The Duality of State and Observation in Probabilistic Transition Systems

Monica Dinculescu[1], Christopher Hundt[2], Prakash Panangaden[3], Joelle Pineau[3], and Doina Precup[3]

[1] Morgan Stanley, Montreal, Canada
[2] Google Inc., Mountain View, California, USA
[3] McGill University, Montréal, Québec

Abstract. In this paper we consider the problem of representing and reasoning about systems, especially probabilistic systems, with hidden state. We consider transition systems where the state is not completely visible to an outside observer. Instead, there are observables that partly identify the state. We show that one can interchange the notions of state and observation and obtain what we call a dual system. In the case of deterministic systems, the double dual gives a minimal representation of the behaviour of the original system. We extend these ideas to probabilistic transition systems and to partially observable Markov decision processes (POMDPs).

1 Introduction

Learning and planning under uncertainty is a crucial focus of modern AI research. In the AI literature there is much discussion of the role of "state" and there is a point of view that asserts that "a canonical notion of state does not exist." According to this view, state is merely a sufficient statistic for predicting the future. This view has found its most articulate and developed exposition in what is called the "predictive representation of state" or PSRs [18, 24]. The present paper arose from an attempt to understand PSRs from a foundational point of view as well as to understand certain well-known learning algorithms [33] that seem implicitly to use some ideas from duality.

The main point of the present paper is the presence of a duality between state and observation which seems to lie at the heart of the PSR representation. The use of the word "duality" of course evokes connections with the duality between logic and transition systems and of the many dualities known in mathematics: Stone duality and its many variants and extensions, Priestley duality, Gelfand duality and general concrete dualities in category theory; see, for example, the excellent monograph of Johnstone [19] for a general categorical discussion. Duality has been very important in systems theory and expresses a relation between observability and controllability [21].

Much of the work in AI planning and learning under uncertainty is based on the framework of Partially Observable Markov Decision Processes (POMDPs) [20]. In this framework, problems are modeled using discrete states and actions.

Actions cause stochastic transitions between states. At each time step, a stochastic observation is also generated, based on the current state and the previous action. Much work has been devoted to planning in POMDPs when a model of the system (in terms of the stochastic transitions between states and the probability distributions over observations) is known. Unfortunately, learning POMDPs from data is a very difficult problem. One standard algorithmic solution is expectation maximization (EM) [11], but for POMDPs this approach is plagued by local minima (more so than for other probabilistic models) and works poorly in practice unless a good initial model of the system is used. History-based methods [25] often work better in practice, but are less general. A lot of recent research has been devoted to finding alternative representations for such systems, e.g., diversity-based representation [33], predictive state representations (PSRs) [24] and TD-networks [37]. These approaches aim to combine the generality of POMDPs with the ease of learning of history-based methods. The key idea underlying all of these approaches is that the state of the system is not considered as predefined; instead, it is viewed as a sufficient statistic for making (probabilistic) predictions about future trajectories. However, the models themselves are different and their relationships are only partially understood at the moment. It is our hope that the theory of the present paper will serve as a foundation for these different models and bring into focus the commonalities and differences.

In this paper, we develop a duality theory for POMDPs, which unifies much of the existing work on predictive representations. We show how, for any POMDP, one can develop two alternative representations: a dual machine and a double-dual machine. The key idea in the development is that of making measurements on the system, which we call experiments. Experiments are sequences of actions interspersed with observations. They generalize previous notions of tests from the literature on predictive state representations. Both of the alternative representations that we present allow an accurate prediction of the probability of any experiment. The double-dual representation is of particular interest, because it has a *deterministic* transition structure, and no hidden state. Instead, its states can be thought of as "bundles" of predictions for experiments. As such, this representation holds the promise of much better planning and learning algorithms than those currently available. Our work also generalizes similar representations from automata theory [10] and is closely related to the update graph from [33]. We show how existing predictive representations can be viewed from the perspective of this framework. We also discuss the implications of these alternative representations for learning algorithms, approximate planning algorithms as well as working with continuous observations. A preliminary version of these ideas has appeared in [17].

The main technical result of the present paper is that when one constructs the *double dual* one obtains a *minimal behaviourally equivalent* version of the original system. Of course, as written this cannot be quite right! One should get an isomorphic object when one goes back and forth across a dual situation. Nevertheless, this is what happened with the construction that we present in

this paper. This shows that what we have cannot be a pure Stone-type duality. It took a long time to fit these results into a coherent categorical picture. Of course, in the probabilistic setting the systems are infinite so "minimal" does not mean fewest states but is best expressed as a couniversal property; we will come back to this point in the conclusions.

The categorical version of the results of the present paper are being written up in a separate paper by the first author and Nick Bezhanishvili and Clemens Kupke. In that work the authors deal with weighted automata as well. The main idea there is that there is a duality between the transition system and an appropriate algebra *with additional operators*. Thus, for ordinary automata, one has a duality between the automata and boolean algebras equipped with "modal" operators while for the probabilistic case one has C^*-algebras equipped with additional operators. It turns out that the problem of minimizing a transition system can be seen in the dual category as the problem of finding a 0-generated subalgebra. Thus going to the dual, finding the 0-generated subalgebra and returning via the duality automatically minimizes the transition system.

2   Background

In this section we review the definitions of the various kinds of transition systems that we work with in this paper. The type of system most used in applications to planning under uncertainty and learning are partially observable Markov decision processes (POMDPs). There are, however, a number of simpler situations where the duality phenomenon also occurs and we will discuss duality for these systems before going on to POMDPs [20].

The crucial ingredient is the interplay between the dynamics, i.e. the state-to-state transitions, and the observations. In automata theory, the concept of partially observable system is often not made explicit in the definitions. In systems theory [21], and in artificial intelligence, one typically has observations that only partially reveal the state. One, thus, has to reason about behavioural equivalences between systems; bisimulation [26, 29, 30] is the most common of these equivalences[1]. In process algebra one has a similar situation: one does not see the state, only whether actions are "accepted" or "rejected" in a given state. These approaches have equivalent modelling power and concepts like bisimulation can be defined in both settings. In this paper we will always assume that all actions are possible in every state and the behavioural equivalences will take the observations into account. The usual example from process algebra [26, 27] showing the difference between trace equivalence and bisimulation can be mimicked in this setting.

The observations that we use are to be thought of as Boolean valued. Intuitively, one thinks of a black box with a number of buttons (the actions) and a number of lights (the observations). One can press any button and induce some

---

[1] Though [29] is often cited for bisimulation it is only mentioned there in passing and Park never wrote a paper on bisimulation; the slides [30] are the closest thing to a proper citation for Park's contribution.

kind of internal state transition - which will be invisible - but some of the lights may light up. In the most general case the observations depend on the action and the target state or source state of each transition. In the AI literature it is more usual to consider the observation as depending on the target (posterior) state and we will do that here when we discuss POMDPs. We discuss everything in the context of discrete state spaces so we avoid all measure-theoretic complications and we can work with state-to-state transition probabilities.

Of course, this is a first step. Eventually we will be interested in the case where the observables are real-valued – as in the case of rewards – or indeed vectors of real-values. We will also be interested in the case where the state space is continuous. In that case we will be concerned with developing a good approximation theory so that one can have tractable representations of continuous systems.

