Towards Symbolic Factual Change in DEL

Malvin Gattinger

Institute for Logic, Language & Computation, University of Amsterdam

Abstract. We extend symbolic model checking for Dynamic Epistemic Logic (DEL) with factual change. Our transformers provide a compact representation of action models with pre- and postconditions, for both S5 and the general case. The method can be implemented using binary decision diagrams and we expect it to improve model checking performance. As an example we give a symbolic representation of the Sally-Anne false belief task.

Keywords: Epistemic Logic, Symbolic Model Checking, Factual Change

1 Introduction

Symbolic representation is a solution to the state explosion problem in model checking. The idea is to not store models explicitly in memory, but to find more compact representations which still allow the evaluation of formulas. In [3] it was shown that S5 Kripke models for Dynamic Epistemic Logic (DEL) can be encoded symbolically using *knowledge structures*. They are of the form (V, θ, O) where V is a set of atomic propositions called *vocabulary*, θ is a boolean formula called *state law* and $O_i \subseteq V$ are observational variables for each agent. Notably, this symbolic representation preserves the truth of all DEL formulas, including higher-order knowledge (see Section 3 below).

The framework was generalized in [4] in two ways: From equivalences to arbitrary relations and from announcements to action models. The latter can be represented by *knowledge transformers* of the form (V^+, θ^+, O^+). Analogous to the product update on Kripke models [1], applying a transformer to a structure yields a new structure. However, knowledge transformers only change what agents know and not what is the case — they do not provide a symbolic equivalent of postconditions for factual change as studied in [5].

In this paper we combine the two generalizations and add the missing components to treat factual change. The result are *belief transformers with factual change* which for simplicity we will just call *transformers*.

Possible worlds in a Kripke model get their meaning but not their identity via a valuation function. In particular we can assign the same atomic truths to different possible worlds. In contrast, all states of a knowledge structure satisfy different atomic propositions and can thus be identified with their valuation. This is what makes structures symbolic and efficient to implement, but it complicates the idea of changing facts, as the following minimal example shows.

Example 1. Consider a coin lying on a table with heads up: p is true and this is common knowledge. Suppose we then toss it randomly and hide the result from agent a but reveal it to agent b. Figure 1 shows a Kripke model of this update.

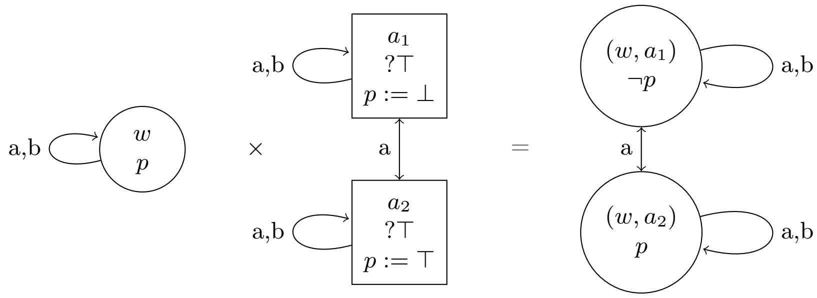

Fig. 1. Factual change on Kripke models.

It is easy to find the following structures that are equivalent to the initial and the resulting model, but how can we symbolically describe the update which transforms one into the other?

$$\begin{array}{ll} & (V = \{p\},\ \theta = p,\ O_a = \{p\},\ O_b = \{p\}) \\ \times & \qquad\qquad\qquad ??? \\ = & (V = \{p\},\ \theta = \top,\ O_a = \varnothing\ ,\ O_b = \{p\}) \end{array}$$

The name of a resulting world (w, a_1) makes clear that it "comes from" w. But a state like $\varnothing$ does not reveal its history or any relation to $\{p\}$. For purely epistemic actions this is not a problem — we only *add* propositions from V^+ to the state to distinguish different epistemic events. But for factual change propositions from V have to be modified and we need a way to *remove* them from states.

Our solution is to *copy propositions*: We store the old value of p in a fresh variable p°. Then we rewrite the state law and observations using substitutions.

We proceed as follows. Sections 2 and 3 summarize the relevant parts of [4], generalized to belief transformers. We then add factual change in Section 4 and show that transformers are equivalent to action models in Section 5. The Sally-Anne task illustrates our framework in Section 6 and we finish with further questions in Section 7.

Definition 1 (Languages and Notation). *We fix a finite set of agents I denoted by i, j, etc. and use the letters V or X for sets of atomic propositional variables denoted by p, q, etc.*

For any set of propositions X we write $\mathcal{L}_B(X)$ for the boolean language given by the BNF $\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi$ where $p \in X$. Similarly, let $\mathcal{L}(X)$ be the epistemic language over X given by $\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \Box_i\varphi$ where $i \in I$.

Primes and circles denote fresh variables, for example p' and p°. For sets of variables let $X' := \{x' \mid x \in X\}$ and $X^\circ := \{x^\circ \mid x \in X\}$. We also extend this notation to formulas recursively, for example $(\Box_i p \wedge q)' = (\Box_i p' \wedge q')$.

We write $[p/\psi]\varphi$ for the result of substituting ψ for p in φ. Given two sets of the same size A and B of atomic propositions, and implicitly assuming an enumeration $A = \{a_1, \ldots, a_k\}$ and $B = \{b_1, \ldots, b_k\}$ we write $[B/A]\,\varphi$ for the result of substituting a_i for b_i in φ in parallel for all i.

