On the Tractability of Comparing Informational Structures*

Cédric Dégremont[1], Lena Kurzen[2], and Jakub Szymanik[3]

[1, 3]Institute of Artificial Intelligence, University of Groningen
[2]Institute for Logic, Language and Computation, University of Amsterdam

{cedric.uva | lena.kurzen | jakub.szymanik}@gmail.com

1   Introduction

Epistemic modal logics and their extensions are concerned with global and abstract problems in reasoning about information. One of the features of that approach is its struggle for flexibility: it aims at designing logical systems that can model a large variety of epistemic scenarios [9, 4]. Hence, it is not surprising that the trade-off between expressivity and complexity has been one of the central problems in the epistemic logic literature. Logics need to be quite complex to account for a wide range of problems and it is not a surprise that there are many intractability results in the literature (see e.g., [13] and [5] for a survey).

One of the aims of this paper is to initiate the mapping of the tractability border among the epistemic tasks rather than epistemic logics. As a result, we can identify a theoretical threshold in the difficulty of reasoning about information, as was already done in the context of reasoning with quantifiers (see [19, 20]). In order to do this, we shift our perspective: Instead of investigating the complexity of a given logic that may be used to describe a problem, we turn towards a complexity study of that concrete problem itself, determining what computational resources are needed in order to perform the reasoning. Focusing on specific problems, things may be much easier since concrete problems involved in the study of multi-agent interaction are rarely as general as e.g. satisfiability. In most cases, checking whether a given property is satisfied in a given (minimal) epistemic scenario is sufficient. Hence, many problems turn out to be tractable. Still, we will see that even in this perspective there are some intractable problems. This feasibility border in epistemic tasks seems to be an interesting new topic for a formal study. Moreover, in principle the cognitive plausibility of the border could be empirically assessed by checking whether it correlates with the difficulties faced by human subjects (cf. [23, 21]). So in a sense, we aim to initiate a search for an appropriate perspective and complexity measures that describe in plausible ways the cognitive difficulties agents face while interacting. Certain experimental results in the economics literature [24, 10] explore similar directions. In general, the approach we have described in this paper focuses exclusively on the abstract information structure leaving out any concept of preferences and strategic reasoning.

In this paper we investigate the computational complexity of various decision problems that are relevant for interactive reasoning in epistemic modal logic frameworks. In particular, we explore the complexity of manipulating and comparing information structures possessed by different agents. For instance, we are interested in how difficult it is to answer the following questions.

- Is one agent's information strictly less refined than another agents' information?

- Do two agents have the same knowledge/belief about each other's knowledge/belief?

- Given two agents, is it possible to give some information to one of them such that as a result

    - both agents have similar information structures? (cf. [22].)

    - one of them has more refined information than the other?

For determining the computational complexity of the different problems, we use complexity results from graph theory (see e.g. [12]). Thus, we also clarify the computational impact of assuming S5 accessibility relations in epistemic models, i.e., the impact of assuming partition-based information structures on the complexity of various problems.

After giving the preliminaries in Section 2, we discuss four types of epistemic tasks and their computational complexity: informational similarity (Section 3.1), informational symmetry (Section 3.2) and two kinds of informational manipulation (Section 3.3 and 3.4). Omitted proofs can be found in the appendix. Section 4 concludes.

2 Preliminaries

2.1 Modeling information

We use relational structures from epistemic logic for modeling information (cf. [6, 9]). Kripke models can compactly represent the information agents have about the world and about the information possessed by the other agents. In what follows, $N = \{1, \ldots, n\}$ is a fixed finite set of agents and PROP is a countable set of propositional variables.

Definition 2.1 (Kripke Models). A *Kripke model* $\mathcal{M}$ based on a set of agents N is of the form $(W, (R_i)_{i \in N}, V)$, where $W \neq \emptyset$, for each $i \in N$, R_i is a binary relation on W, and $V : \text{PROP} \to \wp(W)$.

It is frequently assumed that information structures are partition-based [1, 9, 16]:

Definition 2.2 (Epistemic Models). An *epistemic model* is a Kripke model such that for all $i \in N$, R_i is an equivalence relation. (We usually write $\sim_i$ instead of R_i).

We write $|\mathcal{M}|$ to refer to the size of the model $\mathcal{M}$, and $Dom(\mathcal{M})$ to refer to the domain of $\mathcal{M}$. We refer to a pair $(\mathcal{M}, w)$ with $w \in Dom(\mathcal{M})$ as a pointed model. Intuitively R_i encodes i's uncertainty: if sR_it, then if the actual world were s then i would consider it possible that the actual world is t. For any non-empty set $G \subseteq N$, we write R_G^* for the reflexive transitive closure of $\bigcup_{i \in G} R_i$.

2.2 Comparing models and reasoning about submodels

In what follows, we need a reasonable notion of two models being similar. In addition to the notion of isomorphism, we make use of the notions of simulation, simulation equivalence and bisimulation.

Definition 2.3 (Simulation). We say that a pointed Kripke model $(\mathcal{M}, s)$, where $\mathcal{M} = (W, (R_i)_{i \in N}, V)$ and $s \in W$, is simulated by another pointed model $(\mathcal{M}', s')$ (which we denote by $(\mathcal{M}, s) \sqsubseteq (\mathcal{M}', s')$) such that $\mathcal{M}' = (W', (R_i')_{i \in N}, V')$ with $s' \in W'$ if there exists a binary relation $Z \subseteq W \times W'$ such that sZs' and for any two states x, x' whenever xZx' then for all $i \in N$:

1. x, x' verify the same proposition letters.

2. if xR_iz in $\mathcal{M}$ then there exists some $z' \in W'$ with $x'R_i'z'$ and zZz'.

We say that $\mathcal{M} = (W, (R_i)_{i \in N}, V)$ is simulated by $\mathcal{M}' = (W', (R_i')_{i \in N}, V')$ (denoted by $\mathcal{M} \sqsubseteq \mathcal{M}'$) if there are $s \in W$ and $s' \in W'$ such that $(\mathcal{M}, s) \sqsubseteq (\mathcal{M}', s')$. We say that a simulation $Z \subseteq W \times W'$ is *total* if for every $s \in W$, there is some $t \in W'$ such that sZt, and for every $t \in W'$, there is some $s \in W$ such that sZt. If $\mathcal{M}$ is simulated by $\mathcal{M}'$ by means of a total simulation, we say $\mathcal{M} \sqsubseteq_{total} \mathcal{M}'$. Moreover, we say that $\mathcal{M} = (W, (R_i)_{i \in N}, V)$ and $\mathcal{M}' = (W', (R_i')_{i \in N}, V')$ are simulation equivalent if $\mathcal{M}$ simulates $\mathcal{M}'$ and $\mathcal{M}'$ simulates $\mathcal{M}$. The following notion is stronger than simulation equivalence.

Definition 2.4 (Bisimulation)**.** A local bisimulation between two pointed Kripke models with set of agents N, $(\mathcal{M}, s)$ with $\mathcal{M} = (W, (R_i)_{i \in N}, V)$ and $(\mathcal{M}', t)$ with $\mathcal{M}' = (W', (R_i')_{i \in N}, V')$ is a binary relation $Z \subseteq W \times W'$ such that sZs' and also for any worlds x, x' whenever xZx' then for all $i \in N$:

1. x, x' verify the same proposition letters.

2. if xR_iu in $\mathcal{M}$ then there exists $u' \in W'$ with $x'R_i'u'$ and uZu'.

3. if $x'R_i'u'$ in $\mathcal{M}'$ then there exists $u \in W$ with xR_iu and uZu'.

