Look and Learn: Extracting Information from Actions of Others[1]

Rasmus K. Rendsvig

Abstract

Dynamics pertaining to learning from the actions of others are modeled with focus on a running example, showing how we may come to conclusions regarding an ontic fact solely by observing others act on their beliefs regarding the fact in question. The tools used are elements from dynamic epistemic logic, with the addition of *decision* and *interpretation rules*, allowing local, rule-based choice and reasoning about such. The paper reports on work in progress.

1 Introduction

We learn by watching others.[1] We mimic language and movements, absorb norms and customs. In general, we may learn a lot from watching the actions of others, and in many day-to-day practical settings, observing those around us is an easy way of obtaining information: people leaving during dinner may indicate the direction of the bathroom, great sales may be taken to indicate a great book, and a jaywalker that crossing the street is safe despite a red light.

Though extraction of information from action is a fundamental learning method,[2] the deductive mechanics involved in a reconstruction or explication of such inferences is complex, drawing on underlying assumptions about intentions, plans and rationality, as well as beliefs and higher-order beliefs about such, on behalf of those involved. Seeing a performed action does not *a priori* tell us much, but coupled with suitable beliefs about the possible purpose of the action together with a belief that the actor is rational, leaves us with tentative information regarding the actor's beliefs. Given further beliefs about the verisimilitude of the beliefs of the actor may yield a belief regarding a proposition crucial to performing a given action; e.g. coming to believe that the bathroom is in the back.

Taking the actions of others as indications of either prevailing norms or ontic facts has been described as accepting *social proof* on the matter [5]. Social proof is often transmitted non-linguistically, i.e. by the actions chosen.

In this paper, the dynamics of learning from the actions of others by social proof is discussed by modeling a storyline using tools from *Dynamic Epistemic Logic* [1]. In this process, we introduce *decision rules*, a variant of *knowledge-based programs* [9], suitable for specifying agent actions in the DEL framework, *interpretation rules*, facilitating abductive inference from actions to beliefs, and a *belief adoption policy* allowing inter-agent belief transmission.

The former two devices are introduced due to the *localized* perspective of DEL. Where extensive game trees and branching-time epistemic temporal logic would allow choices to be made on the basis of rationality and expected utility, and interpretation of actions to be done in the style of forward induction, this is not possible in DEL as the terminal histories that grant pay-offs simply does not exist. Each DEL model is restricted to a specific point in time, with no embedded information pertaining to future developments. This 'shortcoming'

[1]A previous version of this paper was presented at GIRL'13@LUND (Games, Interactive Rationality and Learning, Lund 2013), the participants of which are thanked for helpful comments.

[2]Though far from all extraction of information via *social proof* ensures healthy learning; see e.g. [10] for examples of how social proof may lead to 'irrational' group behavior.

is extremely useful, as it allows one to work with small, easily digestible models. It is the main goal of this paper to introduce modules suitably localized for DEL that allow modeling choice and interpretation.

The paper is structured around modeling the following storyline. Two agents are standing at a red light, both wanting to cross the street. However, neither crosses as they do not know whether to do so is safe. This requires representation of knowledge, for which *epistemic plausibility models* are used. The first agent, call her a, looks about and concludes that crossing is indeed safe and so walks. The belief update is modeled using *action plausibility models* and *action-priority update* [1], whereas the choice of what to do following update is determined by the agents *decision rules*, for which a formalism is introduced. The second agent, b, now contemplates and then crosses the street without orienting himself. Why? His action is due to the epistemic influence a's action has on b. Seeing a's orientation act and subsequent choice to walk makes b reason that a believes doing so is safe, something b now takes a's word (well, action) for. To facilitate b's reasoning from a's action to a's belief, *action interpretation rules* are introduced, specifying a hierarchy of abductive hypotheses. To what degree b should adopt the perceived beliefs of a may again be controlled by a set of decision rules (and a *belief adoption policy*), and so may b's decision to cross the street or not. In the second to last section, it is shown how the model may be extended to allow for multiple initial walkers. The main technical and conceptual achievement of the paper is the specification of decision and interpretation rules, allowing local, rule-based choice and reasoning about such. The aim is not to construct an explanatory model of jaywalking, but to present *an addition to DEL* that allows us to model dynamics involving decision, action interpretation and hence social proof, with the future aim of modeling informational phenomena from e.g. social psychology.

2 Epistemic States and Belief Update Transitions

The initial state in the jaywalker storyline involves two agents and their mutual lack of knowledge regarding the safety of street crossing. This static epistemic state may be represented using an *epistemic plausibility model.*

Epistemic Plausibility Models. A *(finite) epistemic plausibility frame* (EPF) is a structure $(S, \leq_i)_{i\in\mathcal{A}}$ where S is a finite set of *worlds* with typical elements s, t, and $\leq_i$ is a *preorder* on S for each *agent* $i \in \mathcal{A}$. A *(pointed) epistemic plausibility model* (EPM) is an EPF augmented with i) a *valuation set* $\|\Phi\|$ consistent of a *doxastic proposition* P for every *atom* in Φ, where P is a map assigning to every EPF S a subset $P_{\mathsf{S}} \subseteq S$, and ii) a designated state s_0 called *the actual world.* An EPM is typically denoted $\mathbf{S} = (S, \leq_i, \|\Phi\|, s_0)_{i\in\mathcal{A}}$.

Given an EPM, the *indistinguishability relation* for agent i is the equivalence relation $\sim_i := \leq_i \cup \geq_i$. Further, the *information cell* of agent i at state s is $\mathcal{K}_i[s] = \{t : s \sim_i t\}$ and the *plausibility cell* of agent i at state s is $\mathcal{B}_i[s] = Min_{\leq_i}\mathcal{K}_i[s] = \{t \in \mathcal{K}_i[s] : t \leq_i s', \text{ for all } s' \in \mathcal{K}_i[s]\}$. The plausibility cell $\mathcal{B}_i[s]$ contains the worlds the agent find *most plausible* from the information cell $\mathcal{K}_i[s]$ and represent the "doxastic appearance" [1, p. 25] of s to i.[3] Notice that $s \leq_i t$ means that s is at *least as plausible* as t for i.