A boolean assignment is identified with its set of true propositional variables and we write $\vDash$ for the standard satisfaction relation. Boolean quantification is used as follows: $\forall p \varphi := [p/\top]\,\varphi \wedge [p/\bot]\,\varphi$. For any $A = \{p_1, \ldots, p_n\}$, let $\forall A \varphi := \forall p_1 \forall p_2 \ldots \forall p_n \varphi$. To abbreviate that a specific subset of propositions is true, let $A \sqsubseteq B := \bigwedge A \wedge \bigwedge \{\neg p \mid p \in B \setminus A\}$.

Definition 1 describes operations on boolean *formulas* which might not be efficient in practice. In any actual implementation of our methods those should be replaced with operations on boolean *functions* represented as Binary Decision Diagrams (BDDs) [7]. For example, boolean quantification should *not* be implemented as an abbreviation but can instead be done efficiently by eliminating quantified variables from the BDD.

2 Kripke Models and Action Models

We quickly state the standard definitions for Kripke semantics of DEL. For a general introduction see [9] and for details on factual change see [5].

Definition 2. *A* Kripke model *for V is a tuple $\mathcal{M} = (W, \mathcal{R}, \pi)$ where W is the set of worlds, $\mathcal{R}_i \subseteq W \times W$ is a relation for each i and $\pi : W \to \mathcal{P}(V)$ is a valuation function. A* pointed *Kripke model is a pair $(\mathcal{M}, w)$ where $w \in W$.*

We interpret $\mathcal{L}(V)$ on pointed Kripke models as follows.

1. $(\mathcal{M}, w) \vDash p$ *iff* $p \in \pi(w)$
2. $(\mathcal{M}, w) \vDash \neg\varphi$ *iff not* $(\mathcal{M}, w) \vDash \varphi$
3. $(\mathcal{M}, w) \vDash \varphi \wedge \psi$ *iff* $(\mathcal{M}, w) \vDash \varphi$ *and* $(\mathcal{M}, w) \vDash \psi$
4. $(\mathcal{M}, w) \vDash \Box_i \varphi$ *iff for all* $v \in W$ *: If* $w \mathcal{R}_i v$ *then* $(\mathcal{M}, v) \vDash \varphi$

The following definition describes action models and how they can be applied to Kripke models. Our definition of postconditions differs from the standard in [5] because we only allow boolean formulas. This however does not change the expressivity [10].

Definition 3. *An* action model *is a tuple $\mathcal{A} = (A, R, \mathsf{pre}, \mathsf{post})$ where A is a set of atomic events, $R_i \subseteq A \times A$ a relation for each i, $\mathsf{pre} : A \to \mathcal{L}(V)$ is a precondition function and $\mathsf{post} : A \times V \to \mathcal{L}_B(V)$ a postcondition function.*

The product update *is defined by $\mathcal{M} \times \mathcal{A} := (W^{new}, \mathcal{R}_i^{new}, \pi^{new})$ where*

- $W^{new} := \{(w, a) \in W \times A \mid \mathcal{M}, w \vDash \mathsf{pre}(a)\}$
- $\mathcal{R}_i^{new} := \{((w, a), (v, b)) \mid \mathcal{R}_i wv \text{ and } R_i ab\}$
- $\pi^{new}((w, a)) := \{p \in V \mid \mathcal{M}, w \vDash \mathsf{post}_a(p)\}$

An action *is a pair $(\mathcal{A}, a)$ where $a \in A$.*

3 Belief Structures and Belief Transformers

We now present the definitions of belief structures and belief transformers. The key idea is that instead of explicitly listing worlds we use a symbolic representation: The set of worlds and the valuation function are replaced by a vocabulary and a boolean formula called the state law. The set of states is then implicitly given as all boolean assignments that satisfy the state law, i.e. a subset of the powerset of the vocabulary. Moreover, also the epistemic relations can be encoded using boolean formulas and the goal is to interpret the language on the resulting structures without ever computing or listing the full set of states. For more details and proofs, we refer the reader to [4].

Definition 4. *A* belief structure *is a tuple* $\mathcal{F} = (V, \theta, \Omega)$ *where* V *is a finite set of atomic propositions called* vocabulary, $\theta \in \mathcal{L}_B(V)$ *is the* state law *and* $\Omega_i \in \mathcal{L}_B(V \cup V')$ *are called* observations. *Any* $s \subseteq V$ *such that* $s \vDash \theta$ *is called a* state *of* $\mathcal{F}$. *A pair* $(\mathcal{F}, s)$ *where* s *is a state of* $\mathcal{F}$ *is called a* scene.

We interpret $\mathcal{L}(V)$ *on scenes as follows.*

1. $(\mathcal{F}, s) \vDash p$ *iff* $s \vDash p$.
2. $(\mathcal{F}, s) \vDash \neg\varphi$ *iff not* $(\mathcal{F}, s) \vDash \varphi$
3. $(\mathcal{F}, s) \vDash \varphi \wedge \psi$ *iff* $(\mathcal{F}, s) \vDash \varphi$ *and* $(\mathcal{F}, s) \vDash \psi$
4. $(\mathcal{F}, s) \vDash \Box_i\varphi$ *iff for all* $t \subseteq V$: *If* $t \vDash \theta$ *and* $(s \cup t') \vDash \Omega_i$ *then* $(\mathcal{F}, t) \vDash \varphi$.