We say that $\mathcal{M} = (W, (R_i)_{i \in N}, V)$ and $\mathcal{M}' = (W', (R_i')_{i \in N}, V')$ are bisimilar ($\mathcal{M} \underline{\leftrightarrow} \mathcal{M}'$) if there are $s \in W$ and $s' \in W'$ such that $(\mathcal{M}, s) \underline{\leftrightarrow} (\mathcal{M}', s')$. A bisimulation $Z \subseteq Dom(\mathcal{M}) \times Dom(\mathcal{M}')$ is *total* if for every $s \in Dom(\mathcal{M})$, there is some $t \in Dom(\mathcal{M}')$ such that sZt, and for every $t \in Dom(\mathcal{M}')$, there is some $s \in Dom(\mathcal{M})$ such that sZt. Then we write $\mathcal{M} \underline{\leftrightarrow}_{total} \mathcal{M}'$.

To reason about informational structures that can be obtained by providing agents with new information, we use the notions of submodel and generated submodel.

Definition 2.5 (Submodel)**.** We say that $\mathcal{M}'$ is a submodel of $\mathcal{M}$ iff $W' \subseteq W$, $\forall i \in N$, $R_i' = R_i \cap (W' \times W')$, $\forall p \in \text{PROP}$, $V'(p) = V(p) \cap W'$.

The notion of *induced subgraph* is just like that of a submodel without the condition for the valuations. The notion of *subgraph* is weaker than that of an induced subgraph as it allows that $R_i' \subset R_i \cap W' \times W'$).

Definition 2.6 (Generated submodel)**.** We say that $\mathcal{M}' = (W', (R_i)'_{i \in N}, V')$ is a generated submodel of $\mathcal{M} = (W, (R_i)_{i \in N}, V)$ iff $W' \subseteq W$ and $\forall i \in N$, $R_i' = R_i \cap (W' \times W')$, $\forall p \in \text{PROP}$, $V'(p) = V(p) \cap W'$ and if $w \in W'$ and wR_iv then $v \in W'$. The submodel of $\mathcal{M}$ generated by $X \subseteq W$ is the smallest generated submodel $\mathcal{M}'$ of $\mathcal{M}$ with $X \subseteq Dom(\mathcal{M}')$.

We write $\mathcal{K}_i[w] := \{v \in W \mid wR_iv\}$ to denote i's information set at w and $R_G^*[w] := \{v \in W \mid wR_G^*v\}$. This notion is generalized by the concept of horizon:

Definition 2.7 (Horizon)**.** The *horizon* of i at $(\mathcal{M}, w)$ (notation: $(\mathcal{M}, w)^i$) is the submodel generated by $\mathcal{K}_i[w]$.

This paper will not use syntactic notions. In terms of intuition, the important definition is that of knowledge K_i: agent i knows ϕ at w if ϕ is true in all states that i considers possible at w. In equivalent semantic terms: i knows E if $E \subseteq \mathcal{K}_i[w]$. E is common knowledge in a group G at w iff $E \subseteq R_G^*[w]$.

2.3 Tractability

Some problems, although computable, nevertheless require too much time or memory to be feasibly solved by a realistic computational device. Computational complexity theory investigates the resources (time, memory, etc.) required for the execution of algorithms and the inherent difficulty of computational

problems [17]. In particular, we want to identify efficiently solvable problems and draw a line between tractability and intractability. In general, the most important distinction is that between problems which can be computed in polynomial time with respect to their size, and those which are believed to have only exponential time algorithmic solutions. The class of problems of the first type is called PTIME (P for short); one can demonstrate that a problem belongs to this class if one can show that it can be computed by a deterministic Turing machine in polynomial time. Problems belonging to the second class are referred to as NP-hard. They are at least as difficult as problems belonging to the NPTIME (NP) class; this is the class of problems which can be computed by nondeterministic Turing machines in polynomial time. NP-complete problems are NP-hard problems belonging to NPTIME, hence they are intuitively the most difficult problems among the NPTIME problems.

3 Complexity of comparing and manipulating information

3.1 Information similarity

The first natural question we would like to address is whether an agent in a given situation has similar information to the one possessed by some other agent (in a possibly different situation). One very strict way to understand such similarity is through the use of isomorphism.

For the general problem of checking whether two Kripke models are isomorphic, we can give tight complexity bounds, as this problem is polynomially equivalent to graph isomorphism. The graph isomorphism problem is neither known to be NP-complete nor to be tractable and the set of problems with a polynomial-time reduction to the graph isomorphism problem is called GI.

Decision Problem 3.1 (Kripke model isomorphism)**.**
Input: *Pointed Kripke models* $(\mathcal{M}_1, w_1)$, $(\mathcal{M}_2, w_2)$.

Question: *Are* $(\mathcal{M}_1, w_1)$ *and* $(\mathcal{M}_2, w_2)$ *isomorphic, i.e. is it the case that* $(\mathcal{M}_1, w_1) \cong (\mathcal{M}_2, w_2)$?

Fact 3.2. *Kripke model isomorphism is GI-complete.*

However, isomorphism is arguably a too restrictive notion of similarity. Bisimilarity is a weaker but still a very natural concept of similarity for relational structures. Here the question arises as to whether working with $S5$ models – a common assumption in the epistemic logic and interactive epistemology literature – rather than arbitrary Kripke structures has an influence on the complexity of the task.

Decision Problem 3.3 (Epistemic model bisimilarity)**.**
Input: *Two pointed multi-agent epistemic S5 models* $(\mathcal{M}_1, w_1), (\mathcal{M}_2, w_2)$.

Question: *Are the two models bisimilar, i.e.* $(\mathcal{M}_1, w_1) \underline{\leftrightarrow} (\mathcal{M}_2, w_2)$?

In [3], it has been shown that deciding bisimilarity is P-complete for finite labelled transition systems. It follows that epistemic models bisimilarity is also in P.

Fact 3.4. *Multi-agent epistemic S5 model bisimulation can be done in polynomial time with respect to the size of the input* $(|\mathcal{M}_1| + |\mathcal{M}_2|)$.

Thus, multi-agent epistemic S5 model bisimilarity is in P. Now, of course the question arises if it is also P-hard.[1]
Open problem *Is multi-agent epistemic model (S5) bisimulation P-hard?*

---

[1] We conjecture that we can show P hardness using methods of simulating an arbitrary relation R by a combination of two equivalence relations $\sim_1$ and $\sim_2$ as follows: we replace each wRv by $w \sim_1 z \sim_2 v$, for a new state z.

Without any assumptions on the accessibility relations for the agents, we immediately get P-completeness for multi-agent Kripke models as the problem is equivalent to bisimilarity for finite labelled transition systems.

The picture. Deciding whether two models are bisimilar is tractable for S5 epistemic models, and in case of the arbitrary Kripke structures it is among the hardest tractable problems. Kripke model isomorphism lives on the tractability border. It is open whether isomorphism for (partition-based) epistemic models is tractable and whether epistemic S5 model bisimilarity is P-complete, which we indeed conjecture to be the case.

3.2 Informational symmetry: knowing what others know

The preceding notions of similarity are very strong. In the context of analyzing epistemic interactions between agents, weaker notions of similarity are of interest. In general, the information that agents have about each other's information state plays a crucial role. We will now analyze the problem of deciding whether two agents' views about the interactive epistemic structure, and in particular about the knowledge of other agents, are equivalent. A first reading is simply to fix some fact $E \subseteq W$ and ask whether E is common knowledge in a group G. Clearly this problem is tractable.