Doxastic Propositions. Where $P, Q \in \Phi$, let the set Prop of doxastic propositions be given by $\varphi := \top \,|\, \bot \,|\, P \,|\, \neg\varphi \,|\, \varphi \wedge \psi \,|\, B_i\varphi \,|\, K_i\varphi$, specified for individual models by $\top_{\mathbf{S}} = S, \bot = \emptyset, P_{\mathbf{S}} := P_{\mathsf{S}}, (\neg\varphi)_{\mathbf{S}} := S \backslash \varphi_{\mathbf{S}}, (\varphi \wedge \psi)_{\mathbf{S}} := \varphi_{\mathbf{S}} \cap \psi_{\mathbf{S}}, (K_i\varphi)_{\mathbf{S}} := \{s \in S : \mathcal{K}_i[s] \subseteq \varphi_{\mathbf{S}}\}, (B_i\varphi)_{\mathbf{S}} := \{s \in S : \mathcal{B}_i[s] \subseteq \varphi_{\mathbf{S}}\}$. Boolean connectives are defined as usual. A proposition P *is true*

[3] The definition of EPMs is based on [1], with slight alterations and omissions due to page constraints. The notation for information and plausibility cells are adopted from [6].

at state s in model $\mathbf{S}$ iff $s \in P_\mathbf{S}$, also written $\mathbf{S}, s \models P$. *Entailment* is given by $\varphi \models \psi$ iff $\varphi_\mathbf{S} \subseteq \psi_\mathbf{S}$ for all $\mathbf{S}$.

Note two things: first, every doxastic proposition specifies a regular modal logical proposition (i.e. a set of worlds) relative to each model. Doxastic propositions simply allow looking at the same (doxastic) proposition across models. Second, K_iP and B_iP are equivalent to standard notions from epistemic logic: a world is a K_iP/B_iP-world just in case every world in i's information/plausibility cell is a P-world. For all intents, K_i and B_i reflect respectively S5 and S4 modalities with $K_i\varphi \rightarrow B_i\varphi$ valid.

Initial Uncertainty. The initial state of uncertainty may be represented as an EPM $\mathbf{S_0} = (S_0, \leq_i, \|\Phi\|, s)_{i\in\{a,b\}}$ with $S_0 = \{s,t\}$, taking the atom $S \in \Phi$, read 'it is safe to cross the street', to have assignment $S_{S_0} = \{s\}$. By the latter, $(\neg S)_{S_0} = \{t\}$. This model is illustrated in Figure 1.[4]

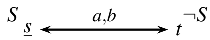

Fig. 1. The EPM $\mathbf{S_0}$. Labeled arrows represent plausibility relations, with $s \longleftarrow_i t$ depicting that $s \leq_i t$. Reflexive loops are omitted for all but singleton information cells. States are labeled with their true atoms or the true negations of atoms (here S at s and $\neg S$ at t). The actual world in underlined.

In $\mathbf{S_0}$, neither a nor b knows whether it is safe to cross or not, as both $s, t \in \mathcal{K}_a[s] = \mathcal{K}_b[s]$. Hence $\mathbf{S_0}, s \models \neg K_iS \wedge \neg K_i\neg S$, for $i \in \{a,b\}$. Neither do either agent believe that it is safe to cross, as s and t are equi-plausible for both, as indicated by the bidirectional arrow.

To capture the event that a orientates herself about safety, we use models very similar in structure to EPMs, but where states represent *ongoing actions/events*, rather than static worlds.

Action Plausibility Models. A *(pointed) action plausibility model* (APM) is a structure $\mathbf{E} = (\Sigma, \preceq_i, pre, post, \sigma_0)_{i\in\mathcal{A}}$ where Σ is a finite set of *actions* with typical elements σ, τ, each $\preceq_i$ is a pre-order on Σ, $pre : \Sigma \longrightarrow \mathsf{Prop}$ is a *precondition map* and $post : \Sigma \longrightarrow \mathsf{Prop}$ a *postcondition map* such that $post(\sigma) = \psi$ where $\psi \in \{\top, \bot\}$ or $\psi = \bigwedge_1^n \varphi_n$ with $\varphi_i \in \{P, \neg P : P \in \Phi\}$. Finally, σ_0 is *the actual event*. Note that $(\Sigma, \preceq_i)_{i\in\mathcal{A}}$ is an epistemic plausibility frame.[5]

Just as every world in an EPM represents a possible state of affairs, specified by the world's true propositions, so every action in an APM represents a possible *change*. *What* change is specified by the pre- and postconditions; preconditions determine what is required for the given action to take place, i.e. what conditions a world must satisfy for an action to executable in that world, and postconditions what *factual* change the action brings about.[6] The event where a checks traffic is captured by the APM $\mathbf{E_1}$, illustrated in Figure 2.

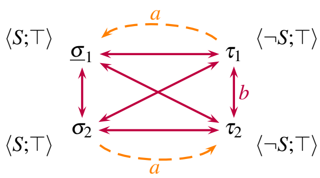

Fig. 2. The APM $\mathbf{E_1}$ representing a's orientation act, and both agents uncertainty about what is happening. State labels $\langle\varphi;\psi\rangle$ specify pre- and postconditions.

Expl.: Agent a checks whether it is safe to cross or not. In fact, she sees that it is safe (σ_1), a judgment she trusts, but not completely (hence $\sigma_1 \preceq_a \tau_1$). b sees that a looks, but cannot tell what she sees. As the act of observation changes no ontic facts: postconditions are empty ($\top$) for all events.

---

[4] More atoms are introduced later; all are assumed false at $\mathbf{S_0}$.

[5] Again, this presentation follows [1], with the addition of postconditions as formulated in [7, 4].

[6] To exemplify, the action 'agent a plays a Queen' is only executable when a has a Queen on hand (precondition), and brings the factual change that the given Queen has now been played (postcondition).

$\mathbf{E_1}$ includes uncertainty for both agents, but this may be restricted by looking at *doxastic programs* over $\mathbf{E_1}$. A doxastic program is the action model equivalent of a proposition, i.e. a subset of all actions in the models' event space: $\Gamma \subseteq \Sigma$. Over $\mathbf{E_1}$, the program $\Gamma_1 = \{\sigma_1, \tau_1\}$ captures the event where a sees it's safe, but is still uncertain, and b can tell that a sees either that it's safe or not, that she is not certain, but leans towards σ_1. $\Delta_1 = \{\sigma_1, \tau_2\}$ captures that a sees it's safe, and is certain about this, and b sees that a either sees that it's safe or not and that a is certain about which, while b cannot tell which is seen. For the present case, either Δ_1 or $\mathbf{E_1}$ seems the reasonable modeling choice[7], depending on how much credit is given to a's eyesight, and focus will be on the latter. To incorporate this new information into $\mathbf{S_0}$, the *APU product* of the two models is taken.