We write $(\mathcal{F}, s) \equiv_V (\mathcal{F}', s')$ *iff these two scenes agree on all formulas of* $\mathcal{L}(V)$.

An interesting property of belief structures is that on a given structure all epistemic formulas have boolean equivalents. The following translation reduces model checking to boolean operations which is not possible on Kripke models.

Definition 5. *For any belief structure* $\mathcal{F} = (V, \theta, \Omega)$ *and any formula* $\varphi \in \mathcal{L}(V)$ *we define its* local boolean translation $\|\varphi\|_\mathcal{F}$ *as follows.*

1. *For any primitive formula, let* $\|p\|_\mathcal{F} := p$.
2. *For negation, let* $\|\neg\psi\|_\mathcal{F} := \neg\|\psi\|_\mathcal{F}$.
3. *For conjunction, let* $\|\psi_1 \wedge \psi_2\|_\mathcal{F} := \|\psi_1\|_\mathcal{F} \wedge \|\psi_2\|_\mathcal{F}$.
4. *For belief, let* $\|\Box_i\psi\|_\mathcal{F} := \forall V'(\theta' \rightarrow (\Omega_i \rightarrow (\|\varphi\|_\mathcal{F})'))$

Theorem 1 (from [4]). *Definition 5 preserves and reflects truth. That is, for any formula* φ *and any scene* $(\mathcal{F}, s)$ *we have that* $(\mathcal{F}, s) \vDash \varphi$ *iff* $s \vDash \|\varphi\|_\mathcal{F}$.

The following definition was only hinted at in [4]. Belief transformers are like knowledge transformers, but instead of observed propositions O_i^+ we use boolean formulas Ω_i^+ to encode arbitrary relations on $\mathcal{P}(V^+)$.

Definition 6. *A* belief transformer *for* V *is a tuple* $\mathcal{X} = (V^+, \theta^+, \Omega^+)$ *where* V^+ *is a set of atomic propositions such that* $V \cap V^+ = \varnothing$, $\theta^+ \in \mathcal{L}(V \cup V^+)$ *is a possibly epistemic formula and* $\Omega_i^+ \in \mathcal{L}_B(V \cup V^+)$ *is a boolean formula for each* $i \in I$. *A* belief event *is a belief transformer together with a subset* $x \subseteq V^+$, *written as* $(\mathcal{X}, x)$.

The belief transformation *of a belief structure* $\mathcal{F} = (V, \theta, \Omega)$ *with* $\mathcal{X}$ *is defined by* $\mathcal{F} \times \mathcal{X} := (V \cup V^+, \theta \wedge ||\theta^+||_{\mathcal{F}}, \{\Omega_i \wedge \Omega_i^+\}_{i \in I})$*. Given a scene* $(\mathcal{F}, s)$ *and a belief event* $(\mathcal{X}, x)$*, let* $(\mathcal{F}, s) \times (\mathcal{X}, x) := (\mathcal{F} \times \mathcal{X}, s \cup x)$*.*

The resulting observations are boolean formulas over the new double vocabulary $(V \cup V') \cup (V^+ \cup V^{+\prime}) = (V \cup V^+) \cup (V \cup V^+)'$, describing a relation between the new states which are subsets of $V \cup V^+$.

4 Belief Transformers with Factual Change

We now define transformation with factual change, adding the components V_- and θ_-. Note that the belief transformers without factual change as discussed in the previous section are exactly those transformers where $V_- = \varnothing$.

Definition 7. *A* transformer *for* V *is a tuple* $\mathcal{X} = (V^+, \theta^+, V_-, \theta_-, \Omega^+)$ *where*

- V^+ *is a set of fresh atomic propositions such that* $V \cap V^+ = \varnothing$*,*
- θ^+ *is a possibly epistemic formula from* $\mathcal{L}(V \cup V^+)$*,*
- $V_- \subseteq V$ *is the* modified subset *of the original vocabulary,*
- $\theta_- : V_- \to \mathcal{L}_B(V \cup V^+)$ *maps modified propositions to boolean formulas,*
- $\Omega_i^+ \in \mathcal{L}_B(V^+ \cup V^{+\prime})$ *are boolean formulas for each* $i \in I$*.*

To transform $\mathcal{F} = (V, \theta, \Omega_i)$ *with* $\mathcal{X}$*, let* $\mathcal{F} \times \mathcal{X} := (V^{new}, \theta^{new}, \Omega_i^{new})$ *where*

1. $V^{new} := V \cup V^+ \cup V_-^\circ$
2. $\theta^{new} := [V_-/V_-^\circ] (\theta \wedge \|\theta^+\|_{\mathcal{F}}) \wedge \bigwedge_{q \in V^-} (q \leftrightarrow [V_-/V_-^\circ] (\theta_-(q)))$
3. $\Omega_i^{new} := ([V_-/V_-^\circ] [(V_-)'/(V_-^\circ)'] \Omega_i) \wedge \Omega_i^+$

An event *is a pair* $(\mathcal{X}, x)$ *where* $x \subseteq V^+$*. Given* $(\mathcal{F}, s)$ *and* $(\mathcal{X}, x)$*, let* $(\mathcal{F}, s) \times (\mathcal{X}, x) := (\mathcal{F} \times \mathcal{X}, s^{new})$ *where the new actual state is* $s^{new} := (s \setminus V_-) \cup (s \cap V_-)^\circ \cup x \cup \{p \in V_- \mid s \cup x \vDash \theta_-(p)\}$*.*

To explain this definition, let us consider the components one by one.

