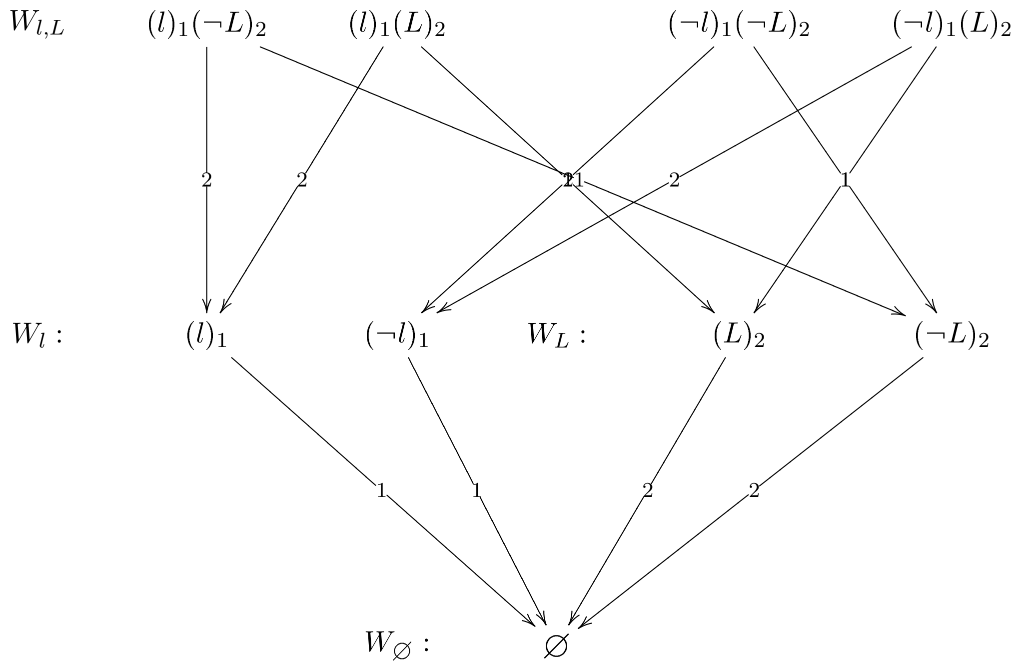


For agent 2 we have :



Remark 3.2.6 *We can actually transform the above IA-model into a HMS-style multi-state model as:*

[11] Where the relations are equivalence relation. And we omit the higher-order relations which are not interesting.



where A_c in the IA-model determines the projections between multiple levels and the relations for imaginary agents in the IA-model can be translated into the possibility relations for agent within each level. We now leave the precise correspondence results to the further study.

From this example, we can see that the non-redundant indexed model with awareness gives us a compact way of modelling interactive unawareness, under the constraint that agents only have uncertainties about primitive possibilities.

Chapter 4

Conclusion

This thesis aimed to propose an alternative semantics for multi-agent doxastic/epistemic logics. We started from the intuition that when thinking about others' information states, we actually have some "imaginary agents" in our mind, who may differ from the actual persons in the possibility relations they have. By including extra possibility relations for each imaginary agent, we can explicitly represent such imaginations of agents about each other in the so called indexed models. Accordingly, the indexed semantics interprets nested-modality formulas context-dependently to capture the intuition about the imaginary agents. In the indexed semantics, a modal operator has different explicit meanings under different scopes of other operators. The similar idea of context dependent Kripke semantics also appears in some recent works [Gab02][Gab04][BE06], where the meaning of a modality in a formula sometimes depends on the path of worlds we passed to evaluate that formula.

We have shown some advantages of our indexed semantics approach. The explicit imaginary relations make modelling much easier. We have proved that we can always have a desired non-redundant indexed model for the situations in which agents don't have uncertainties about others' possibility relations. On the other hand, the criteria for having a non-redundant classical model is much stronger and not very intuitive. By taking the advantage of such non-redundant indexed model, we give an intuitive and succinct way for modelling interactive unawareness under some constraints. As we have shown, the generalization and uniform substitution no longer preserve validity on all the class of indexed frames. Sometimes, this is useful, for we can have more subtle axioms in the logic. Some complete logic systems have been given which can be used for different purposes.

However, we have to admit that the indexed model/semantics approach is not the ideal alternative of the classical possible-worlds model/semantics. It is most useful and best understood in modelling situations in which agents don't have higher order uncertainties. Actually, the problem lies in the way we think of the imaginary agents. In fact, if agent 1 has uncertainties about how agent 2 considers the primitive possibilities, then there are actually more than one imaginary agents in agent 1's mind, each has its own certain information. This suggests a possible way to keep the non-redundant model and deal with higher-order uncertainties at the same time, namely, to add more imaginary agents. For example, if agent 1 thinks there are two possible "versions" 2.1 and 2.2 of agent 2, then we could revise the semantics to interpret any formula in the shape of $\Box_1\Box_2\phi$ as $\Box_1(\Box_{2.1}\phi \wedge \Box_{2.2}\phi)$.

Moreover, we should notice that although we indeed reduce the size of models on one hand, we also have to pay the price for adding complex relations on the other hand. The logic without GEN and SUB looks complicated and it is hard to obtain straightforward completeness proofs.

What we have explored in this thesis, are just the basic results about indexed model and semantics. We hope those results could help people to make use of the old tools to solve new questions in our approach. Actually, there is much more to discuss and compare with the classical approach. Many topics in classical approach are interesting to be reconsidered in the new approach. For example, here are some interesting questions that should not be very hard to answer:

- Can we find a straightforward variation of bisimulation for indexed models?
- Can we find a way of building the "indexed canonical model"?
- How do we add common knowledge?
- Are all the important classical properties of frames still definable by modal formulas in the indexed set-up?[1]
- To give a intuitive interpretations for the first-order correspondents of T_i, 4_i and 5_i.
- To give a suitable semantics for indexed model with awareness and compare it with the HMS approach.

[1] It seems classical transitivity is not definable in a modal formula.

- Is it easy to do updates on indexed models? It seems we can not eliminate points as we did for classical models since one point in an indexed model may stand for different possibilities for different agents.

All of the above questions deserve careful discussions which we hope to do in the future.

Bibliography

[BE06] Denis Bonnay and Paul Égré, *A non-standard semantics for inexact knowledge with introspection*, Manuscript Undersubmission. Available at http://jeannicod.ccsd.cnrs.fr/ijn_00000679, March 2006.

[BRV] Patrick Blackburn, Maarten de Rijke, and Yde Venema, *Modal logic*.

[FHV91] Ronald Fagin, Joseph Y. Halpern, and Moshe Y. Vardi, *A model-theoretic analysis of knowledge*, J. ACM **38** (1991), no. 2, 382–428.

[Gab02] Dov Gabbay, *A theory of hypermodal logics: Mode shifting in modal logic*, Journal of Philosophical Logic **31** (2002), no. 3, 211–243.

[Gab04] D. Gabbay, *Reactive kripke semantics and arc accessibility*, Initial Draft, July 2004.

[HMS03] Aviad Heifetz, Martin Meier, and Burkhard C. Schipper, *Interactive unawareness and speculative trade*, Bonn econ discussion papers, University of Bonn, Germany, July 2003.

[MR99] Salvatore Modica and Aldo Rustichini, *Unawareness and partitional information structures*, Games and Economics Behavior **27** (1999).

[Sad05] Tomasz Sadzik, *Kownledge awareness and probabilistic beliefs*, Unpublished Manuscript, November 7 2005.