Action-Priority Update Product. The *action-priority update* is a binary operation $\otimes$ with first argument an EPM $\mathbf{S}$ and second argument a doxastic program $\Gamma \subseteq \Sigma$ over some APM $\mathbf{E}$ with action space Σ. The *APU product* is an EPM $\mathbf{S} \otimes \Gamma = (S \otimes \Gamma, \leq_i^{\uparrow}, \|\Phi\|^{\uparrow}, (s_0, \sigma_0))$ where the updated state space is $S \otimes \Gamma = \{(s, \sigma) \in S \times \Gamma : \mathbf{S}, s \models pre(\sigma)\}$; each updated pre-order $\leq_i^{\uparrow}$ is given by $(s, \sigma) \leq_i^{\uparrow} (t, \tau)$ iff either $\sigma \prec_i \tau$ and $s \sim_i t$, or else $\sigma \simeq_i \tau$ and $s \leq_i t$;[8] the valuation set $\|\Phi\|^{\uparrow}$ is given by the following: for every atom $P \in \Phi$, $P_{\mathbf{S}\otimes\Gamma} = (\{(s, \sigma) : s \in P_{\mathbf{S}}\} \backslash \{(s, \sigma) : post(\sigma) \models \neg P\}) \cup \{(s, \sigma) : post(e) \models P\}$ for states $(s, \sigma) \in S \otimes \Sigma$. Finally, (s_0, σ_0) is the new actual world.

The APU product gives priority to new information encoded in Γ over the old beliefs from $\mathbf{S}$ by the 'anti-lexicographic' specification of $\leq_i^{\uparrow}$ that gives priority to the APM plausibility relation $\preceq_i$. The definition further clarifies the role of pre- and postconditions; if a world does not satisfy the preconditions of an action, then the given state-action pair does not survive the update, and if postconditions are specified, these override earlier ontic facts, else leave all as was. The definition is based on [1] for the APU product with the valuation clause from [7, 4].

After Orientation. Updating $\mathbf{S_0}$ with $\mathbf{E_1}$ produces the EPM $\mathbf{S_1} := \mathbf{S_0} \otimes \mathbf{E_1}$, depicted in Figure 3, with a frame structure identical to that of $\mathbf{E_1}$ with $S_{\mathbf{S_0}\otimes\mathbf{E_1}} = \{(s, \sigma_1), (s, \sigma_2)\}$ and $\neg S_{\mathbf{S_0}\otimes\mathbf{E_1}} = \{(t, \sigma_1), (t, \sigma_2)\}$. In the actual state (s, σ_1), the following are all satisfied: $B_a S \wedge \neg K_a S$, $\neg B_b S \wedge \neg B_b \neg S$, $\neg B_b B_a S \wedge \neg B_b B_a \neg S$, $K_b(B_a S \vee B_a \neg S) \vee \neg K_b B_a S \wedge \neg K_b B_a \neg S$. That is, a believes S, but doesn't know; b believes neither S nor $\neg S$; nor does b have specific beliefs about a's beliefs about S; though he knows that a believes either S or $\neg S$.

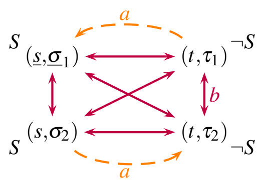

Fig. 3. The updated EPM $\mathbf{S_1} := \mathbf{S_0} \otimes \mathbf{E_1}$.

3 Acting on Beliefs: Decision Rules

Agent a stands to choose between either walking or staying put, her choice done *publicly* and her action post-factually represented by an atom, W_a or $\overline{W}_a$.[9] *Ex post*, the action should

[7] Γ_1 is too strong for present purposes as it would entail that b learns that a believes S following the update. This would skip chapters of the jaywalker story being told.

[8] $\preceq_i$ is from $\mathbf{E}$ and $\leq_i$ from $\mathbf{S}$. $\sigma \prec_i \tau$ denotes ($\sigma \preceq_i \tau$ and not $\sigma \succeq_i \tau$), $\sigma \simeq_i \tau$ denotes ($\sigma \preceq_i \tau$ and $\sigma \succeq_i \tau$).

[9] On the propositions W_a and $\overline{W}_a$: as the event of walking takes place *in* an APM, the action *has been executed* in the resulting APU product model. Hence a past-tense reading of W_a is in order: a has walked/a

hence be known to all. A suitable APM for both actions is depicted in Fig. 4. Which is actually executed depends on which singleton doxastic program, $\{\gamma_1\}$ or $\{\gamma_2\}$, is chosen .

Fig. 4. The APM $\mathbf{E_2}$, representing a's two options: a may either choose to walk (γ_1) or may choose to not walk (γ_2). In either case, b has no uncertainty about what the ongoing event is: he clearly sees whether a walks or not.

Given her belief that S, for a to act reasonably, it is intuitively clear that the next transition should be an update with γ_1. However, simply performing this update *as modeler* does not present a with much of a *choice*. Put differently, if we as modelers have to inspect the model and hand pick a next update for each agent action, the agents are not very *autonomous*: their decision architecture is not incorporated in the sequence model, but only in the mind of the modeler.

One way incorporate the decision architecture of agents suitable for epistemic logic is the *knowledge-based programs* of [9], being simple directions of the form '**if** $B_a S$, **do** W_a', specifying an action based on local epistemic state. Defining such a rule for each relevant belief allows for the definition of various *agent types* with choices specified for also counterfactual situations. These rules may then be considered constituent parts of the sequence model, or 'system', exemplified below.[10] As DEL terminology includes doxastic programs, the term *transition rules* will be used to denote the version of [9]'s programs here tailored to the DEL framework.

Transition Rules. A *transition rule* $\mathcal{T}$ is an expression $\varphi \rightsquigarrow [X]\psi$ where $\varphi, \psi \in \mathsf{Prop}$. Call φ the *trigger* and ψ the *effect*. Typically, the trigger will be a proposition of the form $B_i\varphi'$ or $K_i\varphi'$, and if so, call the expression a *decision rule for agent i*. If EPM $\mathbf{S}$ satisfies the trigger of some transition rule $\mathcal{T}$ at $\mathbf{S}$'s actual world, $\mathcal{T}$ is said to be *active* in $\mathbf{S}$ (else *inactive*).