Fact 3.5. *Given a pointed model $(\mathcal{M}, w)$, some $E \subseteq Dom(\mathcal{M})$ and $G \subseteq N$, deciding whether E is common knowledge in the group G at w can be done in polynomial time.*

Proof. From reachability for R_G^*. $\qquad\square$

However, instead of fixing some specific fact of interest, the question might be whether a situation is symmetric with respect to two given agents, say Alice and Bob. In other words, is the interactive informational structure from Alice's perspective similar to how it is from Bob's perspective?

Definition 3.6. We write $\mathcal{M}[i/j]$ to be the model obtained by switching labels between i and j.

Definition 3.7. We say that two pointed multi-agent epistemic models $(\mathcal{M}, s)$ and $(\mathcal{M}', s')$ (with set of agents N) are flipped bisimiliar for agents $i, j \in N$, $(\mathcal{M}, s)\underline{\leftrightarrow}_f^{(i,j)}(\mathcal{M}', s')$, iff $(\mathcal{M}, s)\underline{\leftrightarrow}(\mathcal{M}'[i/j], s')$.

A natural question is the relation of flipped bisimulation to the fact that all knowledge of both agents is common knowledge. The following is immediate:

Observation 3.8. *If in $\mathcal{M}$, $\sim_{\{1,2\}}^* \subseteq \sim_j$ for $j \in \{1, 2\}$, then for all $w \in Dom(\mathcal{M})$, $(\mathcal{M}, w)\underline{\leftrightarrow}_f^{(1,2)}(\mathcal{M}, w)$.*

Is other direction true? Locally, even on S5 models, flipped self-bisimulation is a much weaker requirement: it does not even imply that (shared) knowledge of facts is common knowledge:

Fact 3.9. *There exists a pointed S5 epistemic model which is a, b-flipped bisimilar to itself, where the two agents know that p (with $p \in \text{PROP}$), and p is not common knowledge between a and b.*

But required globally of every state, we do have the following converse:

Fact 3.10. *Let $\mathcal{M}$ be a transitive model with $w \in Dom(\mathcal{M})$. Whenever the submodel $\mathcal{M}'$ of $\mathcal{M}$ generated by $\{w\}$ is such that every state is $0, 1$-flipped bisimilar to itself, then for any $p \in \text{PROP}$, if j knows p at w (i.e., $V(p) \subseteq K_j[w]$) for some $j \in \{0, 1\}$ then p is common knowledge between 0 and 1 at w.*

Let us recall the notion of horizon (see Definition 2.7). It is the submodel generated by the information set of the agent: the *horizon* of i at (M, w) (notation: $(M, w)^i$) is the submodel generated by $\mathcal{K}_i[w]$.

Decision Problem 3.11 (Flipped multi-agent epistemic model horizon bisimilarity)**.**
Input: Two pointed multi-agent epistemic models $(\mathcal{M}, w)$, $(\mathcal{M}', w')$, two agents i, j.

Question: Are the horizons of agents i and j in $(\mathcal{M}, w)$ and $(\mathcal{M}', w')$ respectively flipped bisimilar for i, j, namely is it the case that: $(\mathcal{M}, w)^i \underline{\leftrightarrow} (\mathcal{M}', w')^j [i/j]$?

Fact 3.12. *Flipped multi-agent epistemic S5 model horizon bisimilarity is in P. Given a multi-agent epistemic model $(\mathcal{M}, w)$, it is trivial to decide if for two agents i, j it holds that $(\mathcal{M}, w)^i \underline{\leftrightarrow} (\mathcal{M}, w)^j [i/j]$.*

Proof. We can use a polynomial algorithm for Kripke model bisimilarity. Horizons of two agents at the same point in a model are always equal in S5 because of reflexivity of the accessibility relations. □

Fact 3.13. *Without any assumptions on the accessibility relations, the computational complexity of multi-agent Kripke model flipped horizon bisimilarity is P-complete.*

Proof. Follows from [3] and the fact that in general the horizons of two agents can be disjoint. □

The picture. Deciding horizon flipped bisimilarity in Kripke models is among the hardest tractable problems. It is trivial for partition-based models. Deciding whether a fact is commonly known is tractable.

3.3 Can we reshape an agent's mind into some desired informational state?

So far, we have been comparing agents' informational states within models. The next interesting problem is to decide whether new informational states (satisfying desired properties) can be achieved in certain ways. One immediate question is whether one can give some information to an agent (i.e. to restrict her horizon) such that after the update her horizon is bisimilar to the horizon of some other agent. Concretely, we would like to know if there is any type of information that could reshape some agent's information to fit some desired new informational state or at least be similar to it. We will thus investigate the task of checking whether there is a submodel that has certain properties. This means that we determine if it is possible to purposely refine a model in a certain way. This question is in line with problems addressed by arbitrary public announcement logic and arbitrary event modal logic [2, 11, 22].[2]

We start with the problem of checking whether there is a submodel of one model that is bisimilar to another one. On graphs, this is related to the problem of deciding if one contains a subgraph bisimilar to another. Note that in the problem referred to in the literature as "subgraph bisimulation" [8], the subgraph can be any graph whose vertices are a subset of the vertices of the original graph, and the edges can be any subset of the edges of the original graph restricted to the subset of vertices. To be more specific, the problem investigated in [8] is the following:

> Given two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, is there a graph $G_2' = (V_2', E_2')$ with $V_2' \subseteq V_2$ and $E_2' \subseteq E_2$ such that there is a total bisimulation between G_2' and G_1?

Since we want to investigate the complexity of reasoning about epistemic interaction using modal logic, we are interested in subgraphs that correspond to *relativization* in modal logic: induced subgraphs. This leads us to an investigation of *induced* subgraph bisimulation.

Decision Problem 3.14 (Induced subgraph bisimulation)**.**
Input: Two finite graphs $G_1 = (V_1, E_1)$, $G_2 = (V_2, E_2)$, $k \in \mathbb{N}$.

Question: *Is there an induced subgraph of G_2 with at least k vertices that is bisimilar to G_1, i.e. is there some $V' \subseteq V_2$ with $|V'| \geq k$ and $(V', E_2 \cap (V' \times V')) \underline{\leftrightarrow}_{total} G_1$?*

---

[2]Note that in the current work, we focus on the semantic structures only and do not require that the submodel can be characterized by some formula in a certain epistemic modal language.

Even though the above problem looks very similar to the original subgraph bisimulation problem (NP-hardness of which is shown by reduction from Hamiltonian Path), NP-hardness does not follow immediately.[3] Nevertheless, we can show NP-hardness by reduction from Independent Set.

Proposition 3.15. *Induced subgraph bisimulation is NP-complete.*

Now, an analogous result for Kripke models follows. The intuitive interpretation here (with an epistemic/doxastic interpretation of the accessibility relation) is whether it is possible to 'gently' restrict one model without letting its domain get smaller than k such that afterwards it is bisimilar to another model. The intuition is that we would like the new information to change as minimally as possible the informational state of the target agent.

Decision Problem 3.16 (Submodel bisimulation for Kripke models)**.**
Input: *Kripke models $\mathcal{M}_1$, $\mathcal{M}_2$ with set of agents N, $k \in \mathbb{N}$.*
Question: *Is there a submodel $\mathcal{M}_2'$ of $\mathcal{M}_2$ with $|Dom(\mathcal{M}_2')| \geq k$ such that $\mathcal{M}_1$ and $\mathcal{M}_2'$ are totally bisimilar i.e. $\mathcal{M}_1 \underline{\leftrightarrow}_{total} \mathcal{M}_2'$?*

Corollary 3.17. *Submodel bisimulation for Kripke models is NP-complete.*

As we are interested in the complexity of reasoning about the interaction of epistemic agents as it is modeled in (dynamic) epistemic logic, let us now see how the complexity of induced subgraph bisimulation changes when we make the assumption that models are partitional, i.e. that the relation is an equivalence relation, as it is frequently assumed in the AI or interactive epistemology literature. We will see that this assumption makes the problem significantly easier.