2.1 Kripke Automata

We begin with ordinary automata enriched with a notion of observations associated with each state; the observations are deterministic.

Definition 1. *A **deterministic Kripke automaton** (DKA) is a quintuple*

$$\mathcal{K} = (S, \mathcal{A}, \mathcal{O}, \delta : S \times \mathcal{A} \rightarrow S, \gamma : S \rightarrow 2^{\mathcal{O}}).$$

S is a set of states, $\mathcal{A}$ is a set of actions, $\mathcal{O}$ is a set of observations, δ is a transition function and γ is an observation function associated with the states.

The idea here is that there are a number of observations that will be made in a given state; in terms of our previous imagery, for each state there is a set of lights that are turned on in each state. If we call the observations "propositions" instead, this is essentially the definition of Kripke structure [12] used in model checking, except that there are labels on transitions as well, and we have made the transitions deterministic. The automata studied in undergraduate courses [16, 23, 35] are special cases where there is a single observation called "accept."

Note that the observations are associated with the states, like in a Moore machine. Given a state there is a set of observations that are *always* made in that state. One can view γ as a relation; we will switch between γ as defined above and $\hat{\gamma} \subseteq S \times \mathcal{O}$ where $\hat{\gamma}(s, \omega) \Leftrightarrow \omega \in \gamma(s)$. We will usually not even bother to write $\hat{\gamma}$ and just use γ for whatever version is most apt to the situation at hand.

One can take the view that with a state there is always a single observation: the complete description of which lights are on. The picture given above can be encoded in this view by taking the set of observations to be $2^{\mathcal{O}}$. Furthermore, the latter view is a special case of our definition: we are just restricting γ to be a function.

While these two views are the same in the non-probabilistic case, they differ sharply in the probabilistic case. If we were just to give a probability for a given

observation in a state, we could not express *correlations* between observations. Thus, when we come to the probabilistic case, we will insist that γ determines a distribution over all possible observations. Of course these observations could be "structured" in some way and we could analyze aspects of this structure. We are planning to explore these ideas in future work.

2.2   Probabilistic Systems

In systems theory, one often considers systems where the transitions and the observations are probabilistic. This gives a probabilistic version of Kripke automata; the interpretation of γ is, as we explained in the paragraph above, generalized to a distribution over the observations for each state.

Definition 2. *A **partially observable probabilistic automaton** (POPA) is a quintuple*

$$\mathcal{H} = (S, \mathcal{A}, \mathcal{O}, \tau : S \times \mathcal{A} \times S \to [0,1], \gamma : S \times \mathcal{O} \to [0,1])$$

where $\tau(s, a, \cdot)$ defines a probability distribution on possible target states and $\gamma(s, \omega)$ is the probability of observing ω in state s. We will often write $\tau_a(s, \cdot)$ for $\tau(s, a, \cdot)$. We write $\tau_a(s, X)$, where $X \subseteq S$, for $\sum_{t \in X} \tau_a(s, t)$.

In each state we can observe possibly several (or no) observations. The number $\gamma(s, \omega)$ is the probability that one sees ω, as opposed to not seeing it, given that one is in the state s. The function $\omega \mapsto \gamma(s, \omega)$ is not necessarily a probability distribution on $\mathcal{O}$. It defines a probability distribution on $2^{\mathcal{O}}$. Some of our constructions on partially observable probabilistic automata will yield deterministic variants. The interesting case is where the transitions are deterministic, but the observations are still probabilistic.

Definition 3. *A **deterministic automaton with stochastic observations** (DASO) is a quintuple*

$$\mathcal{J} = (S, \mathcal{A}, \mathcal{O}, \delta : S \times \mathcal{A} \to S, \gamma : S \times \mathcal{O} \to [0,1])$$

where $\gamma(s, \omega)$ is the probability of observing ω in state s.

Here the transitions are deterministic, hence, given by a function, but the states can only be partially known through the observations, which are stochastic.

In POMDPs the observations are associated with the transitions rather than with the states.

Definition 4. *A **Partially   Observable   Markov   Decision   Process** (POMDP) is a quintuple*

$$\mathcal{M} = (S, \mathcal{A}, \mathcal{O}, P : S \times \mathcal{A} \times S \to [0,1], \gamma : \mathcal{A} \times S \times \mathcal{O} \to [0,1])$$

where, as before, S is a set of states, $\mathcal{A}$ is a set of actions, $\mathcal{O}$ is a set of observations, P is the transition probability function and γ gives the observation probabilities. For each $s \in S$ and $a \in \mathcal{A}$, the function $\omega \mapsto \gamma(a, s, \omega)$ is a distribution on $\mathcal{O}$.

Note that here we are requiring that γ defines a distribution. This is done to match the definition used in the extant literature for POMDPs. However, when we construct duals and double duals this will no longer be the case. The number $\gamma(a, s, \omega)$ is the probability that one sees the observation ω, given that the system takes the action a and given that *one ends up in s* after the transition is complete[2]. According to our definition, transitions are triggered by the actions in $\mathcal{A}$; there is no attempt to model probability distributions over the actions. They are intended to be actions chosen by an adversary or scheduler (the verification viewpoint) or actions chosen by a policy external to the system definition (the planning viewpoint). The triple $(S, \mathcal{A}, P)$ forms a labelled Markov process [14] (LMP). If we had a reward as well, it would be a Markov decision process (MDP) [32]. The fact that the state is only partially observable is captured by the fact that $\mathcal{O}$ is different from S and there is no given bijection between S and $\mathcal{O}$.

3   Duality for Kripke Automata

We begin with a simple example which is instructive and gives a foretaste of the rest of the paper. We work with deterministic Kripke automata, i.e. ordinary automata enhanced with a notion of observation, and establish a very pleasing duality between state and observation. The situation is rather special, in that the duality relation is very tight: the stochastic case does not have all the features of the present case. The results of the present section are a slight generalization of a technique known to Brzozowski [10] in 1962. We defer a comparison to Brzozowski's work to the end of this section.

We recall the definitions. Let $\mathcal{K} = (S, \mathcal{A}, \mathcal{O}, \delta, \gamma)$ be a deterministic Kripke automaton. Here S is the set of states, $\mathcal{A}$ an alphabet of input symbols, $\mathcal{O}$ is a set of primitive observations, δ is the transition function $\delta : S \times \mathcal{A} \rightarrow S$ and $\gamma : S \rightarrow 2^{\mathcal{O}}$ is a labelling function. One can as well think of the elements of $\mathcal{O}$ as *propositions* capturing basic properties of the states or as *observations* – boolean-valued in this case – that one can make of the states. One can equally well think of γ as having the type $S \times \mathcal{O} \rightarrow \{T, F\}$. We will emphasize the notion of observation and testing rather than the equivalent notion of proposition and modal formula.

Thinking of the elements of $\mathcal{O}$ as basic observations, we can use them to define a family of tests. We define a *test t* according to the following grammar:

$$t ::== \omega \in \mathcal{O} \mid (a) \cdot t$$

where $a \in \mathcal{A}$.

We say that a state s satisfies or *passes* ω, written $s \models \omega$, if $\omega \in \gamma(s)$ or, equivalently, $\gamma(s, \omega) = T$. We say $s \models (a) \cdot t$ if $\delta(s, a) \models t$. We define $[\![t]\!]_{\mathcal{K}} = \{s \in S | s \models t\}$. Clearly this is exactly the same as defining the tests as modal formulas and defining satisfaction as above.