In deciding whether to cross the street, the only determining factor for a is whether she believes to do so is safe or not. These beliefs are naturally correlated with walking/not walking by decision rules $\mathcal{D}_1$ and $\mathcal{D}_2$:

$$\begin{aligned} \mathcal{D}_1 : \quad & B_a S \rightsquigarrow [X]W_a \\ \mathcal{D}_2 : \quad & B_a \neg S \rightsquigarrow [X]\overline{W}_a \end{aligned}$$

Here, $\mathcal{D}_1$ is read "if a believes S, then choose the next action such that after it, a will have walked." Note that any set of decision rules may be taken as defining an *agent type*; some may act oddly, some even have inconsistent directions, and some, like the type given by $\{\mathcal{D}_1, \mathcal{D}_2\}$, will capture a reasonable behavior pattern.

Dynamic Modalities. Note that *decision rules are not doxastic propositions*: the "modality" $[X]$ has no interpretation, and construed as a formula, $\mathcal{D}_1$ has no truth conditions. Instead, transition rules are *prescriptions* for choosing the next action model.[11] The choice of model is made by implementing a transition rule over an EPM $\mathbf{S}$ and a set $\mathbf{G}$ of doxastic programs over one or more APMs using *dynamic modalities.* For any program Γ over APM $\mathbf{E}$, $[\Gamma]$ is a dynamic modality, and the doxastic proposition $[\Gamma]\varphi$ is given by

chose to walk. With the latter reading, it is natural to differentiate between $\neg W_a$ and $\overline{W}_a$, with the first stating that a has not chosen to walk, and the latter that a has chosen not to walk. This makes $\neg W_a \wedge \neg\overline{W}_a$ consistent, as one would expect (e.g.: in $\mathbf{S_0}$, a has chosen *nothing*). It is however natural to require that $(W_a \wedge \overline{W}_a)_{\mathbf{S}} = \emptyset$ for all $\mathbf{S}$.

[10] No definition of *system* is given here; see [14] for details and an application to *informational cascades.*

[11] Or more generally, the next *model transformer*, i.e. function from EPM to EPM, of which many types may be found in the DEL literature. Any APM together with the APU product induces such a function.

$([\Gamma]\varphi)_{\mathbf{S}} := \{s \in S : \forall \sigma \in \Gamma,$ if $(s,\sigma) \in S \otimes \Gamma$ then $(s,\sigma) \in \varphi_{\mathbf{S}\otimes\Gamma}\}$. That is, a world s from $\mathbf{S}$ is a $[\Gamma]\varphi$-world iff every resolution of Γ over s is a φ-world in $\mathbf{S} \otimes \Gamma$.

Solutions and Next APM Choice. A set of transition rules dictates the choice for next APM by finding the transition rule(s)'s *solution*. A *solution* to $\mathcal{T} = \varphi \rightsquigarrow [X]\psi$ over pointed EPM $(\mathbf{S}, s)$ is a doxastic program Γ such that $\mathbf{S}, s \models \varphi \to [\Gamma]\psi$. Γ is a solution to the set $\boldsymbol{T}= \{\mathcal{T}_1, ..., \mathcal{T}_n\}$ with $\mathcal{T}_i = \varphi_i \rightsquigarrow [X]\psi_i$ over $(\mathbf{S}, s)$ if $\mathbf{S}, s \models \bigwedge_1^n (\varphi_i \to [\Gamma_i]\psi_i)$, i.e. if Γ is a solution to all $\mathcal{T}_i$ over $(\mathbf{S}, s)$ *simultaneously.*[12] Finally, a set of doxastic programs $\mathbf{G}$ is a solution to $\boldsymbol{T}$ over $\mathbf{S}$ iff for every t of $\mathbf{S}$, there is a $\Gamma \in \mathbf{G}$ such that Γ is a solution to $\boldsymbol{T}$ over $(\mathbf{S}, t)$.[13,14]

Where $\mathbf{G}$ is a solution to $\boldsymbol{T}$ over $\mathbf{S}$, let the *next APM choice of* $(\mathbf{S}, s_0)$ be a solution to $\boldsymbol{T}$ over $(\mathbf{S}, s_0)$, with s_0 the actual world of $\mathbf{S}$.

If $\mathbf{G}$ is a solution to $\boldsymbol{T}$ over $\mathbf{S}$, then for each state from $\mathbf{S}$, the transition rules in $\boldsymbol{T}$ will specify one (or more) programs from $\mathbf{G}$ as the next choice. A deterministic choice will be made if $\mathbf{G}$ is selected suitably, in the sense that it contains a unique Γ for each s. Note that a next choice specified for each state makes the evolving model sequence sensitive which state is actual. Hereby next choice becomes 'localized': change which state is actual, and a different choice may be made.

Example: Looping System. Consider the very simple 'system', consistent of an EPM $\mathbf{S}$ with $s_0 \in P_{\mathbf{S}}$, and APM $\mathbf{E} \in \mathsf{APM}$ with $pre(\sigma_0) = P$, $post(\sigma_0) = \neg P$, and the set $\boldsymbol{T}= \{\mathcal{T}_0, \mathcal{T}_1\}$ of transition rules:

$$\mathcal{T}_0 = P \rightsquigarrow [X]\neg P \qquad \mathcal{T}_1 = \neg P \rightsquigarrow [X]P$$

$\mathbf{S}$:

s_0
P

$\mathbf{E}$:

σ_0	σ_1
$\langle P; \neg P\rangle$	$\langle \neg P; P\rangle$

With $\Gamma_0 = \{\sigma_0\}$ and $\Gamma_1 = \{\sigma_1\}$, $\mathbf{G} = \{\Gamma_0, \Gamma_1\}$ is a solution to $\boldsymbol{T}$ over $\mathbf{S}$. For $\mathcal{T}_1$, $\mathbf{S}, s_0 \models \neg P \to [\Gamma_0]P$ as $s_0 \notin (\neg P)_{\mathbf{S}}$. For $\mathcal{T}_0$, it is easy to check that $\mathbf{S} \otimes \Gamma, (s_0, \sigma_0) \models \neg P$, why $\mathbf{S}, s_0 \models [\Gamma_0]\neg P$. As Γ_0 is unique, this is chosen as next update. It should be easy to see that $\mathbf{G}$ is also a solution to $\boldsymbol{T}$ over $\mathbf{S} \otimes \Gamma_0$, where Γ_1 is chosen. Further re-application of $\boldsymbol{T}$ loops the system.