Proposition 3.18. *If for graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, E_1 and E_2 are reflexive, transitive and symmetric, then induced subgraph bisimulation for G_1 and G_2 can be solved in linear time.*

Assuming the edge relation in a graph to be an equivalence makes induced subgraph bisimulation a trivial problem because, unless its set of vertices is empty, every such graph is bisimilar to the graph $(\{v\}, \{(v,v)\})$. But for S5 models this is of course not the case, as the bisimulation takes into account the valuation. Nevertheless, we will now show that also for single agent S5 models, the problem of submodel bisimulation is significantly easier than in the case of arbitrary single agent Kripke models. To be more precise, we will distinguish between two problems:

The first problem is *local* single agent S5 submodel bisimulation. Here we take as input two pointed S5 models. Then we ask whether there is a submodel of the second model that is bisimilar to the first one. Thus, the question is whether it is possible to restrict one of the models in such a way that there is a state in which the agent has exactly the same information as in the situation modeled in the other model.

Decision Problem 3.19 (Local S5 submodel bisimulation for single agent epistemic models)**.**
Input: *A pointed S5 epistemic model $(\mathcal{M}_1, w)$ with $\mathcal{M}_1 = (W_1, \sim_1, V_1)$ and $w \in W_1$, and an S5 epistemic model $\mathcal{M}_2 = (W_2, \sim_2, V_2)$.*
Question: *Is there a submodel $\mathcal{M}_2' = (W_2', \sim_2', V_2')$ of $\mathcal{M}_2$ such that $(\mathcal{M}_1, w) \underline{\leftrightarrow} (\mathcal{M}_2', w')$ for some $w' \in Dom(\mathcal{M}_2')$?*

Proposition 3.20. *Local submodel bisimulation for single agent pointed epistemic models is in P.*

The second problem we consider is *global* S5 submodel bisimulation, where the input are two models $\mathcal{M}_1$ and $\mathcal{M}_2$ and we ask whether there exists a submodel of $\mathcal{M}_2$ such that it is totally bisimilar to $\mathcal{M}_1$.

---

[3]For Induced Subgraph Bisimulation, a reduction from Hamiltonian Path seems to be more difficult, as does a direct reduction from the original subgraph bisimulation problem.

Decision Problem 3.21 (Total S5 submodel bisimulation for single agent epistemic models)**.**
Input: Two S5 epistemic models $\mathcal{M}_1 = (W_1, \sim_1, V_1), \mathcal{M}_2 = (W_2, \sim_2, V_2)$.

Question: Is there a submodel $\mathcal{M}'_2 = (W'_2, \sim'_2, V'_2,)$ of $\mathcal{M}_2$ such that $\mathcal{M}_1 \underline{\leftrightarrow}_{total} \mathcal{M}'_2$?

We can show that even though the above problem seems more complicated than local submodel bisimulation, it can still be solved in polynomial time. The proof uses the fact that finding a maximum matching in a bipartite graph can be done in polynomial time (see e.g. [18]).

Theorem 3.22. *Total submodel bisimulation for single agent epistemic models is in P.*

Now, the question arises whether the above results also hold for the multi-agent case.

Decision Problem 3.23 (Global submodel bisimulation for multi-agent pointed epistemic models)**.**
Input: Two epistemic models $\mathcal{M}_1 = (W_1, (\sim_{1i})_{i \in N}, V_1), \mathcal{M}_2 = (W_2, (\sim_{2i})_{i \in N}, V_2)$, for N being a finite set (of agents), and $k \in \mathbb{N}$.

Question: Is there a submodel $\mathcal{M}'_2 = (W'_2, (\sim'_{2i})_{i \in N}, V'_2)$ of $\mathcal{M}_2$ such that $\mathcal{M}_1 \underline{\leftrightarrow}_{total} \mathcal{M}'_2$?

We conjecture that using similar ideas to those outlined in footnote 1 to show that the above problem is NP-complete for models with at least two agents.

The picture. Induced subgraph bisimulation is intractable (NP-complete) and so is submodel bisimulation for arbitrary Kripke models. For S5 models, induced subgraph bisimulation is tractable, and so are local submodel bisimulation and total submodel bisimulation in the single agent case. We think that NP-completeness can be shown for the case of at least two agents.

3.4 Simulation vs Bisimulation

In dynamic systems with diverse agents, an interesting question is whether it is possible to give some information to one agent such that afterwards she knows at least as much as some other agent. This is captured by an asymmetric notion, that of simulation. With this difference, the question can be raised of the effect on tractability and intractability of requiring simulation versus requiring bisimulation. With this motivation, we would like to explore the problem of induced subgraph simulation.

Decision Problem 3.24 (Induced subgraph simulation)**.**
Input: Two finite graphs $G_1 = (V_1, E_1)$, $G_2 = (V_2, E_2)$, $k \in \mathbb{N}$.

Question: Is there an induced subgraph of G_2 with at least k vertices that is simulated by G_1, i.e., is there some $V' \subseteq V_2$ with $|V'| \geq k$ and $(V', E_2 \cap (V' \times V')) \sqsubseteq_{total} G_1$?

Proposition 3.25. *Induced subgraph simulation is NP-complete.*

In [7], it has been shown that given two graphs it is also NP-complete to decide if there is a subgraph (not necessarily an induced one) of one such that it is simulation equivalent to the other graph. Here, we show that this also holds if the subgraph is required to be an induced subgraph.

Decision Problem 3.26 (Induced subgraph simulation equivalence)**.**
Input: Two finite graphs $G_1 = (V_1, E_1)$, $G_2 = (V_2, E_2)$, $k \in \mathbb{N}$.

Question: Is there an induced subgraph of G_2 with at least k vertices that is similar to G_1, i.e. is there some $V' \subseteq V_2$ with $|V'| \geq k$ and $(V', E_2 \cap (V' \times V')) \sqsubseteq_{total} G_1$ and $G_1 \sqsubseteq_{total} (V', E_2 \cap (V' \times V'))$?

Proposition 3.27. *Induced subgraph simulation equivalence is NP-complete.*

As a corollary of the two previous propositions, we get that for arbitrary Kripke models both submodel simulation and submodel equivalence are NP-complete. We conjecture that for single agent S5, we can use similar methods as used in the proof of Theorem 3.22. Let us conclude with an interesting open question, as to whether the results from [15] also hold for epistemic models.

Open problem *Is deciding simulation (equivalence) of epistemic models at least as hard as deciding bisimilarity?*

The picture. Induced subgraph simulation and equivalence are both intractable (NP-complete). The same holds for Kripke model simulation (equivalence). It remains to be investigated for epistemic models.