---

[2] It would be more natural, perhaps, to make this depend on the source state; we are following the convention used by AI researchers [20].

We define an equivalence relation $\sim_{\mathcal{K}}$ between *tests* as $t_1 \sim_{\mathcal{K}} t_2$ if $[\![t_1]\!]_{\mathcal{K}} = [\![t_2]\!]_{\mathcal{K}}$. Note that this allows us to identify an equivalence class for t with the set of states $[\![t]\!]_{\mathcal{K}}$ that satisfy t. Note that another way of defining this equivalence relation is

$$t \sim_{\mathcal{K}} t' := \forall s \in S.s \models t \iff s \models t'.$$

We also define an equivalence $\equiv$ between *states* in $\mathcal{K}$ as $s_1 \equiv s_2$ if for all tests t on $\mathcal{K}$, $s_1 \models t \iff s_2 \models t$. The equivalence relations $\sim$ and $\equiv$ are clearly closely related: they are the hinge of the duality between states and observations.

We say that $\mathcal{K}$ is *reduced* if it has no $\equiv$-equivalent states. Since there is more than just one observation, in general the relation $\equiv$ is finer than the usual equivalence of automata theory.

Finally, we say that two DKAs $\mathcal{K} = (S, \mathcal{A}, \mathcal{O}, \delta, \gamma)$ and $\mathcal{K}' = (S', \mathcal{A}', \mathcal{O}', \delta', \gamma')$ are *isomorphic* if $\mathcal{A} = \mathcal{A}'$, $\mathcal{O} = \mathcal{O}'$, and there exists a bijection $\phi : S \to S'$ such that, for all $s \in S$, $\gamma(s) = \gamma'(\phi(s))$ and for all $a \in \mathcal{A}$ $\phi(\delta(s, a)) = \delta'(\phi(s), a)$.

We define the dual construction as follows.

Definition 5. *Let $\mathcal{K}$ be a Kripke automaton $\mathcal{K} = (S, \mathcal{A}, \mathcal{O}, \delta, \gamma)$. Let T be the set of $\sim_{\mathcal{K}}$-equivalence classes of tests on $\mathcal{K}$. We define $\mathcal{K}' = (S', \mathcal{A}, \mathcal{O}', \delta', \gamma')$ as follows:*

- $S' = T = \{[\![t]\!]_{\mathcal{K}}\}$
- $\mathcal{O}' = S$
- $\delta'([\![t]\!]_{\mathcal{K}}, a) = [\![(a) \cdot t]\!]_{\mathcal{K}}, \forall [\![t]\!]_{\mathcal{K}} \in S', a \in \mathcal{A}$
- $\forall [\![t]\!]_{\mathcal{K}} \in S' \ \gamma'([\![t]\!]_{\mathcal{K}}) = [\![t]\!]_{\mathcal{K}}$ or $\gamma'([\![t]\!]_{\mathcal{K}}, s) = (s \models t)$

The somewhat strange-looking definition of γ' is to be understood as follows. In the machine $\mathcal{K}'$ the observations one can make of the state $[\![t]\!]$ are those states of $\mathcal{K}$ (which are the observations of $\mathcal{K}'$) that satisfy the test t; this set is exactly $[\![t]\!]$. We have interchanged the states and the observations; more precisely we have interchanged equivalence classes of tests - based on the observations - with the states. We have made the states of the old machine the observations of the dual machine. To see the remarkable effect of this interchange we consider the double dual. We will see that the double dual is the minimal automaton with the same behaviour.

Now consider $\mathcal{K}'' = (\mathcal{K}')'$, the dual of the dual. Its states are equivalence classes of $\mathcal{K}'$-tests. Each such class is identified with a set $[\![t']\!]_{\mathcal{K}'}$ of $\mathcal{K}'$-states by which tests in that class are satisfied, and each $\mathcal{K}'$-state is an equivalence class of $\mathcal{K}$-tests. Thus we can look at states in $\mathcal{K}''$ as collections of $\mathcal{K}$-test equivalence classes. It is with this perception in mind that we construct the *Sat* function. To avoid confusion, we will write $\hat{s}$ for a state $s \in S$ when viewed as an observation of $\mathcal{K}'$.

Definition 6. *Let $\mathcal{K}''$ be the double dual of $\mathcal{K}$. For any state $s \in S$ of $\mathcal{K}$ we define $Sat(s) = \{[\![t]\!]_{\mathcal{K}} | s \models t\}$.*

The next lemma shows that these sets are always states of the double dual. The following notation will be useful. A state s of $\mathcal{K}$ is an observation of $\mathcal{K}'$; recall that $\hat{s}$ is s viewed as an observation of $\mathcal{K}'$.

Lemma 7. *Let $s \in S$ be any state in the original automaton $\mathcal{K}$. Then $Sat(s)$ is a state in $\mathcal{K}''$.*

Proof. The observations in $\mathcal{K}'$ are the states in $\mathcal{K}$, i.e. elements of S. Recall that $\hat{s} \in \mathcal{O}'$ is the observation associated with s. Then $[\![\hat{s}]\!]_{\mathcal{K}'}$ is a state in $\mathcal{K}''$, and

$$[\![\hat{s}]\!]_{\mathcal{K}'} = \{[\![t]\!]_{\mathcal{K}} | [\![t]\!]_{\mathcal{K}} \models \hat{s}\} = \{[\![t]\!]_{\mathcal{K}} | s \in [\![t]\!]_{\mathcal{K}}\} = \{[\![t]\!]_{\mathcal{K}} | s \models t\} = Sat(s).$$

∎

In fact *all* the states of the double dual have this form.

Lemma 8. *Let $s'' = [\![t]\!]_{\mathcal{K}'} \in S''$ be any state in $\mathcal{K}''$. Then $s'' = Sat(s_t)$ for some state $s_t \in S$.*

Proof. The proof is by induction on the length of t. The base case is settled by Lemma 7.

Now suppose $t = (a) \cdot t'$ for some t'. Then, by the inductive hypothesis, there is some $s_{t'}$ such that $[\![t']\!]_{\mathcal{K}'} = Sat(s_{t'})$. Let $s_t = \delta(s_{t'}, a)$. Then

$$[\![t]\!]_{\mathcal{K}'} = \{[\![r]\!]_{\mathcal{K}} | [\![r]\!]_{\mathcal{K}} \models t\} = \{[\![r]\!]_{\mathcal{K}} | [\![r]\!]_{\mathcal{K}} \models (a) \cdot t'\} = \{[\![r]\!]_{\mathcal{K}} | \delta'([\![r]\!]_{\mathcal{K}}, a) \models t'\}$$
$$= \{[\![r]\!]_{\mathcal{K}} | [\![(a) \cdot r]\!]_{\mathcal{K}} \models t'\} = \{[\![r]\!]_{\mathcal{K}} | [\![(a) \cdot r]\!]_{\mathcal{K}} \in [\![t']\!]_{\mathcal{K}'}\}$$
$$= \{[\![r]\!]_{\mathcal{K}} | [\![(a) \cdot r]\!]_{\mathcal{K}} \in Sat(s_{t'})\} = \{[\![r]\!]_{\mathcal{K}} | s'_t \models (a) \cdot r\}$$
$$= \{[\![r]\!]_{\mathcal{K}} | \delta(s_{t'}, a) \models r\} = \{[\![r]\!]_{\mathcal{K}} | s_t \models r\} = Sat(s_t).$$

We have used the induction hypothesis in the first equality of the penultimate line. ∎

Now an immediate consequence of the definitions is:

Observation 9. $s \equiv s'$ *if and only if* $Sat(s) = Sat(s')$.