Choosing to Walk. For $\mathbf{DR}_a = \{\mathcal{D}_1, \mathcal{D}_2\}$ over $\mathbf{S_0} \otimes \mathbf{E_1}$, the set $\mathbf{G}_a$ consistent of singleton programs $\{\gamma_1\}, \{\gamma_2\}$ over APM $\mathbf{E_2}$, as depicted in Fig. 4, is a solution. Consulting the actual world of $\mathbf{S_1}$, only $\{\gamma_1\}$ is a solution to $\mathbf{DR}_a$: as $(s, \sigma_1) \models B_a S$, the antecedent of $\mathcal{D}_2$ is false, and hence $\{\gamma_1\}$ is a solution to this rule. For $\mathcal{D}_1$, $\{\gamma_1\}$ will ensure the satisfaction of the consequent by the postconditions of γ_1. Further, $\{\gamma_1\}$ is the unique solution as may easily be checked. Hence the (deterministic) next APM choice of $(\mathbf{S_1}, (s, \sigma_1))$ is $\{\gamma_1\}$.

In short, given her belief that it is safe, a's decision rules dictates the choice to walk. Clearly, this choice could simply have been dictated by the modeler; the inclusion of decision rules however allows the choice to be made autonomously by the 'system'. This also goes for counter-factual cases: had a had a different belief, she would have chosen differently.

[12] Note the analogy with numerical equations; for both $2 + x = 5$ and $\{2 + x = 5, 4 + x = 7\}$, $x = 3$ is the (unique) solution.

[13] A broader solution space may possibly be used, replacing the role of doxastic programs with *model transformers*, i.e. maps $f(\mathbf{S}) = \mathbf{S}'$. A doxastic program applied using the APU product is such an f. The action model solution space may then be refined by using a notion of *action model equivalence* and *action emulation* [8] suitable altered to APMs with postconditions.

[14] The latter is defined thusly to ensure that $\mathbf{G}$ includes actions for both active and inactive decision rules. Inactive decision rules does not affect the choice, but if solutions to these are not included, agents would have no possible choice in counter-factual situations (or runs with other initial states).

Transition Rules, Runs, Protocols and Systems. The primary role of transition rules is to provide a specification of the next APM to be added to the model sequence, where the choice is determined by what propositions the latest model in the sequence satisfies. As such, transition rules may be seen as partial functions from locally satisfied conditions to actions, *a la* the *programs* of [9]. Notions of *runs, protocols* and *systems* based on the local next model choice provided by a set of transition rules may therefore be definable. One important obstacle for doing so is *concurrent choice*, i.e. cases with multiple active rules. Apart from the obvious problem of two active decision rules having inconsistent effects, there is a challenge in "joining solutions". Where each decision rule in a set **DR** of active rules with consistent effects each has a solution in **G**, it is not guaranteed that **G** will be a set solution to **DR** as two (or more) rules may each pick a different $\Gamma \in \mathbf{G}$. Moreover, it is not obvious how a "joint solution" should be constructed from individual choices.[15] In the ensuing, these problems will not arise as only one decision rule will be active at the time and no attempt to define the mentioned notions is made. It is hoped that putting the present approach in relation to recent work on DEL protocols (see e.g. [3, 6]) will shed light on these issues.

4 From Action to Belief: Abduction and Adoption

Updating $\mathbf{S_0} \otimes \mathbf{E_1}$ with a's choice $\{\gamma_1\}$ produces the EPM $\mathbf{S_2} := \mathbf{S_1} \otimes \{\gamma_1\}$, identical to $\mathbf{S_1}$ except that all states in $\mathbf{S_2}$ satisfy W_a. The model is depicted in Fig. 5.

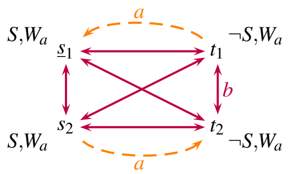

Fig. 5. The EPM $\mathbf{S_2} := \mathbf{S_1} \otimes \{\gamma_1\}$: no change has occurred but the switch in truth value of W_a. From this it follows that both agents now know that a has walked.

For future reference, it is noted that $s_1 = ((s, \sigma_1), \gamma_1)$, $s_2 = ((s, \sigma_2), \gamma_1)$ and $t_1 = ((t, \tau_1), \gamma_1)$, $t_2 = ((t, \tau_2), \gamma_1)$.

Following a's action, her choice to walk is known to both agents, but b has not received any new information about a's beliefs pertaining to safety. Neither does b have any means of deducing these beliefs given the introduced formal framework. Such a deduction would require e.g. the ability to rationalize by *forward induction*, which require information from both past play and future possibilities [2] couched in a game framework representing preferences, rationality and more (see e.g. [16]). Though structures akin to game trees can be defined using EPMs, APMs and protocols [6], a simpler, more superficial construct may be used to facilitate the reasoning. The suggested approach utilizes an 'inverse' version of decision rules, brute forcing conclusions about belief from observations about action.

In making decisions, our beliefs about the relevant state of affairs dictate our action, up to error and human factors. Hence the route from beliefs to actions can be mapped as a function. As this function will often not be a injection, moving from actions to beliefs is not as straight forward, as multiple different belief states may result in the same action. Having to provide a rationalization of a given action will therefore often include abductive reasoning. An abductive hypothesis to rationalize an action allows inferring as explanation of the action a belief state of the acting agent. Below, such hypotheses are called *interpretation rules.*

Interpretation Rules. An *interpretation rule* is a doxastic proposition $\varphi \to [\mathbf{S}]B_i\psi$, where we call φ the *basis* and ψ the *content.* The idea behind interpretations rules is that

[15] See [15] for a suggestion in relation to *the bystander effect*, utilizing Cartesian products of APMs.

on the basis of an action (e.g. W_a), agents must deduce something about the content of i's beliefs (e.g. that B_iS). Doxastic propositions involving the modality of the consequent are given by $([\mathbf{S}]\chi)_{\mathbf{S'}} := \{s' \in S' : s' \in s \text{ and } (s, \mathbf{S}) \models \chi\}$, where $s' \in s$ means that s is a *predecessor*[16] of s'. Hence $[\mathbf{S}]\chi$ is true in $(\mathbf{S'}, s')$ just in case s''s predecessor in $\mathbf{S}$ was a χ-world. The modality is included to respect the temporal aspect introduced by updates, and $\mathbf{S}$ is to be substituted with the EPM based on which i made the choice in question.