First, the new vocabulary contains $V_-^\circ = \{p^\circ \mid p \in V_-\}$. These are fresh copies of the modified subset. We use them to keep track of the old valuation.

Second, the new state law: A state in the resulting structure needs to satisfy the old state law and the event law encoding the preconditions. For modified propositions the old values have to be used, hence we apply a substitution to both laws in the left conjunct. Modified propositions are then overwritten in the right conjunct, using θ_- which encodes postconditions. As in Definition 3, postconditions are evaluated in the old model, hence we also substitute here.

Third, for the new observations we replace modified variables by their copies. Two substitutions are needed because Ω_i^{new} is in a double vocabulary. Old observations induce new ones via the state law. For example, if q was flipped publicly, then $q \leftrightarrow \neg q^\circ$ is part of the new state law and observing whether q is equivalent to observing whether $\neg q^\circ$, i.e. having observed q in the original structure. In the simpler S5 setting we would use $O_i^{\text{new}} := ([V_-/V_-^\circ] O_i) \cup O_i^+$.

Finally, the new actual state s^{new} is the union of, in this order: propositions in the old state that have not been modified ($s \setminus V_-$), copies of the modified propositions that were in the old state $(s \cap V_-)^\circ$, those propositions labeling the actual event x and the modified propositions whose precondition was true in the old state $\{p \in V_- \mid s \cup x \vDash \theta_-(p)\}$.

Example 2. We can now model the coin flip from Example 1 as follows. Because we use the more general belief (instead of knowledge) structures, the initial structure now has boolean formulas Ω_i instead of observational variables O_i:

$$(V = \{p\},\ \theta = p,\ \Omega_a = p \leftrightarrow p',\ \Omega_b = p \leftrightarrow p')$$

The following transformer models the coin flip visible to b but not to a:

$$(V^+ = \{q\},\ \theta^+ = \top,\ V_- = \{p\},\ \theta_-(p) := q,\ \Omega_a^+ = \top,\ \Omega_b^+ = q \leftrightarrow q')$$

The result of applying the latter to the former is this:

$$(V = \{p, q, p^\circ\},\ \theta = p^\circ \wedge (p \leftrightarrow q),\ \Omega_a = p^\circ \leftrightarrow p^{\circ\prime},\ \Omega_b = (p^\circ \leftrightarrow p^{\circ\prime}) \wedge (q \leftrightarrow q'))$$

Example 3. A publicly observable change $p := \varphi$ for a propositional formula φ is modeled by:

$$(V^+ = \varnothing,\ \theta^+ = \top,\ V_- = \{p\},\ \theta_-(p) := \varphi,\ \Omega_i^+ = \top)$$

DEL does not have temporal operators and agents never know the past explicitly. Hence the old valuation is often irrelevant and the product update on Kripke models does this "garbage collection" better than our transformation. But we can eliminate propositions outside the original V using the following Lemma. A more thorough analysis of minimizing knowledge structures will be future work.

Lemma 1. *Suppose $\mathcal{F}$ uses the vocabulary $V \cup \{p\}$ and $p \notin V$ is determined by the state law (i.e. $\theta \rightarrow p$ or $\theta \rightarrow \neg p$ is a tautology). Then we can remove p from the state law and observational BDDs to get a new structure $\mathcal{F}'$ using the vocabulary V such that $(\mathcal{F}, s) \equiv_V (\mathcal{F}', s \setminus \{p\})$.*

Example 4. The result from Example 2 is $\equiv_{\{p,q\}}$ equivalent to:

$$(V = \{p, q\},\ \theta = p \leftrightarrow q,\ \Omega_a = \top,\ \Omega_b = q \leftrightarrow q')$$

5 Equivalence and Expressiveness

We now show that transformers describe exactly the same class of updates as action models. The main ingredients for the proof are the following Lemma and two Definitions of how to go from transformers to action models and back.

Lemma 2 (from [4]). *Suppose we have a belief structure $\mathcal{F} = (V, \theta, \Omega)$, a finite Kripke model $\mathcal{M} = (W, \pi, \mathcal{R})$ for the vocabulary $X \subseteq V$ and a function $g : W \rightarrow \mathcal{P}(V)$ such that*

C1 *For all* $w_1, w_2 \in W$ *and* $i \in I$ *we have that* $g(w_1)(g(w_2)') \vDash \Omega_i$ *iff* $w_1 \mathcal{R}_i w_2$.
C2 *For all* $w \in W$ *and* $p \in X$, *we have that* $p \in g(w)$ *iff* $p \in \pi(w)$.
C3 *For every* $s \subseteq V$, s *is a state of* $\mathcal{F}$ *iff* $s = g(w)$ *for some* $w \in W$.