Now from Lemmas 7 and 8 and the observation we have the following corollary.

Corollary 10. *If $\mathcal{K}$ is reduced then Sat is a bijection from S to S''.*

Proof. Lemma 7 shows that $Sat : S \to S''$. The fact that $\mathcal{K}$ is reduced means that $s_1 \neq s_2 \implies Sat(s_1) \neq Sat(s_2)$, which by Observation 9 implies that Sat is injective. Lemma 8 shows that Sat is surjective. Thus Sat is a bijection from S to S''. ∎

The statement of the corollary can be strengthened to show that we actually have an isomorphism of DKAs between $\mathcal{K}$ and $\mathcal{K}''$. We know that the action set $\mathcal{A}$ is the same for both $\mathcal{K}$ and $\mathcal{K}''$, but the observation set $\mathcal{O}$ is not equal to the observation set $\mathcal{O}''$. However, we can transform $\mathcal{K}''$ by just restricting to the original observations in the following way: recall that an observation in $\mathcal{K}''$ is a state in $\mathcal{K}'$, which is an equivalence class of tests in $\mathcal{K}$. We let $\mathcal{O}$ be the observations for $\mathcal{K}''$ and we define a new observation function γ''_T for $\mathcal{K}''$ as

$$\forall s'' \in S'' \ \omega \in \gamma''_T(s'') \iff [\![\omega]\!]_{\mathcal{K}} \in \gamma''(s'').$$

Now we can define $\mathcal{K}''_T$.

Definition 11. $\mathcal{K}_T'' = (S'', \mathcal{A}, \mathcal{O}, \delta'', \gamma_T'')$, *with γ_T'' defined as above.*

This allows us to establish the required isomorphism.

Theorem 12. *Suppose $\mathcal{K}$ is reduced. Then $\mathcal{K}$ is isomorphic to $\mathcal{K}_T''$.*

Proof.   We see that $\mathcal{A}$ and $\mathcal{O}$ are the same in both $\mathcal{K}$ and $\mathcal{K}_T''$. We use *Sat* as our bijection from S to S''. It remains only to verify the properties of *Sat*. First, the observations: for all $s \in S$ and all $\omega \in \mathcal{O}$ the following holds:

$$\begin{aligned}
\omega \in \gamma(s) &\iff s \models \omega \\
&\iff [\![\omega]\!]_\mathcal{K} \in Sat(s) \\
&\iff [\![\omega]\!]_\mathcal{K} \in \gamma''(Sat(s)) \\
&\iff \omega \in \gamma_T''(Sat(s)).
\end{aligned}$$

We now check the transitions. For any $s \in S$ and any $a \in \mathcal{A}$, let t be such that $Sat(s) = [\![t]\!]_{\mathcal{K}'}$. Then

$$\begin{aligned}
\delta''(Sat(s), a) &= \delta''([\![t]\!]_{\mathcal{K}'}, a) \\
&= [\![(a) \cdot t]\!]_{\mathcal{K}'} \\
&= \{[\![r]\!]_\mathcal{K} | [\![r]\!]_\mathcal{K} \models (a) \cdot t\} \\
&= \{[\![r]\!]_\mathcal{K} | [\![(a) \cdot r]\!]_\mathcal{K} \models t\} \\
&= \{[\![r]\!]_\mathcal{K} | [\![(a) \cdot r]\!]_\mathcal{K} \in [\![t]\!]_{\mathcal{K}'}\} \\
&= \{[\![r]\!]_\mathcal{K} | [\![(a) \cdot r]\!]_\mathcal{K} \in Sat(s)\} \\
&= \{[\![r]\!]_\mathcal{K} | s \models (a) \cdot r\} \\
&= \{[\![r]\!]_\mathcal{K} | \delta(s, a) \models r\} \\
&= Sat(\delta(s, a)).
\end{aligned}$$

Thus *Sat* establishes an isomorphism between $\mathcal{K}$ and $\mathcal{K}_T''$.   ∎

Thus the double dual construction produces a machine which is - in a very strong sense - the "minimal version" of the original machine. What we mean by minimal is that no further reduction or collapsing of states is possible. We will expand on this in the conclusions.

   In the stochastic case we will not get such a tight correspondence but this gives a preview of what will happen there. In fact analogous results work for the nondeterministic case; in this case the double dual is the minimized version of the equivalent deterministic machine.

An Extended Example. We will explain the concept of duality on a concrete example, using the finite automaton below, where $\mathcal{S} = \{s_1, \ldots, s_6\}$, $\mathcal{A} = \{0, 1\}$ and $\mathcal{O} = \{\mathbf{G}, \mathbf{R}, \mathbf{Y}\}$.
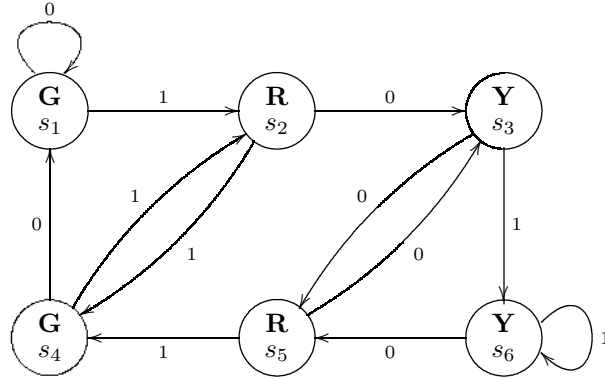
In the above machine $010\mathbf{Y}$, $\mathbf{R}$, and $110\mathbf{R}$ are examples of tests. Suppose that we start with state s_1 in our example. If we follow the sequence of actions 0010, we end up in state s_3, which gives an observation of $\mathbf{Y}$. Thus, we can say that $s_1 \models 0010\mathbf{Y}$. It is easy to see that $s_4 \models 01\mathbf{R}$, $s_1 \models 011\mathbf{G}$ and $s_6 \models 1100\mathbf{Y}$.

Consider the test $0\mathbf{G}$ in the above example. We notice that only s_1 and s_4 satisfy it. In order to find other tests equivalent to it, we should look at tests that only s_1 and s_4, and no other states, satisfy. Other such tests are $\mathbf{G}, 00\mathbf{G}, 000\mathbf{G} \ldots$ etc. Thus, we can say that $[\mathbf{G}] = [0\mathbf{G}] = [00\mathbf{G}] \ldots$. Similarly, we find that only the states s_2 and s_5 satisfy the equivalent tests $[\mathbf{R}] = [11\mathbf{R}] = [101\mathbf{R}] \ldots$, and the states s_3 and s_6 satisfy $[\mathbf{Y}] = [1\mathbf{Y}] = [100\mathbf{Y}] \ldots$.