A set of interpretation rules is implemented using an APM where the preconditions of each state is a conjunction of interpretation rules with different bases. Hereby each state represents a different way of interpreting the possible 'action propositions' included in the bases (W_a and $\overline{W}_a$ in the running example). The plausibility order of the APM is defined depending on what picture agents should obtain about how agent i makes decisions, about i's "type".

Why did a cross the street? A plausible explanation for b of a's choice to walk would be that she believed it safe, but a could for all b knows be of an odd agent type that walks when they believe it to be unsafe. The latter, however, might seem implausible to b. Likewise, a plausible explanation of a not walking could be that she believed it unsafe, and less plausibly that a is odd and chooses to not walk if safe. Assuming that b only considers it possible that a is either consistently odd or not odd at all (normal), and that a knows whether she is odd or normal, these interpretation rules can be implemented by updating $\mathbf{S_2}$ with $\mathbf{E_3}$, depicted in Fig. 5.

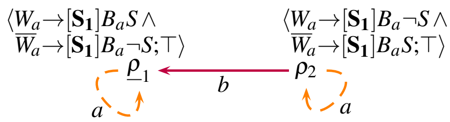

Fig. 6. The APM $\mathbf{E_3}$ representing a and b's interpretations of a's possible actions. There is no uncertainty for a, but b considers it possible that a is either normal (ρ_1), or that she is odd (ρ_2).

From Then to Now. In $\mathbf{S_2}$, only in the states s_1, t_1 has agent a behaved normally and only in states s_2, t_2 has she behaved oddly. Only s_1, t_1 satisfy $[\mathbf{S_1}]B_aS$: their respective predecessors (s, σ_1) and (t, τ_1) satisfy B_iS is $\mathbf{S_1}$. This is not the case for s_2, t_2: their predecessors both belong to $(B_a\neg S)_{\mathbf{S_1}}$. Hence $(pre(\rho_1))_{\mathbf{S_1}} = \{s_1, t_1\}$ and $(pre(\rho_2))_{\mathbf{S_1}} = \{s_2, t_2\}$, why $\mathbf{S_3} := \mathbf{S_2} \otimes \mathbf{E_3}$ again contains four states, as illustrated in Figure 7.

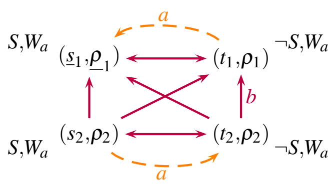

Fig. 7. The EPM $\mathbf{S_3} := \mathbf{S_2} \otimes \mathbf{E_3}$ capturing a and b's epistemic states after b's interpretation of a's action. b now believes that a believes S.

Unrealistically, a now *knows* that b believes that she believes S. This could be eliminated using a larger model. This is noted and ignored.

Notice that b has now formed correct beliefs about both a's beliefs prior to her action $((s_1, \rho_1) \in ([\mathbf{S_1}]B_aS)_{\mathbf{S_3}})$, but also about a's beliefs *in the current state* $(([\mathbf{S_1}]B_aS)_{\mathbf{S_3}} = (B_aS)_{\mathbf{S_3}}\ \ (*))$. Furthermore, *$b$ knows that a's beliefs have not changed* (by $(*)$ and $([\mathbf{S_1}]B_a\neg S)_{\mathbf{S_3}} = (B_a\neg S)_{\mathbf{S_3}}$). This is more than what is enforced by the interpretation rules in general, and is a consequence of the fact that a has not changed her beliefs about S since $\mathbf{S_1}$. It eliminates an interesting temporal difficulty in the current story, namely *which* set of a's beliefs b should consider relevant to his own decision: those upon which a made

[16] When constructing APU products, a state in the product model is an ordered pair (s, σ) of a state s and and action σ. In this pair, s may again be such a pair. We say that a *predecessor* of s' is any s that occurs in any of the ordered pairs of s', including s' itself.

her decision, or her most recent? Clearly, if a learns that it is not safe to walk while doing so, this information would be of importance to b. It should not be portrayed by her action W_a, though, why b could find himself in a situation where he has a belief about a's earlier beliefs, but does not know whether they have changed.

Belief Adoption. Though b has formed the belief $B_b B_a S$ in $\mathbf{S_3}$, he has not changed his opinion regarding S; b neither believes nor disbelieves it. He has however obtained (fallible) "evidence" for the truth of S, namely that a believes it. In case b trusts a's beliefs to be more often correct than not, it seems reasonable for b to accept $B_a S$ as evidence for S if b has nothing else to go on. This belief adoption policy may be captured by the following decision rules:

$$\begin{array}{llll} \mathcal{D}_3: & B_b B_a S & \wedge \neg B_b \neg S & \rightsquigarrow [X] B_b S \\ \mathcal{D}_4: & B_b B_a \neg S & \wedge \neg B_b S & \rightsquigarrow [X] B_b \neg S \end{array}$$

Different adoption policies may be obtained by modifying the triggers. E.g., if one wishes that b adopts a's beliefs no matter what, the two last two conjuncts could be removed. Such rules would capture that b trusts a's judgment higher than his own.

A solution to $\{\mathcal{D}_3, \mathcal{D}_4\}$ over $\mathbf{S_3}$ is the set $\{\Delta_1, \Delta_2\}$ of programs over APM $\mathbf{E_4}$, Fig. 8, with $\Delta_1 = \{\mu_1, \eta_1, \tau\}$ and $\Delta_2 = \{\mu_2, \eta_2, \tau\}$.

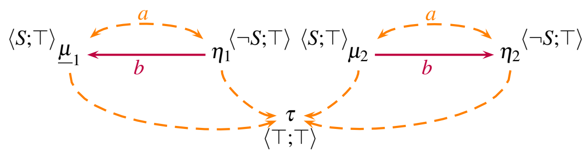

Fig. 8. The APM $\mathbf{E_4}$. The doxastic program $\Delta_1 = \{\mu_1, \eta_1, \tau\}$ captures that b makes a "soft update" with S ($\mu_1 \preceq_b \eta_1$), while it appears to a that nothing is going on (τ).