Then, for every $\mathcal{L}(X)$*-formula* φ *we have* $(\mathcal{F}, g(w)) \vDash \varphi$ *iff* $(\mathcal{M}, w) \vDash \varphi$.

Definition 8 (Act)**.** *Given an event* $(\mathcal{X} = (V^+, \theta^+, V_-, \theta_-, \Omega^+), x)$, *define an action* $(\mathsf{Act}(\mathcal{X}) := (A, \mathsf{pre}, \mathsf{post}, R), a := x)$ *by*

- $A := \mathcal{P}(V^+)$
- $\mathsf{pre}(a) := [a/\top]\,[(V^+ \setminus a)/\bot]\,\theta^+$
- $\mathsf{post}_a(p) := \begin{cases} [a/\top]\,[(V^+ \setminus a)/\bot]\,(\theta_-(p)) & \text{if } p \in V_- \\ p & \text{otherwise} \end{cases}$
- $R_i := \{(a, b) \mid a \cup (b') \vDash \Omega_i^+\}$

Definition 9 (Trf)**.** *Consider an action* $(\mathcal{A} = (A, \mathsf{pre}, \mathsf{post}, R), a_0)$. *Let* $n := \mathsf{ceil}(\log_2 |A|)$ *and* $\ell : A \to \mathcal{P}(\{q_1, \ldots, q_n\})$ *be an injective labeling function using fresh atomic variables* q_k. *Then let* $(\mathsf{Trf}(\mathcal{A}) := (V^+, \theta^+, V_-, \theta_-, \Omega^+), x := \ell(a_0))$ *be the event defined by*

- $V^+ := \{q_1, \ldots, q_n\}$
- $\theta^+ := \bigvee_{a \in A} (\mathsf{pre}(a) \land \ell(a) \sqsubseteq V^\circ)$
- $V_- := \{p \in V \mid \exists a : \mathsf{post}_a(p) \neq p\}$
- $\theta_-(p) := \bigvee_{a \in A} (\ell(a) \sqsubseteq V^+ \land \mathsf{post}_a(p))$
- $\Omega_i^+ := \bigvee_{(a,b) \in R_i} (a \sqsubseteq V^+ \land (b \sqsubseteq V^+)')$

Besides these translations for the dynamic parts, we also use the translations $\mathcal{M}(\cdot)$ and $\mathcal{F}(\cdot)$ from structures to models and vice versa, as given in Definitions 18 and 19 of [4]. Now everything is in place to state and prove our main result. The following generalizes Theorem 4 in [4].

Theorem 2. *(i) Definition 8 is truthful: For any scene* $(\mathcal{F}, s)$, *any event* $(\mathcal{X}, x)$ *and any formula* φ *over the vocabulary of* $\mathcal{F}$ *we have:*

$$(\mathcal{F}, s) \times (\mathcal{X}, x) \vDash \varphi \iff (\mathcal{M}(\mathcal{F}), s) \times (\mathsf{Act}(\mathcal{X}), x) \vDash \varphi$$

(ii) Definition 9 is truthful: For any pointed Kripke model $(\mathcal{M}, w)$, *any action* $(\mathcal{A}, a)$ *and any formula* φ *over the vocabulary of* $\mathcal{M}$ *we have:*

$$(\mathcal{M} \times \mathcal{A}, (w, a)) \vDash \varphi \iff (\mathcal{F}(\mathcal{M}), g_{\mathcal{M}}(w)) \times (\mathsf{Trf}(\mathcal{A}), \ell(a)) \vDash \varphi$$

where $g_{\mathcal{M}}$ *is from* $\mathcal{F}(\mathcal{M})$ *in Definition 19 of [4] and* $\mathsf{Trf}(\mathcal{A})$ *and* ℓ *are from Definition 9 above.*

Proof. By Lemma 2. We first need appropriate functions g.

For part (i), g needs to map worlds of $\mathcal{M}(\mathcal{F}) \times \mathsf{Act}(\mathcal{X})$, i.e. pairs $(s, x) \in \mathcal{P}(V) \times \mathcal{P}(V^+)$ to states of $\mathcal{F} \times \mathcal{X}$, i.e. subsets of $V \cup V^+ \cup V_-^\circ$. Let $g(s, x) := (s \setminus V_-) \cup (s \cap V_-)^\circ \cup x \cup \{p \in V_- \mid s \cup x \vDash \theta_-(p)\}$ which is exactly s^{new} from Definition 7 above. We now prove C1 to C3 from Lemma 2.