As we have said before, an equivalence class of tests is identified by the set of states that satisfy these tests. Then, in our example, the equivalence classes of tests are: $t_1 = \{s_1, s_4\}$, $t_2 = \{s_2, s_5\}$ and $t_3 = \{s_3, s_6\}$.

We use the usual labelled transition notion: $s_1 \xrightarrow{a} s_2$ when a transition on action a has source s_1 and target s_2. Notice that not only $s_1 \xrightarrow{0} s_1$, and $s_4 \xrightarrow{0} s_1$, but $s_1 \xrightarrow{1} s_2$, and $s_4 \xrightarrow{1} s_2$ as well. Furthermore, these states have the same observations. This means that s_1 and s_4 have the same behaviour, and thus any test satisfied by one must also be satisfied by the other. Thus, $s_1 \equiv s_4$. Following the same reasoning, we can say that $s_2 \equiv s_5$ and $s_3 \equiv s_6$.

We now construct the **dual machine** of our example described above. Recall that $\mathcal{K} = (S, \mathcal{A}, \mathcal{O}, \delta, \gamma)$. The dual of $\mathcal{K}$ is $\mathcal{K}' = (S', \mathcal{A}, \mathcal{O}', \delta', \gamma')$ where:

$$S' = T = \{[t]\}$$
$$\mathcal{O}' = S$$
$$\delta'([t], a) = [at]$$
$$\gamma'([t], s) = [t]$$

The states of the new machine are the equivalence classes of the original machine, and the new transition function is defined to work with them; the observations are the states of the original machine. What observations do we see? Since the

new states of the dual machine are the equivalence classes $[t]$ of the old one, then the observations that we see should be the states that satisfy t; this is just what is given by the new observation function γ'.

We now construct the dual machine. In the original machine, we had determined that we have 3 equivalence classes: $[t_1]$, $[t_2]$ and $[t_3]$. Thus, the new machine has those three states. In addition, the observations are the states of the original machine that satisfy each of the equivalence classes, so transition arrows aside, the dual looks like:

We now add the transition arrows to the dual.

$$\delta([t_1], 0) = [0t_1]$$

We know that $[t_1] = [\mathbf{G}]$, therefore $[0t_1] = [0\mathbf{G}]$. To find out what equivalence class this test belongs to, we have to go through the test backwards. In order to see an observation of $\mathbf{G}$, we have to be in either state s_1 or s_4. The action that produced the $\mathbf{Y}$ observation was 0, so consider what happens if we take a 0 action backwards from s_1: $s_1 \overset{0}{\leftarrow} s_1$ or $s_1 \overset{0}{\leftarrow} s_4$. Note that there is no state s such that $s_4 \overset{0}{\leftarrow} s$. Thus, whatever equivalence class contains $[0\mathbf{G}]$ is identified by the states $\{s_1, s_4\}$. This is just t_1 itself, so we have a 0 transition arrow from $[t_1]$ to itself. Computing the transition arrows in this way we get the dual machine below.
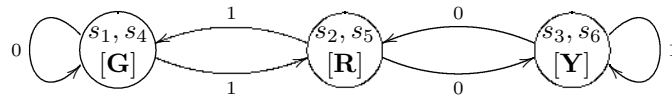
In this special case, the underlying transition graph of the double dual will be isomorphic to the transition graph of the dual but that is just an artefact of this example.

Brzozowski's Algorithm for Minimization. Brzozowski [10] discovered the following intriguing algorithm for minimizing finite state automata viewed as acceptors. Take the transitions and reverse the arrows. In addition interchange the accepting and non-accepting states; the resulting machine is not deterministic, of course. Determinize this machine in the usual way then reverse the result, flip the accepting states and the non accepting states again and determinize again. Remarkably, this gives the minimal deterministic automaton. The reverse operation is exactly our duality construction for the special case of one observation. Brzozowski does not present the duality construction using the logic that we have made explicit, but it is clearly there implicitly. Our presentation is essential for the generalization to the stochastic case. Of course this algorithm can blow up exponentially in the intermediate stage (the construction of the dual).

Despite this it appears to be useful in practice. We should point out that reachability plays a key role in Brzozowski's algorithm so the correspondence is not perfect. A recent paper [8] gives the precise categorical treatment of Brzozowski's algorithm.

4   A Simple Duality for Partially Observable Probabilistic Automata

In this section we work with partially observable probabilistic automata which is a prelude to the treatment of partially observable Markov decision processes (POMDPs). We will develop a duality with deterministic automata with stochastic observations (DASOs) using the probabilistic analogues of the simple formulas of the last two sections. In the AI literature these are called "e-tests" [24], where the e signifies that there is a single observation made at the *end* of a sequence of actions. With these e-tests one does get a pleasant duality theory, but the dual automaton loses much of the information of the original automaton. Nevertheless, this simple duality does capture many aspects of the original behaviour.

We recall the definition of a partially observable probabilistic automaton as $\mathcal{H} = (S, \Sigma, \mathcal{O}, \tau, \gamma)$, where $\tau : S \times \Sigma \times S \rightarrow [0,1]$ is the *transition function* and $\gamma : S \times \mathcal{O} \rightarrow [0,1]$ is the observation probability function. In our setup the observations are taken from a discrete set. It is not hard to develop the theory with observations taking on, for example, real values, but we will not do that in the present paper because that would involve us in measure-theoretic considerations.

We can define a dual using the same inductive definition for tests as in the deterministic case, but with a probabilistic semantics where a state satisfies a test with a given probability. Thus, the meaning of tests $[\![t]\!]$ should be considered as functions assigning probabilities to states. We will use s to stand for a typical state and ω for a typical observation. The definitions are:

$$[\![\omega]\!]_{\mathcal{H}}(s) = \gamma(s, \omega)$$
$$[\![(a) \cdot t]\!]_{\mathcal{H}}(s) = \sum_{s'} \tau(s, a, s')[\![t]\!]_{\mathcal{H}}(s')$$

We can define an equivalence relation on these tests by $t_1 \sim t_2$ if and only if $[\![t_1]\!] = [\![t_2]\!]$: thus, the equivalence class of t is completely determined by $[\![t]\!]$. We will just use $[\![t]\!]$ rather than the equivalence class $[t]$.

We define the dual $\mathcal{H}' = (S', \Sigma, \mathcal{O}', \tau', \gamma')$ as follows:
- $S' = \{[\![t]\!]_{\mathcal{H}} | t \in \mathcal{F}\}$, where $\mathcal{F}$ is the collection of formulas.
- $\mathcal{O}' = S$
- $\gamma'([\![t]\!]_{\mathcal{H}}, s) = [\![t]\!]_{\mathcal{H}}(s)$
- $\tau([\![t]\!]_{\mathcal{H}}, a, [\![(a) \cdot t]\!]_{\mathcal{H}}) = 1$ (0 otherwise)

Note that the transition function is now completely deterministic: it can be written in the much more perspicuous form $\tau([\![t]\!], a) = [\![(a) \cdot t]\!]$. Thus, the duality construction has made the system deterministic.