4 Conclusions and Further Work

In this work, we have identified concrete epistemic tasks related to the comparison and manipulation of informational states of agents in possibly different situations. Interestingly, our complexity analysis shows that the preceding problems live on both sides of the border between tractability and intractability:

Problem	Tractable?	Comments
Kripke model isomorphism	unknown	in GI
Epistemic model bisimilarity	Yes	Conjecture: P-hard for ≥ 2 agents
Flipped horizon bisimilarity	Yes	P-complete for arbitrary models
Kripke submodel bisimulation	No	NP-complete for arbitrary models;
		in linear time for S5
Local S5 submodel bisimulation	Single agent: Yes	unknown
Total S5 submodel bisimulation	Single agent: Yes	Conjecture: NP-complete for ≥ 2 agents
Kripke submod. simulation (equiv.)	No	Conjecture: in P for single agent S5

Table 1: Summary of the results and open questions.

As such, this work is a first step towards mapping out the complexity of concrete epistemic problems based on epistemic modeling. It would be interesting to systematize this approach to a larger class of problems. Further work to complete the picture includes the open problems that we mentioned in our analysis in Section 3. Solving them would clarify the border between tractability and intractability in the domain of epistemic reasoning tasks. This would then also shed some light on the more general question as to what is the impact of the assumption of S5 on the complexity of certain problems from graph theory. It would moreover clarify whether for some epistemic tasks, moving from single agent to multi-agent scenarios has the consequence of crossing the border between tractability and intractability.

How would we like to interpret our results? One conclusion, we can draw from our case study is that assuming partition-based information structures simplifies epistemic tasks of comparing and manipulating informational structures. In particular, we saw that comparing agents's informational structures via bisimulation is tractable in the multi-agent case, meaning that it should be relatively easy to say whether Alice's information is strictly less refined than Bob's. Furthermore, deciding whether two agents have symmetric knowledge about each other's knowledge should be also in principle easy (PTIME for S5 models and P-complete for arbitrary models). Finally, we proved that things are getting harder if one wants to know wether a certain manipulation of agents' knowledge is possible. Deciding whether the information structure of Alice is more refined than that of Bob is in general intractable, independently of choosing bisimilarity or isomorphism as our notion of similarity. However, the problem becomes easy if one assumes that agents's knowledge can be modeled by equivalence relations. On the other hand,

substituting bisimulation by simulation gives rise to an interesting open problem whether computing simulation equivalence of epistemic models is a at least as hard as deciding their bisimilarity.

References

[1] Robert J. Aumann. Interactive epistemology I: Knowledge. *International Journal of Game Theory*, 28(3):263–300, 1999.

[2] Philippe Balbiani, Alexandru Baltag, Hans van Ditmarsch, Andreas Herzig, Tomohiro Hoshi, and Tiago de Lima. 'knowable' as 'known after an announcement'. Technical Report IRIT/RR-2008-2-FR, IRIT, University of Toulouse 3, 2008. URL: `ftp://ftp.irit.fr/IRIT/LILAC/2008_Report_Balbiani_et_al.pdf`.

[3] José L. Balcázar, Joaquim Gabarró, and Miklos Santha. Deciding bisimilarity is P-complete. *Formal Aspects of Computing*, 4(6A):638–648, 1992.

[4] Alexandru Baltag and Lawrence S. Moss. Logics for epistemic programs. *Synthese*, 139(2):165–224, 2004.

[5] Johan van Benthem and Eric Pacuit. The tree of knowledge in action: towards a common perspective. In I. Hodkinson G. Governatori and Y. Venema, editors, *Advances in Modal Logic*, volume 6. College Publications, 2006.

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

[7] Lorenzo De Nardo, Francesco Ranzato, and Francesco Tapparo. The subgraph similarity problem. *IEEE Transactions on Knowledge and Data Engineering*, 21(5):748–749, 2009.

[8] Agostino Dovier and Carla Piazza. The subgraph bisimulation problem. *IEEE Transactions on Knowledge and Data Engineering*, 15(4):1055–1056, 2003.

[9] Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. *Reasoning About Knowledge.* MIT Press, Cambridge, 1995.

[10] Nick Feltovich. Reinforcement-based vs. beliefs-based learning in experimental asymmetric-information games. *Econometrica*, 68:605–641, 2000.

[11] Tim French and Hans van Ditmarsch. Undecidability for arbitrary public announcement logic. In Carlos Areces and Robert Goldblatt, editors, *Advances in Modal Logic*, pages 23–42. College Publications, 2008.

[12] Michael R. Garey and David S. Johnson. *Computers and Intractability; A Guide to the Theory of NP-Completeness.* W. H. Freeman & Co., New York, NY, USA, 1990.

[13] Joseph Y. Halpern and Moshe Y. Vardi. The complexity of reasoning about knowledge and time. I. Lower bounds. *Journal of Computer and Systems Science*, 38(1):195–237, 1989.

[14] Monika R. Henzinger, Thomas A. Henzinger, and Peter W. Kopke. Computing simulations on finite and infinite graphs. In *FOCS '95: Proceedings of the 36th Annual Symposium on Foundations of Computer Science*, pages 453–462. IEEE Computer Society Press, 1995.

[15] Antonín Kučera and Richard Mayr. Why is simulation harder than bisimulation? In *CONCUR '02: Proceedings of the 13th International Conference on Concurrency Theory*, pages 594–610, London, UK, 2002. Springer-Verlag.

[16] Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory.* MIT Press, 1994.

[17] Christos H. Papadimitriou. *Computational Complexity.* Addison Wesley, November 1993.

[18] Christos H. Papadimitriou and Kenneth Steiglitz. *Combinatorial optimization: algorithms and complexity.* Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1982.

[19] Ian Pratt-Hartmann and Lawrence S. Moss. Logics for the relational syllogistic. *The Review of Symbolic Logic*, 2(04):647–683, 2009.

[20] Jakub Szymanik. Computational complexity of polyadic lifts of generalized quantifiers in natural language. *Linguistics and Philosophy*, 33:215–250, 2010.

[21] Jakub Szymanik and Marcin Zajenkowski. Comprehension of simple quantifiers. Empirical evaluation of a computational model. *Cognitive Science: A Multidisciplinary Journal*, 34(3):521–532, 2010.

[22] Hans van Ditmarsch and Tim French. Simulation and information: Quantifying over epistemic events. In *Knowledge Representation for Agents and Multi-Agent Systems: First International Workshop, KRAMAS 2008, Sydney, Australia, September 17, 2008, Revised Selected Papers*, pages 51–65, Berlin, Heidelberg, 2009. Springer-Verlag.

[23] Rineke Verbrugge. Logic and social cognition. The facts matter, and so do computational models. *Journal of Philosophical Logic*, 38(6):649–680, 2009.

[24] Roberto Weber. Behavior and learning the "dirty faces" game. *Experimental Economics*, 4:229–242, 2001.

A   Proofs of selected theorems

Fact: 3.9 There exists a pointed S5 epistemic model which is a, b-flipped bisimilar to itself, where the two agents knows that p (with $p \in \text{PROP}$), and p is not common knowledge between a and b.

Proof. Consider the model $\mathcal{M} = \langle W, \sim_a, \sim_b, V \rangle$ with $W = \{-2, -1, 0, 1, 2\}$, $\sim_a$ is the smallest equivalence relation on W containing $\{(-2, -1), (0, 1)\}$, $\sim_b$ is the smallest equivalence relation on W containing $\{(-1, 0), (1, 2)\}$, and $V(p) = \{-1, 0, 1\}$. It is easy to check that both Alice and Bob knows p at 0: $\mathcal{K}_a[0] = \{0, 1\} \subseteq V(p)$ and $\mathcal{K}_b[0] = \{-1, 0\} \subseteq V(p)$. Also p is not common knowledge between Alice and Bob at 0, indeed $0 \sim_a 1 \sim_b 2$ and $2 \notin V(p)$. Now it remains to show that $\mathcal{M}, 0$ is a, b-flipped bisimilar to itself. The flipped bisimulation is defined as $Z = \{(n, 0 - n) \mid n \in W\}$. It is easily checked by inspection that Z is indeed a a, b-flipped bisimulation. $\square$

Fact 3.10: Let $\mathcal{M}$ be a transitive model with $w \in Dom(\mathcal{M})$. If the submodel $\mathcal{M}'$ of $\mathcal{M}$ generated by $\{w\}$ is such that every state is $0, 1$-flipped bisimilar to itself. Then for any $p \in \text{PROP}$, if j knows p at w, i.e., $(V(p) \subseteq K_j[w]$, for some $j \in \{0, 1\}$, then p is common knowledge between 0 and 1) at w.