For C1, take any two worlds (s,x) and (t,y). We need to show $g(s,x)(g(t,y))' \vDash \Omega_i^{\mathrm{new}}$ iff $\mathcal{R}_i^{\mathrm{new}}(s,x)(t,y)$. For this, note the following equivalences. We have $g(s,x)(g(t,y))' \vDash \Omega_i^{\mathrm{new}}$ iff

$$\begin{aligned}&(s \setminus V_-) \cup (s \cap V_-)^\circ \cup x \cup \{p \in V_- \mid s \cup x \vDash \theta_-(p)\}\\ \cup\ &((t \setminus V_-) \cup (t \cap V_-)^\circ \cup y \cup \{p \in V_- \mid t \cup y \vDash \theta_-(p)\})'\\ \vDash\ &\left[V_-/V_-^\circ\right]\left[(V_-)'/(V_-^\circ)'\right] \Omega_i \wedge \Omega_i^+\end{aligned}$$

Here V_- and V_-' do not occur in the formula, as old epistemic relations do not depend on new values of modified propositions. Hence we can drop the subsets of V_- and V_-' to obtain the equivalent condition

$$(s\setminus V_-)\cup(s \cap V_-)^\circ\cup x\cup(t\setminus V_-)'\cup(t^\circ\cap V_-^\circ)'\cup y' \vDash \left[V_-/V_-^\circ\right]\left[(V_-)'/(V_-^\circ)'\right] \Omega_i\wedge\Omega_i^+$$

in which we can split both sides into separate vocabularies:

$$(s\setminus V_-)\cup(s \cap V_-)^\circ\cup(t\setminus V_-)'\cup(t^\circ\cap V_-^\circ)' \vDash \left[V_-/V_-^\circ\right]\left[(V_-)'/(V_-^\circ)'\right] \Omega_i \text{ and } x\cup y' \vDash \Omega_i^+$$

Now undo the $\circ$-substitution on both sides in the first conjunct to see that it is equivalent to $s\cup t' \vDash \Omega_i$. Hence the whole condition is equivalent to $\mathcal{R}_i st$ and $R_i xy$ which is exactly $\mathcal{R}_i^{\mathrm{new}}(s,x)(t,y)$ by Definition of $\mathcal{M}(\cdot)$ and Definition 8.

To show C2, take any (s,x) and any $p \in V$. We have to show that $p \in g(s,x)$ iff $p \in \pi^{\mathrm{new}}(s,x) = \{p \in V \mid \mathcal{M}, s \vDash \mathsf{post}_x(p)\}$. There are two cases. First, if $p \notin V_-$, then $\mathsf{post}_x(p) = p$ by Definition 8 and we directly have $p \in g(s,x)$ iff $p \in s$ iff $\mathcal{M}, s \vDash p$ iff $p \in \pi^{\mathrm{new}}(s,x)$. Second, if $p \in V_-$, then $p \in g(s,x)$ iff $s \cup x \vDash \theta_-(p)$ by definition of g and $\mathsf{post}_x(p) = [x/\top]\,[(V^+ \setminus x)/\bot]\,\theta_-(p)$ by Definition 8. Hence we have a chain of equivalences: $p \in g(s,x)$ iff $s \cup x \vDash \theta_-(p)$ iff $s \vDash [x/\top]\,[(V^+ \setminus x)/\bot]\,\theta_-(p)$ iff $\mathcal{M}, s \vDash [x/\top]\,[(V^+ \setminus x)/\bot]\,\theta_-(p)$ iff $p \in \pi^{\mathrm{new}}(s,x)$.

For C3, take any $s^{\mathrm{new}} \subseteq V \cup V^+ \cup V_-^\circ$. We want to show that $s^{\mathrm{new}} \vDash \theta^{\mathrm{new}}$ iff there is an (s,x) such that $g(s,x) = s^{\mathrm{new}}$.

For left-to-right, suppose $s^{\mathrm{new}} \vDash \theta^{\mathrm{new}}$, i.e. $s^{\mathrm{new}} \vDash \left[V_-/V_-^\circ\right](\theta \wedge \|\theta^+\|_{\mathcal{F}}) \ \wedge \ \bigwedge_{q \in V^-}\left(q \leftrightarrow \left[V_-/V_-^\circ\right](\theta_-(q))\right)$. Now first, let $s^{\mathrm{old}} := (s^{\mathrm{new}} \cap \setminus V_-) \cup \{p \in V_- \mid p^\circ \in s^{\mathrm{new}}\}$. We then have $s^{\mathrm{old}} \vDash \theta$, i.e. s^{old} is a state of $\mathcal{F}$ and thus by the definition of $\mathcal{M}(\cdot)$ also a world of $\mathcal{M}(\mathcal{F})$. Second, let $x := s^{\mathrm{new}} \cap V^+$ and note that $s \cup x \vDash \|\theta^+\|_{\mathcal{F}}$. It can now be checked that $g(s,x) = s^{\mathrm{new}}$.

For right-to-left, suppose we have an (s,x) such that $g(s,x) = s^{\mathrm{new}}$. Then we want to show $(s \setminus V_-) \cup (s \cap V_-)^\circ \cup x \cup \{p \in V_- \mid s \cup x \vDash \theta_-(p)\} \vDash \left[V_-/V_-^\circ\right](\theta \wedge \|\theta^+\|_{\mathcal{F}}) \ \wedge \ \bigwedge_{q \in V^-}\left(q \leftrightarrow \left[V_-/V_-^\circ\right](\theta_-(q))\right)$ which indeed follows from $s \vDash \theta$ and Definition 8.