We can now consider the double dual, which has $\mathcal{H}'$-equivalence classes of formulas as states. Since the observations of $\mathcal{H}'$ are the states in $\mathcal{H}$, a basic test on $\mathcal{H}'$ would look like $[\![\hat{s}]\!]_{\mathcal{H}'}$ for some $s \in S$, where, as before, we are writing $\hat{s}$ for the state s regarded as an observation of the dual system. We find that

$$\gamma''([\![\hat{s}]\!]_{\mathcal{H}'}, [\![\omega]\!]_{\mathcal{H}}) = [\![\hat{s}]\!]_{\mathcal{H}'}([\![\omega]\!]_{\mathcal{H}}) = \gamma'([\![\omega]\!]_{\mathcal{H}}, \hat{s}) = [\![\omega]\!]_{\mathcal{H}}(s) = \gamma(s, \omega).$$

Tests interpreted on the double dual, applied to states of the double dual that are of the form $[\hat{s}]$ for some state s of the primal machine, define the same functions as they do on when interpreted on the original:

$$\begin{aligned}
[\![a_1 a_2 \cdots a_k \omega]\!]_{\mathcal{H}''}([\hat{s}]_{\mathcal{H}'}) &= [\![\omega]\!]_{\mathcal{H}''}(\delta''([\hat{s}]_{\mathcal{H}'}, a_1 a_2 \cdots a_k)) \\
&= [\![\omega]\!]_{\mathcal{H}''}([a_k \cdots a_2 a_1 \hat{s}]_{\mathcal{H}'}) \\
&= \gamma''([a_k \cdots a_2 a_1 \hat{s}]_{\mathcal{H}'}, [\![\omega]\!]_{\mathcal{H}}) \\
&= [\![a_k \cdots a_2 a_1 \hat{s}]\!]_{\mathcal{H}'}([\![\omega]\!]_{\mathcal{H}}) \\
&= [\![\hat{s}]\!]_{\mathcal{H}'}(\delta'([\![\omega]\!]_{\mathcal{H}}, a_k \cdots a_2 a_1)) \\
&= \gamma'([\![a_1 a_2 \cdots a_k \omega]\!]_{\mathcal{H}}, \hat{s}) \\
&= [\![a_1 a_2 \cdots a_k \omega]\!]_{\mathcal{H}}(s).
\end{aligned}$$

What we have here is a duality between probabilistic Kripke automata and deterministic automata with probabilistic observations. Once again the duality is mediated by the notion of satisfaction between states and tests and the entire duality theory can be seen as transposing the satisfaction relation.

The formulas that we have considered are very special: observations are made only at the end of a sequence of actions. One can consider tests to be formulas and ask what the effect of adding other logical connectives would be. We will, however, take the view that we are working with "tests" that the system may or may not pass. With this viewpoint it is more natural to consider generalizations that are different from what one would consider by adding more logical connectives to a logic.

As a prelude to the next section, consider what happens with a more general kind of test called an "s-test" [24] in the AI literature. The new feature is that one can make observations after every action. Note, however, that with these more general kinds of tests one does not induce the same functions on the original and double dual. This is because in the double dual the transitions are deterministic so the observations provide no additional information about the state, given the action. A more precise semantics would capture conditional probabilities of a given new state conditioned on the observations made.

5   State Based Duality for POMDPs

In order to obtain a duality theory without a loss of information one needs a more refined notion of experiments, or, equivalently, a richer notion of formulas. The class of s-tests introduced at the end of the last section is not quite the right

concept: it does capture all the system dynamics but, it is an ad-hoc class of tests without a nice algebraic structure. We will work with a larger class of tests in which a series of actions can be alternated with an observation. This allows the concatenation of actions with tests, which is essential for duality.

Definition 13. *We define a POMDP as*

$$\mathcal{M} = (S, \mathcal{A}, \mathcal{O}, \delta_{a \in \mathcal{A}} : S \times S \to [0,1], \gamma_{a \in \mathcal{A}} : S \times \mathcal{O} \to [0,1]).$$

We use the word "tests" almost as before ("e-tests"); we use the word "experiments" in this section for sequences of tests. The formal definitions are as follows.

Definition 14. *A **test** t is a non-empty sequence of actions followed by an observation, i.e. $t = a_1 \cdots a_n \omega$, with $n \geq 1$.*

Definition 15. *An **experiment** is a non-empty sequence of tests $e = t_1 \cdots t_m$ with $m \geq 1$.*

Note that these definitions force one to make an action in order to observe anything. This is a consequence of the way observations are defined; they are associated with actions so we cannot just make an observation in a state. Unlike with POPAs, observations are associated with the *action* and the target state so it makes no sense to regard a simple observation as a test.

In order to proceed with the construction of the dual to a POMDP we extend the definition of the transition function to work on sequences of actions.

Definition 16. *Given a POMDP as in Def. 13 we define a transition function δ_α, where α is a sequence of actions, inductively:*

$$\delta_\epsilon(s, x) = \mathbf{1}_{s=x} \qquad\qquad \forall s, x \in S$$
$$\delta_{a\alpha}(s, x) = \sum_{y \in S} \delta_a(s, y) \delta_\alpha(y, x) \qquad\qquad \forall s, x \in S.$$

We have written $\mathbf{1}_{s=x}$ for the indicator function viewed as a Kronecker distribution; i.e. $\mathbf{1}_{s=x}$ is 0 unless $s = x$ in which case it is 1.

In order to define the meaning of a state satisfying a test, or an experiment, we need to introduce a ternary symbol, because a test will contain at least one action and thus will cause a transition to a new state. We will define the satisfaction relation between states and tests as a ternary symbol $\langle s|t|q \rangle$ which gives the probability that the system starts in s, is subjected to the test t and ends up in the state q.

Definition 17. *We define $\langle s|t|q \rangle$ by induction on t: $\langle s|a\omega|q \rangle = \delta_a(s, q) \cdot \gamma_a(q, \omega)$ and $\langle s|\alpha a\omega|q \rangle = \sum_r \delta_\alpha(s, r)\langle r|a\omega|q \rangle$.*

We use the same notation for an experiment e: $\langle s|e|q \rangle$ is the probability of the system starting in state s, being subject to the experiment e and ending up in the state q.

It is worth clarifying exactly what it means to say that "a system is subjected to an experiment." If we have an experiment

$$e = a_1^{(1)} \dots a_{n_1}^{(1)} \omega_1 a_1^{(2)} \dots a_{n_2}^{(2)} \omega_2 \dots a_1^{(m)} \dots a_{n_m}^{(1)} \omega_m$$

then the system is subjected to the sequence of actions

$$a_1^{(1)} \dots a_{n_1}^{(1)} a_1^{(2)} \dots a_{n_2}^{(2)} \dots a_1^{(m)} \dots a_{n_m}^{(1)}.$$

The number $\langle s|e|x \rangle$ is the probability that we see the observations $\omega_1 \omega_2 \dots \omega_m$ at the appropriate points of the action sequence.

Definition 18. *Given an experiment e the probability $\langle s|e|q \rangle$ is given by the following inductive formula: for a basic experiment $e = t$ the formula is given by Def. 17, for an experiment of the form te we define*

$$\langle s|te|q \rangle = \sum_r \langle s|t|r \rangle \langle r|e|q \rangle.$$

These ternary relations are the fundamental quantities that one can use to carry out the duality constructions. One can define the notion of a state s satisfying test t or experiment e by just summing over the target states. We use the same angle-bracket notation for this.