Proof. We prove the contrapositive. Assume that p is not common knowledge between 0 and 1 at w. It follows that we have a $0, 1$-path of length n with $n \in \omega$ of the form $wR_f(1)w_1 Rf(2) \ldots Rf(n-1)w_{n-1}$ with $w_{n-1} \notin V(p)$ and $f(k) \in \{0, 1\}$ for all $k \in n$. Clearly, all the states in the preceding sequence are in $\mathcal{M}'$ so they must be $0, 1$-flipped bisimilar to themselves and, in particular, to w. Hence, by definition of a flipped bisimulation we have a sequence of the form $wR_f^1(1)w_1^1 Rf^1(2) \ldots Rf^1(n-1)w_{n-1}^1$. Especially, $w_{n-1}^1 \notin V(p)$ and $f^1(k) = |1 - f(k)|$. Iterating the process we can obtain a sequence of the form $wR_f^{n-1}(1)w_1^{n-1} Rf^{n-1}(2) \ldots Rf^{n-1}(n-1)w_{n-1}^{n-1}$ with $w_{n-1}^{n-1} \notin V(p)$ and with f^{n-1} being one of the constant functions of the co-domain $\{0, 1\}$. By transitivity it follows that $w_{n-1}^{n-1} \in K_0[w] \cap K_1[w] \cap \overline{V(p)}$, contradicting the assumption that at least of the agents knew p at w. $\square$

Proposition 3.15: Induced subgraph bisimulation is NP-complete.

Proof. Showing that the problem is in NP is straightforward. Hardness is shown by reduction from Independent Set. First of all, let $I_k = (V_{I_k}, E_{I_k} = \emptyset)$ with $|V_{I_k}| = k$ denote a graph with k vertices and no edges. Given the input of Independent Set, i.e. a graph $G = (V, E)$ and some $k \in \mathbb{N}$ we transform it into (I_k, G), k, as input for Induced Subgraph Bisimulation.

Now, we claim that G has an independent set of size at least k iff there is some $V' \subseteq V$ with $|V'| \geq k$ and $(V', E \cap (V' \times V')) \underline{\leftrightarrow}_{total} I_k$.

From left to right, assume that there is some $S \subseteq V$ with $|S| = k$, and for all $v, v' \in S$, $(v, v') \notin E$. Now, any bijection between S and V_{I_k} is a total bisimulation between $G' = (S, E \cap (S \times S))$ and I_k, since $E \cap (S \times S) = \emptyset$ and $|S| = |V_{I_k}|$.

For the other direction, assume that there is some $V' \subseteq V$ with $|V'| = k$ such that for $G' = (V', E' = E \cap (V' \times V'))$ we have that $G' \underline{\leftrightarrow}_{total} I_k$. Thus, there is some total bisimulation Z between G' and I_k. Now, we claim that V' is an independent set of G of size k. Let $v, v' \in V'$. Suppose that $(v, v') \in E$. Then since G' is an induced subgraph, we also have that $(v, v') \in E'$. Since Z is a total bisimulation, there has to be some $w \in I_k$ with $(v, w) \in Z$ and some w' with $(w, w') \in E_{I_k}$ and $(v', w') \in Z$. But this is a contradiction with $E_{I_k} = \emptyset$. Thus, V' is an independent set of size k of G. The reduction can clearly be computed in polynomial time. This concludes the proof. $\square$

Proposition 3.18: If for graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ it holds that E_1 and E_2 are reflexive, transitive and symmetric, then the problem of induced subgraph bisimulation for G_1 and G_2 can be solved in linear time.

Proof. In this proof, we will use the fact that $G_1 = (V_1, E_1) \underline{\leftrightarrow}_{total} G_2 = (V_2, E_2)$ if and only if it is the case that $V_1 = \emptyset$ iff $V_2 = \emptyset$. Let us prove this. From left to right, assume that $G_1 = (V_1, E_1) \underline{\leftrightarrow}_{total} G_2 = (V_2, E_2)$. Then since we have a total bisimulation, it must be the case that either $V_1 = V_2 = \emptyset$ or $V_1 \neq \emptyset \neq V_2$.

For the other direction, assume that $V_1 = \emptyset$ iff $V_2 = \emptyset$. Now, we show that in this case, $V_1 \times V_2$ is a total bisimulation between G_1 and G_2. If $V_1 = V_2 = \emptyset$, we are done. So, consider the case where $V_1 \neq \emptyset \neq V_2$. Let $(v_1, v_2) \in V_1 \times V_2$, and assume that $(v_1, v_1') \in E_1$ for some $v_1' \in V_1$. Since E_2 is reflexive, we know that there is some $v_2' \in V_2$ such that $(v_2, v_2') \in E_2$. Of course $(v_1', v_2') \in V_1 \times V_2$. The back condition is analogous. Since $V_1 \times V_2$ is total, we thus have $G_1 \underline{\leftrightarrow}_{total} G_2$. Hence, $G_1 = (V_1, E_1) \underline{\leftrightarrow}_{total} G_2 = (V_2, E_2)$ if and only if it is the case that $V_1 = \emptyset$ iff $V_2 = \emptyset$.

Therefore, for solving the induced subgraph bisimulation problem for input G_1 and G_2 with E_1 and E_2 being reflexive, transitive and symmetric and $k \in \mathbb{N}$, all we need to do is to go through the input once and check whether $V_1 = \emptyset$ iff $V_2 = \emptyset$, and whether $|V_2| \geq k$. If the answer to both is *yes* then we know that $G_1 \underline{\leftrightarrow}_{total} G_2$ and since $|V_2| \geq k$, we answer *yes*, otherwise *no*. $\square$

Proposition 3.20: Local submodel bisimulation for single agent pointed epistemic models is in P.

Proof. Given the input of the problem, i.e. a pointed epistemic model $\mathcal{M}_1, w$ with $\mathcal{M}_1 = (W_1, \sim_1, V_1)$, and $w \in W_1$ and an epistemic model $\mathcal{M}_2 = (W_2, \sim_2, V_2)$, we run the following procedure.

1. For all $[w_2] \in W_2/\sim_2$ do the following:
   (a) Initialize the set $Z := \emptyset$.
   (b) for all $w' \in [w]$ do the following
      i. For all $w_2' \in [w_2]$ check if for all $p \in \mathbf{Prop}$ it holds that $w' \in V_1(p)$ iff $w_2' \in V_2(p)$. If this is the case, set $Z := Z \cup (w', w_2')$.
      ii. if there is no such w_2', continue with 1, otherwise we return Z and we stop.

2. In case we didn't stop at 1(b)ii, we can stop now, and return *no*.

In the worst case, this takes $|\mathcal{M}_1| \cdot |\mathcal{M}_2|$ steps.