The interpretation of Δ_1 is that while b makes a soft update with S to align his beliefs with a's, a perceives the ongoing event as vacuous: with both pre- and postconditions set to $\top$, τ represents the event where *nothing* happens. It seems the natural choice that a should stay as unaware of b's revision as possible.

Following Footsteps. Given that $\mathbf{S_3}$ satisfies only the trigger of $\mathcal{D}_3$, the next choice of next update is the program Δ_1. In ensuing EPM $\mathbf{S_4} := \mathbf{S_3} \otimes \Delta_1$, b has updated his beliefs, and the actual world satisfies $S \wedge B_b S$. Provided that b makes his decision to walk or not using decision rules $\mathcal{D}_1$ and $\mathcal{D}_2$ with X ranging over the two programs from $\mathbf{E_2}$ (with $\mathcal{D}_1, \mathcal{D}_2, W_a$ and $\overline{W}_a$ indexed for b), he will choose to follow in the footsteps of a by also crossing the street.

5 Being Led by a Crowd

Extending the above model to one where b is in a bigger crowd is unproblematic, and only two aspects requires answering. First, how is concurrent choice to be implemented and how may such actions be interpreted by b? Second, how is b to treat social proof from a crowd? Here, a different belief adoption policy must be defined.

Concurrent Choice and Interpretation. Assume b is standing in a crowd $\mathcal{C}$ of agents, where each agent has orientated themselves as a did above. Let $\mathbf{C_1}$ be an EPM analogous to $\mathbf{S_1}$ above, in which $\mathcal{C}_S \subseteq \mathcal{C}$ believe it is safe to walk, $\mathcal{C}_{\neg S}$ believe is not and b is agnostic whether S, knows every other agent believes either S or $\neg S$, but does not know which. This state is easily constructed.

Assuming that every $i \in \mathcal{C}$ acts in accordance with $\mathcal{D}_1 : B_i S \rightsquigarrow [X]W_i$ and $\mathcal{D}_2 : B_i \neg S \rightsquigarrow \overline{W}_i$, a suitable solution space for any configuration of beliefs in $\mathcal{C}$ simply consists of $2^{\mathcal{C}}$ states, one for each possible conjunction over $\{W_i, \overline{W}_i : i \in \mathcal{C}\}$ with exactly one conjunct for each i, each state with such a conjunction as postcondition. Clearly, this model will serve as solution space for any distribution of beliefs about $S/\neg S$ in $\mathcal{C}$, and implementing $\mathcal{D}_1$ and $\mathcal{D}_2$ over $\mathbf{C_1}$ with this solution space will result in all i in $\mathcal{C}_S$ choosing W_i, while all i in $\mathcal{C}_{\neg S}$ will choose $\overline{W}_i$.

The actual world of the EPM after the concurrent choice will satisfy some conjunction over $\{W_i, \overline{W}_i : i \in \mathcal{C}\}$ which b must now interpret. The interpretation rules from the APM $\mathbf{E_3}$ may be reused for each $i \in \mathcal{C}$. A suitable 'abductive hierarchy' of agent type combinations in line with $\mathbf{E_3}$ may be constructed as the lattice with most plausible state that where all agents are 'normal' and least plausible state that where all agents are 'odd' and where $\rho <_b \rho'$ iff ρ has strictly less 'odd' agents than ρ', ρ and ρ' being equi-plausible just in case they have an equal number of 'odds'. Applying this interpretation, b will (as in the above) come to have a correct belief about the beliefs of $\mathcal{C}$ in the ensuing EPM. On this higher-order belief, b may now revise his agnostic stance.

Belief Adoption in the Crowd. Given that b is considers social proof a reasonable source of information and further assuming that he considers the members of $\mathcal{C}$ equally reliable as information gatherers, May's Theorem [11] provides good reason as to why b should apply simple majority voting on the issue of whether he should believe S or $\neg S$. More specifically, a belief adoption policy may be defined that makes b change his beliefs in accordance with the *perceived* majority (under different circumstances, b's higher-order beliefs could have been wrong).

To define the policy, we make use of the notion of *collective belief of degree* $\frac{m}{l}$ *that* φ from [13] given by

$$P_{\frac{m}{l}G}\varphi := \bigvee_{K \in \mathcal{K}} \bigwedge_{i \in K} B_i \varphi$$

where G is a group of agents and $\mathcal{K}$ is the set of all sets $K \subseteq G$ such that $|K| > \frac{m}{l}|G|$. Using degree $\frac{1}{2}$ makes $P_{\frac{G}{2}}\varphi$ true just in case a (strict) majority of G believes φ.

The majority voting belief adoption policy may now be implemented using the decision rules

$$\begin{array}{llll} \mathcal{D}'_3 : & B_b P_{\frac{\mathcal{C}}{2}} S & \wedge \neg B_b \neg S & \rightsquigarrow [X] B_b S \\ \mathcal{D}'_4 : & B_b P_{\frac{\mathcal{C}}{2}} \neg S & \wedge \neg B_b S & \rightsquigarrow [X] B_b \neg S \end{array}$$

with the same solution as above ($\mathbf{E_4}$), only with a indexes replaced by the group $\mathcal{C}$. Employing this policy, b will believe S just in case a strict majority of $\mathcal{C}$ is believed to do so, and *vice versa* for $\neg S$. In case of a tie in $\mathcal{C}$, neither $\mathcal{D}'_3$ or $\mathcal{D}'_4$ is active, and b will not perform an revision, but stay agnostic. Notice further that b's final beliefs will reflect the perceptions of agents in $\mathcal{C}$ during their initial orientation act: if the majority of these were correct, then b will cross the street safely.

6 Conclusion

The story of two jaywalkers crossing the street have been told using a sequence of models as storyboard, leading from their mutual uncertainty about the safety of crossing, to the decision of the second agent to walk. As it happens, the second agent got to cross the street safely, as he believed he would. Taking a bird's eye view of the model sequence, the interesting question now becomes *what made b succeed?* In essence, the answer to this question is that b learned (in a weak sense) from the action of a, which again invites the question of what made the learning successful. For this, several factors where determining.