Definition 19. *We define $\langle s|t \rangle$ to be the probability that a state s satisfies a test t. It is given by the following formula:*

$$\langle s|t \rangle = \sum_q \langle s|t|q \rangle.$$

Similarly $\langle s|e \rangle$ is the probability that s satisfies an experiment e:

$$\langle s|e \rangle = \sum_q \langle s|e|q \rangle.$$

Now we construct the dual and show how to come back. The dual is not a POMDP but a deterministic transition system with stochastic observations.

We proceed as usual by defining an equivalence, this time on experiments; exactly the same definition can be used on tests of course: tests are just simple experiments.

Definition 20. *For experiments e_1, e_2, we say $e_1 \sim_{\mathcal{M}} e_2$ if and only if $\langle s|e_1 \rangle = \langle s|e_2 \rangle$ for all $s \in S$. Then $[e]_{\mathcal{M}}$ is the $\sim_{\mathcal{M}}$-equivalence class of e.*

The construction of the dual proceeds as before, by making equivalence classes of experiments the states of the dual machine; the states of the primal machine become the observations of the dual machine.

Definition 21. *We define the dual as*

$$\mathcal{M}' = (S', \mathcal{A}, \mathcal{O}', \delta' : S' \times \mathcal{A} \to S', \gamma' : S' \times \mathcal{O}' \to [0,1])$$

where

$$S' = \{[e]_{\mathcal{M}}\}$$
$$\mathcal{O}' = S$$
$$\delta'([e]_{\mathcal{M}}, a_0) = [a_0 e]_{\mathcal{M}}$$
$$\gamma'([e]_{\mathcal{M}}, s) = \langle s|e\rangle$$

As noted before this is a deterministic transition system with stochastic observations.

To get the double-dual we have to use the appropriate construction in the space of the dual machines, i.e. in the space of deterministic transition systems with stochastic observations. This is precisely the simple construction of the last section. Thus, we consider only single-test experiments on the dual (i.e. e-tests), but we allow the action sequences to be empty. Then for a given test $\tau = \alpha s$ of the dual machine[3], where α is a sequence of actions, we have $\langle [t]_{\mathcal{M}}|\tau\rangle = \langle s|\alpha^R t\rangle$, where α^R indicates α with the action order reversed. Equivalence is defined analogously: $\tau_1 \sim_{\mathcal{M}'} \tau_2$ if and only if $\langle x|\alpha_1^R t\rangle = \langle y|\alpha_2^R t\rangle$ for all tests t; where $\tau_1 = \alpha_1 x$ and $\tau_2 = \alpha_2 y$.

We now define the double dual as follows.

Definition 22. *Given a POMDP $\mathcal{M}$ and its dual $\mathcal{M}'$ we construct the double dual $\mathcal{M}'' = (S'', \mathcal{A}', \mathcal{O}'', \delta'', \gamma'')$, which is of the same type as the dual and has the same actions, as follows:*

$$S'' = \{[\tau]_{\mathcal{M}'}\}$$
$$\mathcal{O}'' = S'$$
$$\delta''([\tau]_{\mathcal{M}'}, a_0) = [a_0\tau]_{\mathcal{M}}$$
$$\gamma''([\tau]_{\mathcal{M}'}, [e]_{\mathcal{M}}) = \langle [e]_{\mathcal{M}}|\tau\rangle = \langle s|\alpha^R e\rangle \qquad (\tau = \alpha s)$$

We now show that everything is well-defined for the transition functions, which follows more or less immediately from the definitions for observation functions. Note that the ternary symbol is not needed in the actual definition of the dual but it is necessary for the proof that the transition function is well-defined. Essentially the transitions in the dual are given by e goes under an a-action to ae; for this to make sense we need to show that it does not matter which representative of the equivalence class of e is chosen. This is what the next lemma shows.

Lemma 23. *If $e_1 \sim_{\mathcal{M}} e_2$ then $a_0 e_1 \sim_{\mathcal{M}} a_0 e_2$.*

---

[3] We will use the Greek letter τ for tests of the dual machine.

Proof. $e_1 \sim_{\mathcal{M}} e_2$, so $\langle s|e_1 \rangle = \langle s|e_2 \rangle$ for all $s \in S$. For $i = 1, 2$ let $e_i = t_1^{(i)} \cdots t_{m_i}^{(i)}$ and $t_1^{(i)} = \alpha^{(i)} \omega^{(i)} = a_1^{(i)} \cdots a_{n_i}^{(i)} \omega^{(i)}$. Then for any state s,

$$\langle s|a_0 e_1 \rangle = \sum_{x,y \in S} \langle s|a_0 t_1^{(1)}|y \rangle \langle y|t_2^{(1)} \cdots t_{m_1}^{(1)}|x \rangle$$

$$= \sum_{x,y \in S} \delta_{a_0 \alpha^{(1)}}(s, y) \gamma_{a_{n_1}^{(1)}}(y, \omega^{(1)}) \langle y|t_2^{(1)} \cdots t_{m_1}^{(1)}|x \rangle$$

$$= \sum_{x,y,z \in S} \delta_{a_0}(s, z) \delta_{\alpha^{(1)}}(z, y) \gamma_{a_{n_1}^{(1)}}(y, \omega^{(1)}) \langle y|t_2^{(1)} \cdots t_{m_1}^{(1)}|x \rangle$$

$$= \sum_{z \in S} \delta_{a_0}(s, z) \sum_{x,y \in S} \langle z|t_1^{(1)}|y \rangle \langle y|t_2^{(1)} \cdots t_{m_1}^{(1)}|x \rangle$$

$$= \sum_{z \in S} \delta_{a_0}(s, z) \langle z|e_1 \rangle = \sum_{z \in S} \delta_{a_0}(s, z) \langle z|e_2 \rangle = \langle s|a_0 e_2 \rangle.$$

∎

Similarly, for the double dual we have the following lemma.

Lemma 24. *If $\tau_1 \sim_{\mathcal{M}'} \tau_2$ then $a_0 \tau_1 \sim_{\mathcal{M}'} a_0 \tau_2$ for any action a_0.*

Proof. Let $\tau_1 = \alpha_1 x$ and $\tau_2 = \alpha_2 y$ and assume that $\tau_1 \sim_{\mathcal{M}'} \tau_2$, so $\langle x|\alpha_1^R e \rangle = \langle y|\alpha_2^R e \rangle$ for all experiments e. Then for any experiment e,

$$\langle x|(a_0 \alpha_1)^R e \rangle = \langle x|\alpha_1^R a_0 e \rangle = \langle x|\alpha_1^R (a_0 e) \rangle = \langle y|\alpha_2^R (a_0 e) \rangle.$$

∎

Now that we know that these constructions are well defined the duality is captured by the following theorem.