If the procedure has stopped at 2, there is no bisimulation with the required properties. To see this, note that if we stopped in 2, this means that there was no $[w_2] \in W_2/\sim_2$ such that for every state in $[w]$ there is one in $[w_2]$ in which exactly the same propositional letters are true. Thus, since we were looking for a bisimulation that is also defined for the state w, such a bisimulation cannot exist.

If the algorithm returned a relation Z, this is indeed a bisimulation between $\mathcal{M}_1$ and the submodel $\mathcal{M}_2'$ of $\mathcal{M}_2$ where $\mathcal{M}_2' = (W_2', \sim_2', V_2')$, where

$$W_2' = \{w_2 \in W_2 \mid \text{there is some } w_1 \in [w] \text{ such that } (w_1, w_2) \in Z\}$$

and $\sim_2'$ and V_2' are the usual restrictions of $\sim_2$ and V_2 to W_2. This follows from the following two facts: First, for all pairs in Z it holds that both states satisfy exactly the same proposition letters. Second, since Z is total both on $[w]$ and on W_2' and all the states in $[w]$ are connected to each other by $\sim_1$ and all states in W_2' are connected to each other by $\sim_2'$, both the *forth* and *back* conditions are satisfied. This concludes the proof. $\qquad\square$

Theorem 3.22: Total submodel bisimulation for single agent epistemic models is in P.

First we introduce some notation used in the proof.

Notation A.1. *Let $\mathcal{M} = (W, \sim, V)$ be a single agent epistemic model. For the valuation function $V : \mathbf{Prop} \to W$, we define $\hat{V} : W \to 2^{\mathbf{Prop}}$, with $w \mapsto \{p \in \mathbf{Prop} \mid w \in V(p)\}$. Abusing notation, for $X \subseteq W$ we sometimes write $\hat{V}(X)$ to denote $\{\hat{V}(w) \mid w \in X\}$. For $w \in W$, $[w] = \{w' \in W \mid w \sim w'\}$ denotes the equivalence class of w under $\sim$. $W/\sim$ denotes the set of all equivalence classes of W for the relation $\sim$.*

Definition A.2. Given a single agent epistemic model $\mathcal{M} = (W, \sim, V)$, $\mathcal{M}^{min\_cells}$ denote a model obtained from $\mathcal{M}$ by the following procedure:

1. Initialize X with $X := W/\sim$.

2. Go through all the pairs in $X \times X$.

   (a) When you find $([w], [w'])$ with $[w] \neq [w']$ such that $\hat{V}([w]) = \hat{V}([w'])$, continue at 2 with $X := X - [w']$.

   (b) Otherwise, stop and return the model $\mathcal{M}^{min\_cells} := (\bigcup X, \sim', V')$, where $\sim'$ and V' are the usual restrictions of $\sim$ and V to $\bigcup X$.

Fact A.3. *With input $\mathcal{M} = (W, \sim, V)$, the procedure in Definition A.2 runs in time polynomial in $|\mathcal{M} = (W, \sim, V)|$.*

Proof. Follows from the fact that the cardinality of $W/\sim$ is bounded by $|W|$; we only enter step 2 at most $|W|$ times, and each time do at most $|W|^2$ comparisons. $\qquad\square$

Fact A.4. *The answer to total submodel bisimulation for single agent epistemic models (Decision Problem 3.21) with input $\mathcal{M}_1 = (W_1, \sim_1, V_1), \mathcal{M}_2 = (W_2, \sim_2, V_2)$ is yes iff it is with input $\mathcal{M}_1^{min\_cells} = (W_1, \sim_1, V_1), \mathcal{M}_2 = (W_2, \sim_2, V_2)$.*

Proof. From left to right, we just need to restrict the bisimulation to the states of $\mathcal{M}_1^{min\_cells}$. For the other direction, we start with the given bisimulation and then extend it as follows. For the states in a cell $[w']$ which was removed during the construction of $\mathcal{M}_1^{min\_cells}$, can be mapped to the ones of a cell $[w]$ in $\mathcal{M}_1^{min\_cells}$ with the same valuation. $\qquad\square$

Proof. By Fact A.3 and Fact A.4, transforming $\mathcal{M}_1$ into $\mathcal{M}_1^{min\_cells}$ can be done in polynomial time. Thus, w.l.o.g. we can assume that $\mathcal{M}_1$ is already of the right shape; i.e. $\mathcal{M}_1 = \mathcal{M}_1^{min\_cells}$. Given the two models as input, we construct a bipartite graph $G = ((W_1/\sim_1, W_2/\sim_2), E)$ where E is defined as follows.

$$([w_1], [w_2]) \in E \text{ iff } \hat{V}_1([w_1]) \subseteq \hat{V}_2([w_2]).$$

Claim A.5. *There is a submodel $\mathcal{M}'_2$ of $\mathcal{M}_2$ such that $\mathcal{M}_1 \underline{\leftrightarrow}_{total} \mathcal{M}'_2$ iff G has a matching of size $|W_1/\sim_1|$.*

Proof. From left to right, assume that there is a submodel $\mathcal{M}'_2 = (W'_2, \sim'_2, V'_2)$ of $\mathcal{M}_2$ such that $\mathcal{M}_1 \underline{\leftrightarrow}_{total} \mathcal{M}'_2$. Let Z be such a total bisimulation.

Note that since we assumed that $\mathcal{M}_1 = \mathcal{M}^{min\_cells}$ the following holds:

1. For all $([w_1], [w_2]) \in W_1/\sim_1 \times W_2/\sim_2$ it is the case that whenever $Z \cap ([w_1] \times [w_2]) \neq \emptyset$, then for all $[w'_1] \in W_1/\sim_1$ such that $[w'_1] \neq [w_1]$, $Z \cap ([w'_1] \times [w_2]) = \emptyset$.

Thus, the members of different equivalence classes in $W_1/\sim_1$ are mapped by Z to into different equivalence classes of $W_2/\sim_2$.

Now, we construct $\dot{E} \subseteq E$ as follows.

$$([w_1], [w_2]) \in \dot{E} \text{ iff } ([w_1], [w_2]) \in E \text{ and } ([w_1] \times [w_2]) \cap Z \neq \emptyset.$$

Then $|\dot{E}| \geq |W_1/\sim_1|$ because of the definitions E and $\dot{E}$ and the fact that Z is a bisimulation that is total on W_1. Now, if $|\dot{E}| = |W_1/\sim_1|$ then we are done since by definition of $\dot{E}$, for each $[w_1] \in W_1/\sim_1$ there is some $[w_2] \in W_2/\sim_2$ such that $([w_1], [w_2]) \in \dot{E}$. Then it follows from 1, that $\dot{E}$ is indeed a matching.

If $|\dot{E}| > |W_1/\sim_1|$ then we can transform $\dot{E}$ into a matching E' of size $W_1/\sim_1|$: For each $[w_1] \in W_1/\sim_1$, we pick *one* $[w_2] \in W_2/\sim_2$ such that $([w_1], [w_2]) \in \dot{E}$ and put it into E' (note that such a $[w_2]$ always exists because by definition of $\dot{E}$, for each $[w_1] \in W_1/\sim_1$ there is some $[w_2] \in W_2/\sim_2$ such that $([w_1], [w_2]) \in \dot{E}$; moreover because of 1 all the $[w_2] \in W_2/\sim_2$ that we pick will be different). Then the resulting $E' \subseteq \dot{E} \subseteq E \subseteq (W_1/\sim_1 \times W_2/\sim_2)$ is a matching of G of size $|W_1/\sim_1|$. Thus, we have shown that if there is a submodel $\mathcal{M}'_2$ of $\mathcal{M}_2$ such that $\mathcal{M}_1 \underline{\leftrightarrow}_{total} \mathcal{M}'_2$ then G has a matching of size $|W_1/\sim_1|$.