First, a got lucky: in the initial orientation act, a formed the correct belief that crossing was safe. *Ceteris paribus*, if a had had a wrong perception of the event, either both agents would have walked while it was unsafe, or both would have stayed though they could have crossed the street safely. Generalizing from this single case, we may therefore postulate that in learning from the actions of others, the student's beliefs are only as good as the teacher's. In this case, learning from the actions of others merely reiterates the common sense lessons that basing your beliefs on misdirecting information will most likely lead to incorrect beliefs.

Second, b's first choice of abductive hypothesis regarding a's type was correct. Had b's primary interpretation rule not correctly mirrored a's decision rule, b would have come to the wrong conclusion about what a believed. All other being equal, b would have concluded that it was not safe to walk. This is a general consequence of the fact that *actions often underdetermine conclusions about the underlying beliefs*: where actions to beliefs is functional, belief to action may not be.

Third, the correctness of a's belief was stable over sufficient time. Had the street turned unsafe during a's crossing, then a's initial belief would stopped being correct without b being informed hereof. As a result, b would have formed beliefs based on "old" information, causing him to make an unsafe choice. Hence, in attempting to deduce ontic facts from the actions of others, one should try to ensure that conclusions are drawn presently.

Though these three points played important roles in the story, neither are *necessary* for the end result. Pertaining to the third point, time enough could have been made to pass for S to switch truth-value twice. For point one and two, two wrongs could here have made a right.

Future work. In the above, it has been shown how transition- and interpretation rules may be used to model dynamics involving decisions and decision interpretation, allowing subsequent agents to be informed by social proof. Many elements amiss for the presented to be a fully working framework. Most notably, definitions of systems and runs are required to rigidly investigate how the modules used in the modeling affect the overall picture, and how the present approach relates to DEL protocols and ETL [3, 6]. Being able to consider runs allows a macro view, where differences in final results may be mapped to changes in initial conditions, decision rules, etc. Taking a run-based approach would further allow one to investigate the role of interpretation rules. Though handy as a brute force approach to facilitate action to belief conclusions, a deeper analysis involving intentions, plans and rationality will render them obsolete, at the cost of working with large temporal structures, as opposed to the localized approach taken here. How these local and global perspectives relate is an obvious venue for further research.

The idea of treating decision rules as 'equations' requiring solutions from a given set or class of actions models (or model transformers) seems technically potent. It may be used e.g. to classify the 'strength' of an action model class in terms of what decision rules the class will contain solutions to. E.g., the class of action models without postconditions is weaker than that with, as the latter will contain solutions to the active rule $P \to [X]\neg P$ for atomic P, where the former will not. There are many venues of exploration: is there a 'complete' class of action models, containing a solution to $\varphi \to [X]\psi$ for all φ, ψ? Is there a weakest or minimal such? Is there a definable class of non-solvable rules? Etc.

Further, there are three strands of literature that should be consulted. The field of epistemic logic broadly construed contains much work that must surveyed, and relevant recommendations will be much appreciated. Further, game-theoretic work on *forward induction* (as reviewed in e.g. [12]) will be relevant, and thirdly, recent works on *observational learning* (e.g. [17]) from theoretical economy will also provide interesting perspectives. Hopefully, the author will get a chance to learn something from looking at these others' work.

References

[1] A. Baltag and S. Smets. A Qualitative Theory of Dynamic Interactive Belief Revision. In G. Bonanno, W. v.d.Hoek, and M. Wooldridge, editors, *Logic and the Foundations of Game and Decision Theory (LOFT 7)*, pages 9–58. Amsterdam Univ. Press, 2008.

[2] J. v. Benthem. *Logic in Games*. MIT Press, forthcoming.

[3] J. v. Benthem, J. Gerbrandy, T. Hoshi, and E. Pacuit. Merging Frameworks for Interaction. *Journal of Philosophical Logic*, 38(5):491–526, 2009.

[4] T. Bolander and M. Birkegaard. Epistemic planning for single- and multi-agent systems. *Journal of Applied Non-Classical Logics*, 21(1):9–34, 2011.

[5] R. Cialdini. *Influence: The Psychology of Persuasion*. HarperCollins Publishers, 2007.

[6] C. Dégremont. *The Temporal Mind: Observations on the logic of belief change in interactive systems*. PhD thesis, University of Amsterdam, 2010.

[7] H. v. Ditmarsch and B. Kooi. Semantic Results for Ontic and Epistemic Change. In G. Bonanno, W. v.d.Hoek, and M. Wooldridge, editors, *Logic and the Foundations of Game and Decision Theory (LOFT 7)*, pages 87–117. Amsterdam Univ. Press, 2008.

[8] J. Eijck, J. Ruan, and T. Sadzik. Action Emulation. *Synthese*, 185(S1):131–151, Mar. 2012.

[9] R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning About Knowledge*. The MIT Press, 1995.

[10] P. G. Hansen, V. F. Hendricks, and R. K. Rendsvig. Infostorms. *Metaphilosophy*, 44(3):301–326, 2013.

[11] K. O. May. A set of independent necessary and sufficient conditions for simple majority decisions. *Econometrica*, 20(4):680–684, 1952.

[12] A. Perea. *Rationality in Extensive Form Games*. Theory and Decision Library, Series C: Game Theory, Mathematical Programming and Operations Research. Kluwer Academic Publishers, 2001.

[13] C. Proietti and E. J. Olsson. A DDL Approach to Pluralistic Ignorance and Collective Belief. *Journal of Philosophical Logic*, March:1–17, 2013.

[14] R. K. Rendsvig. A Qualitative Model of Informational Cascades. *Submitted*.

[15] R. K. Rendsvig. Pluralistic Ignorance in the Bystander Effect. *Submitted*.

[16] R. K. Rendsvig and V. F. Hendricks. Social Proof in Extensive Games. *Submitted*.

[17] L. Smith and P. N. Sørensen. Observational Learning. In S. N. Durlauf and L. E. Blume, editors, *The New Palgrave Dictionary of Economics*. Palgrave Macmillan, online edition, 2011.

Rasmus K. Rendsvig
Dept. of Media, Cognition and Communication
University of Copenhagen
Njalsgade 80, 2300S Copenhagen, Denmark
Email: `rendsvig@gmail.com`