For part (ii), g should map worlds of $\mathcal{M} \times \mathcal{A}$ to states of $\mathcal{F}(\mathcal{M}) \times \mathsf{Trf}(\mathcal{A})$. Again we use s^{new}, but s and x are given by propositional encodings $g_{\mathcal{M}}(w)$ and $\ell(a)$. Let $g(w,a) := (g_{\mathcal{M}}(w)\setminus V_-)\cup(g_{\mathcal{M}}(w) \cap V_-)^\circ\cup\ell(a)\cup\{p \in V_- \mid g_{\mathcal{M}}(w)\cup\ell(a) \vDash \theta_-(p)\}$. We leave checking C1 to C3 as an exercise to the reader — the proofs are very similar to those in part (i). □

6 Symbolic Sally-Anne

The Sally-Anne false belief task is a famous example used to illustrate and test for a theory of mind. The basic version goes as follows (adapted from [2]):

> Sally has a basket, Anne has a box. Sally also has a marble and puts it in her basket. Then Sally goes out for a walk. Anne moves the marble from the basket into the Box. Now Sally comes back and wants to get her marble. Where will she look for it?

To answer this, one needs to realize that Sally did not observe that the marble was moved and will thus look for it in the basket. We now translate the first DEL modeling of this story from [6] to our framework. This choice is also motivated by a recent interest in the complexity of theory of mind [12,13] where our symbolic representation might provide a new perspective. For simplicity we adopt the naive modeling given in [6], leaving it as future work to also adopt the refinement with edge-conditions and other improvements of the model.

We use the vocabulary $V = \{p, t\}$ where p means that Sally is in the room and t that the marble is in the basket. In the initial scene Sally is in the room, the marble is not in the basket and both of this is common knowledge:

$$(\mathcal{F}_0, s_0) = ((V = \{p, t\},\ \theta = (p \land \neg t),\ \Omega_{\mathrm{S}} = \top,\ \Omega_{\mathrm{A}} = \top), \{p\})$$

The sequence of events is:

$\mathcal{X}_1$: Sally puts the marble in the basket: $((\varnothing, \top, \{t\}, \theta_-(t) = \top, \top, \top), \varnothing)$.
$\mathcal{X}_2$: Sally leaves: $((\varnothing, \top, \{p\}, \theta_-(p) = \bot, \top, \top), \varnothing)$.
$\mathcal{X}_3$: Anne puts the marble in the box, not observed by Sally:
$((\{q\}, \top, \{t\}, \theta_-(t) = (\neg q \to t) \land (q \to \bot), \neg q', q \leftrightarrow q'), \{q\})$.
$\mathcal{X}_4$: Sally comes back: $((\varnothing, \top, \{p\}, \theta_-(p) = \top, \top, \top), \varnothing)$.

We calculate the result in Figure 2, using Lemma 1 to remove superfluous variables. Note that all operations are boolean. Finally, we can check that in the last scene Sally believes the marble is in the basket:

$$\begin{aligned}
& \{p, q\} \vDash \Box_{\mathrm{S}} t \\
\iff & \{p, q\} \vDash \forall V'(\theta' \to (\Omega_{\mathrm{S}} \to t')) \\
\iff & \{p, q\} \vDash \forall \{p', t', q'\}((t' \leftrightarrow \neg q') \land p' \to (\neg q' \to t')) \\
\iff & \{p, q\} \vDash \top
\end{aligned}$$

7 Related and Future Work

We generalized knowledge transformers from [4] to *belief transformers with factual change*. The result is a new symbolic representation of action models with postconditions that can be implemented using binary decision diagrams [7].

$$
\begin{array}{lll}
 & ((\{p,t\},(p\wedge\neg t),\top,\top),p) & \mathcal{F}_0\\
\times & ((\varnothing,\top,\{t\},\theta_-(t)=\top,\top,\top),\varnothing) & \mathcal{X}_1\\
= & ((\{p,t,t^\circ\},(p\wedge\neg t^\circ)\wedge t,\top,\top),\{p,t\}) & \\
 & & \\
\times & ((\varnothing,\top,\{p\},\theta_-(p)=\bot,\top,\top),\varnothing) & \mathcal{X}_2\\
= & ((\{p,t,t^\circ,p^\circ\},(p^\circ\wedge\neg t^\circ)\wedge t\wedge\neg p,\top,\top),\{t,p^\circ\}) & \\
\equiv_V & ((\{p,t\},t\wedge\neg p,\top,\top),\{t\}) & \\
 & & \\
\times & ((\{q\},\top,\{t\},\theta_-(t)=(\neg q\to t)\wedge(q\to\bot),\neg q',q\leftrightarrow q'),\{q\}) & \mathcal{X}_3\\
= & ((\{p,t,q,t^\circ\},t^\circ\wedge\neg p\wedge(t\leftrightarrow((\neg q\to t^\circ)\wedge(q\to\bot))),\neg q',q\leftrightarrow q'),\{q\}) & \\
= & ((\{p,t,q,t^\circ\},t^\circ\wedge\neg p\wedge(t\leftrightarrow\neg q),\neg q',q\leftrightarrow q'),\{q\}) & \\
\equiv_V & ((\{p,t,q\},\neg p\wedge(t\leftrightarrow\neg q),\neg q',q\leftrightarrow q'),\{q\}) & \\
 & & \\
\times & ((\varnothing,\top,\{p\},\theta_-(p)=\top,\top,\top),\varnothing) & \mathcal{X}_4\\
= & ((\{p,t,q,p^\circ\},\neg p^\circ\wedge(t\leftrightarrow\neg q)\wedge p,\neg q',q\leftrightarrow q'),\{p,q\}) & \\
\equiv_V & ((\{p,t,q\},(t\leftrightarrow\neg q)\wedge p,\neg q',q\leftrightarrow q'),\{p,q\}) &
\end{array}
$$

Fig. 2. Sally-Anne on belief structures and transformers.