For the other direction, assume that G has a matching $E' \subseteq E$ with $|E'| = |W_1/\sim_1|$. Then, recalling the definition of E, it follows that for all $[w] \in W_1/\sim$ there is some $[w'] \in W_2/\sim_2$ such that $([w], [w']) \in E'$ and thus $\hat{V}_1([w]) \subseteq \hat{V}_2([w'])$.

Let us define the following submodel $\mathcal{M}'_2$ of $\mathcal{M}_2$. $\mathcal{M}'_2 = (W'_2, \sim'_2, V'_2)$, where

$$W'_2 = \{w_2 \in W_2 \mid \text{there is some } w \in W_1 \text{ such that } \hat{V}_1(w) = \hat{V}_2(w_2) \text{ and } ([w], [w_2]) \in E'\}$$

and $\sim'_2$ and V'_2 are the usual restrictions of $\sim_2$ and V_2 to W'_2.

Now, we define a relation $Z \subseteq W_1 \times W'_2$, which we then show to be a total bisimulation between $\mathcal{M}_1$ and $\mathcal{M}'_2$

$$(w_1, w_2) \in Z \text{ iff } \hat{V}(w_1) = \hat{V}_2(w_2) \text{ and } ([w_1], [w_2]) \in E'.$$

Next, let us show that Z is indeed a bisimulation.

Let $(w_1, w_2) \in Z$. Then, by definition of Z, for every propositional letter p, $w_1 \in V_1(p)$ iff $w_2 \in V_2(p)$. Next, we check the *forth* condition. Let $w_1 \sim_1 w'_1$ for some $w'_1 \in W_1$. Then since $(w_1, w_2) \in Z$, and thus $([w_1], [w_2]) \in E'$, there has to be some $w'_2 \in [w_2]$ such that $\hat{V}_2(w'_2) = \hat{V}_1(w'_1)$. Then since $[w'_1] = [w_1]$ and $[w'_2] = [w_2]$, $([w'_1], [w'_2]) \in E'$. Then $w'_2 \in W'_2$, and $(w'_1, w'_2) \in Z$.

For the *back* condition, let $w_2 \sim_2 w'_2$, for some $w'_2 \in W'_2$. Then by definition of W'_2, there is some $w \in W_1$ such that $\hat{V}_1(w) = \hat{V}_2(w'_2)$ and $([w], [w'_2]) \in E'$. Thus, it follows that $(w, w'_2) \in Z$. Now, we still have to show that $w_1 \sim_1 w$. As the following hold: $([w], [w'_2]) \in E', [w_2] = [w'_2]$, $([w], [w_2]) \in E'$ (because $(w_1, w_2) \in Z$) and E' is a matching, it follows that $[w] = [w_1]$. Thus, $w_1 \sim_1 w$.

Hence, we conclude that Z is a bisimulation. It remains to show that Z is indeed total.

Let $w_1 \in W_1$. Since E' is a matching of size $W_1/\sim_1$, there is some $[w_2] \in W_2/\sim_2$ such that $([w_1],[w_2]) \in E'$. Thus, there is some $w_2' \in [w_2]$ such that $\hat{V}_1(w_1) = \hat{V}_2(w_2')$. This means that $w_2' \in W_2'$ and $(w_1, w_2') \in Z$. So Z is total on W_1.

Let $w_2 \in W_2'$. By definition of W_2', there is some $w \in W_1$ such that $\hat{V}_1(w) = \hat{V}_2(w_2)$ and $([w],[w_2]) \in E'$. Thus, by definition of $Z, (w, w_2) \in Z$. Therefore, Z is indeed a total bisimulation between $\mathcal{M}_1$ and $\mathcal{M}_2'$. This concludes the proof of Claim A.5. $\qquad \square$

Hence, given two models, we can transform the first one using the polynomial procedure of Definition A.2 and then we construct the graph G, which can be done in polynomial time as well. Finally, we use a polynomial algorithm to check if G has a matching of size $M_1^{min\_cells}$. If the answer is yes, we return *yes*, otherwise *no*. This concludes the proof of Theorem 3.22. $\qquad \square$

Proposition 3.25: Induced subgraph simulation is NP-complete.

Proof. Showing that the problem is in NP is straightforward. Hardness is shown by reduction from Independent Set. First of all, let $I_k = (V_{I_k}, E_{I_k} = \emptyset)$ with $|V_{I_k}| = k$ denote a graph with k vertices and no edges. Given the input of Independent Set, i.e. a graph $G = (V, E)$ and some $k \in \mathbb{N}$ we transform it into $(I_k, G), k$, as input for Induced Subgraph Simulation.

Now, we claim that G has an independent set of size at least k iff there is some $V' \subseteq V$ with $|V'| \geq k$ and $(V', E \cap (V' \times V')) \sqsubseteq_{total} I_k$.

From left to right, assume that there is some $S \subseteq V$ with $|S| = k$, and for all $v, v' \in S$, $(v, v') \notin E$. Now, any bijection between S and V_{I_k} is a total simulation (and in fact an isomorphism) between $G' = (S, E \cap (S \times S))$ and I_k, since $E \cap (S \times S) = \emptyset$ and $|S| = |V_{I_k}|$.

For the other direction, assume that there is some $V' \subseteq V$ with $|V'| = k$ such that for $G' = (V', E' = E \cap (V' \times V'))$ we have that $G' \sqsubseteq_{total} I_k$. Thus, there is some total simulation Z between G' and I_k. Now, we claim that V' is an independent set of G of size k. Let $v, v' \in V'$. Suppose that $(v, v') \in E$. Then since G' is an induced subgraph, we also have that $(v, v') \in E'$. Since Z is a total simulation, there has to be some $w \in I_k$ with $(v, w) \in Z$ and some w' with $(w, w') \in E_{I_k}$ and $(v', w') \in Z$. But this is a contradiction with $E_{I_k} = \emptyset$. Thus, V' is an independent set of size k of G. The reduction can clearly be computed in polynomial time. This concludes the proof. $\qquad \square$

Proposition 3.27: Induced subgraph simulation equivalence is NP-complete.

Proof. For showing that the problem is in NP, note that we can use a simulation equivalence algorithm as provided in [14]. Hardness can again be shown by reduction from Independent Set. Given the input for Independent Set, i.e. a graph $G = (V, E)$ and some $k \in \mathbb{N}$, we transform it into two graphs $I_k = (V_{I_k} = \{v_1, \ldots v_k\}, E_{I_k} = \emptyset)$ and G, and we keep the $k \in \mathbb{N}$. This can be done in polynomial time.

Now, we claim that G has an independent set of size k iff there is an induced subgraph of G with k vertices that is similar to I_k. From left to right assume that G has such an independent set S with $S \subseteq V, |S| = k$ and $E \cap S \times S = \emptyset$. Then $(S, \emptyset)$ is isomorphic to I_k since both have k vertices and no edges. Thus, they are also simulation equivalent.

For the other direction, assume that there is an induced subgraph $G' = (V', E')$ with $V' \subseteq V, |V'| = k$ and $E' = (V' \times V') \cap E$ such that G' is simulation equivalent to I_k. Suppose that there are $v, v' \in V'$ such that $(v, v') \in E$. Since G' is an induced subgraph, it must be the case that $(v, v') \in E'$, but since I_k simulates G', this leads to a contradiction since I_k does not have any edges.

This concludes the proof. $\qquad \square$