As mentioned above, restricting postconditions to boolean formulas does not limit the expressivity. The authors of [10] in fact prove the stronger result that postconditions can be restricted to $\top$ and $\bot$. Hence one can also model postconditions as functions of the type $A \to \mathcal{P}(V)$ as done in [6]. We leave it as future work to tune the definition of transformers in a similar way.

An alternative "succinct" representation for Kripke models and action models was recently developed in [8]. Succinct models also describe sets of worlds with boolean formulas, but instead of observational variables or boolean formulas over a double vocabulary they use *mental programs* to encode accessibility relations. Notably, model checking DEL is still in PSPACE when models and actions are represented succinctly. No complexity is known for our structures and transformers so far, but we expect it to be the same as for succinct models and actions.

Finally, the presented ideas are of course meant to be implemented. A natural next step therefore is to extend SMCDEL [11], the implementation of [4], with the presented transformers, working towards a symbolic model checker covering the whole *Logic of Communication and Change* from [5]. This work has been started and experimental modules including the Sally-Anne example are now available at `https://github.com/jrclogic/SMCDEL`.

Additionally, an implementation of [8] would be interesting to compare the performance of both approaches. Benchmark problems can be taken from both the DEL and the cognition literature, see for example [13].

Acknowledgements. Many thanks to Fernando R. Velázquez Quesada, Jan van Eijck and the anonymous reviewers for helpful comments and suggestions.

References

1. Baltag, A., Moss, L.S., Solecki, S.: The logic of public announcements, common knowledge, and private suspicions. In: Bilboa, I. (ed.) TARK'98. pp. 43–56 (1998), https://dl.acm.org/citation.cfm?id=645876.671885
2. Baron-Cohen, S., Leslie, A.M., Frith, U.: Does the autistic child have a "theory of mind"? Cognition 21(1), 37–46 (1985), https://doi.org/10/c2nhzk
3. van Benthem, J., van Eijck, J., Gattinger, M., Su, K.: Symbolic Model Checking for Dynamic Epistemic Logic. In: van der Hoek, W., Holliday, H.W., Wang, W.f. (eds.) Proceedings of The Fifth International Conference on Logic, Rationality and Interaction (LORI-V) in Taipei, Taiwan, October 28-30, 2015. pp. 366–378 (2015), https://doi.org/bzb6
4. van Benthem, J., van Eijck, J., Gattinger, M., Su, K.: Symbolic Model Checking for Dynamic Epistemic Logic – S5 and Beyond. Journal of Logic and Computation (JLC) (to appear), http://homepages.cwi.nl/~jve/papers/16/pdfs/2016-05-23-del-bdd-lori-journal.pdf
5. van Benthem, J., van Eijck, J., Kooi, B.: Logics of communication and change. Information and computation 204(11), 1620–1662 (2006), https://doi.org/d3j48n
6. Bolander, T.: Seeing is believing: Formalising false-belief tasks in dynamic epistemic logic. In: Herzig, A., Lorini, E. (eds.) Proceedings of the European Conference on Social Intelligence (ECSI-2014). pp. 246–263 (2014), http://ceur-ws.org/Vol-1283/paper_14.pdf
7. Bryant, R.E.: Graph-Based Algorithms for Boolean Function Manipulation. IEEE Transaction on Computers C-35(8), 677–691 (1986), https://doi.org/bnrh63
8. Charrier, T., Schwarzentruber, F.: A succinct language for dynamic epistemic logic. In: Proceedings of the 16th Conference on Autonomous Agents and Multiagent Systems. pp. 123–131. AAMAS '17, International Foundation for Autonomous Agents and Multiagent Systems (2017), http://www.aamas2017.org/proceedings/pdfs/p123.pdf
9. van Ditmarsch, H., van der Hoek, W., Kooi, B.: Dynamic epistemic logic, vol. 1. Springer Heidelberg (2007), https://doi.org/fdhbwt
10. van Ditmarsch, H., Kooi, B.: Semantic results for ontic and epistemic change. Logic and the Foundations of Game and Decision Theory (LOFT 7) 3, 87–117 (2008)
11. Gattinger, M.: SMCDEL – An Implementation of Symbolic Model Checking for Dynamic Epistemic Logic with Binary Decision Diagrams, https://github.com/jrclogic/SMCDEL
12. van de Pol, I.: How Difficult is it to Think that you Think that I Think that...? A DEL-based Computational-level Model of Theory of Mind and its Complexity. Master's thesis, University of Amsterdam (2015), https://www.illc.uva.nl/Research/Reports/MoL-2015-14.text.pdf
13. van de Pol, I., van Rooij, I., Szymanik, J.: Parameterized complexity results for a model of theory of mind based on dynamic epistemic logic. In: Proceedings Fifteenth Conference on Theoretical Aspects of Rationality and Knowledge, TARK 2015, Carnegie Mellon University, Pittsburgh, USA, June 4-6, 2015. pp. 246–263 (2015), https://doi.org/bzb7