Theorem 25. *The probability of a state s in the primal satisfying an experiment e, i.e. $\langle s|e \rangle$, is given by $\langle [s]_{\mathcal{M}'}|[e]_{\mathcal{M}} \rangle = \gamma''([s]_{\mathcal{M}'}|[e]_{\mathcal{M}})$, where $[s]_{\mathcal{M}'}$ indicates the equivalence class of the e-test on the dual which has s as an observation and an empty sequence of actions.*

Proof. Note that $[s]_{\mathcal{M}'}$ is an equivalence class of states of tests of the dual, hence a state of the double dual. Recall that the dual is a DASO so we are using the simple duality here. Note further that $[e]_{\mathcal{M}}$ is an equivalence class of experiments on the primal, which is a state of the dual and hence an observation of the double dual. So, by definition of the angle bracket notation this is just $\gamma''([s]_{\mathcal{M}'}|[e]_{\mathcal{M}})$. By the definition of the double dual construction we have

$$\gamma''([\tau]_{\mathcal{M}'}|[e]_{\mathcal{M}}) = \langle s|\alpha^R e \rangle,$$

where $\tau = \alpha s$. In our case α is the empty sequence and τ is just s so we get $\gamma''([s]_{\mathcal{M}'}|[e]_{\mathcal{M}}) = \langle s|e \rangle$. ∎

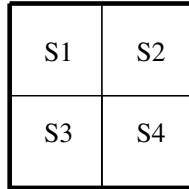| S1	S2
S3	S4

Fig. 1. The navigation domain

Thus, the results of all experiments on the primal can be read off the double dual.

We present two examples illustrating the constructions of this section. The first is taken from [17]. It describes a simple navigation domain as shown in Fig. 1.

The squares represent places where a robot could be. The heavy lines represent walls that cannot be crossed; the walls are painted blue. The robot can take the following actions: N,S,E and W. If it is already at the left end, say in square S_1, and attempts to move west (left) it will just stay where it is. It can also make a observation of the colour of its immediate surroundings. If it is in either of the squares S_1 or S_3 it will see blue with probability 1; in the squares S_2 and S_4 it will see red *or* blue each with probability 0.5; the red reading represents a curtain that it could see on the right corresponding to the lightly marked lines on the extreme right edge of the picture. This is an example with deterministic moves but noisy readings. The observations are state based in this case, but we will present it in the POMDP framework.

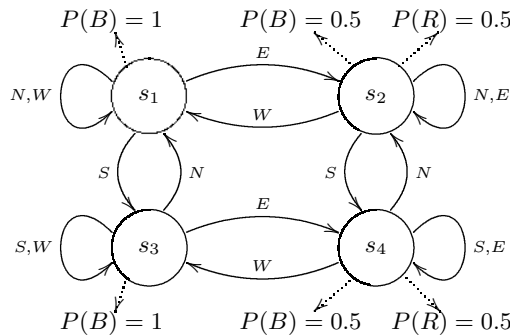The system may be represented by the automaton shown in Fig. 2

Fig. 2. The navigation domain as a transition system

We can calculate a part of the dual as shown in Fig. 3. The whole dual is of course infinite and one cannot write it down explicitly. Here the states are
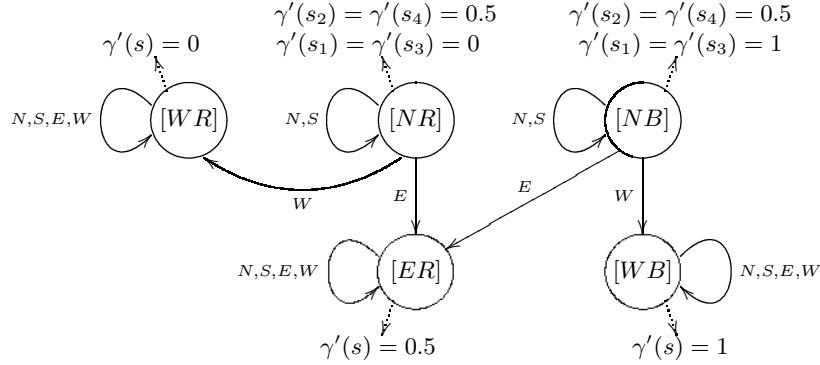
$$\gamma'(s_2) = \gamma'(s_4) = 0.5 \qquad \gamma'(s_2) = \gamma'(s_4) = 0.5$$
$$\gamma'(s) = 0 \qquad \gamma'(s_1) = \gamma'(s_3) = 0 \qquad \gamma'(s_1) = \gamma'(s_3) = 1$$

$$\gamma'(s) = 0.5 \qquad \gamma'(s) = 1$$

Fig. 3. Part of the dual of the navigation domain

labelled by (equivalence classes of) experiments on the original system, but the transitions are labelled by the same actions as in the primal. The old states are now the observations of the dual. We write $\gamma(s) = 0$ as short for $\gamma(s_i) = 0$ for $i = 1, 2, 3, 4$.

The double dual, shown in Fig. 4, will of course collapse it to a minimal representation. Since the navigation domain is up-down symmetric the collapsed version will just have two states as shown below. This version will completely predict all possible experimental outcomes of the original system.
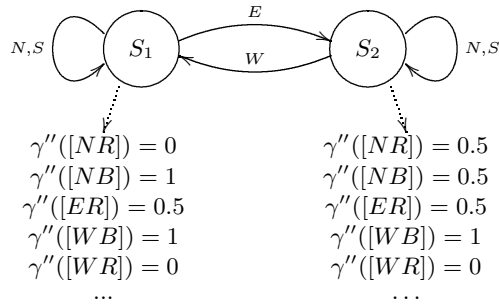

$$\gamma''([NR]) = 0 \qquad \gamma''([NR]) = 0.5$$
$$\gamma''([NB]) = 1 \qquad \gamma''([NB]) = 0.5$$
$$\gamma''([ER]) = 0.5 \qquad \gamma''([ER]) = 0.5$$
$$\gamma''([WB]) = 1 \qquad \gamma''([WB]) = 1$$
$$\gamma''([WR]) = 0 \qquad \gamma''([WR]) = 0$$
$$\cdots \qquad \cdots$$

Fig. 4. The double dual of the navigation domain

Our second example shows a situation where the observations depend on the action as well as the target state. This can also be viewed as a four-square navigation domain. There are two observations o_1 and o_2: exactly one of these observations will be made as a transition is taken so we need only specify the probability of one of them. The actions are N, S, E, W which stand for north, south, east and west respectively. This is also a very symmetric domain with top down symmetry.
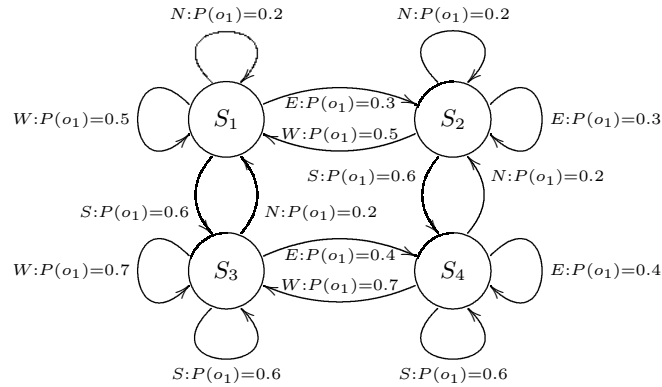
The primal is shown in Fig. 5.

Fig. 5. A POMDP with action dependent observations

The dual is infinite, of course; in Fig. 6 we show a fragment of the dual. We have shown all the states corresponding to tests, i.e. experiments with a single observation and we have suppressed the observations of the dual.

In the table shown in Fig. 7 we display some of the values of γ' for two-observation experiments and all of the values for tests. The first line says: given that that the dual system is in state $[NO_1]$ the observation function for all the s_i is the same and has the value 0.2. Note that, in fact, that for any sequence of N actions followed by O_1 these numbers are the same so all tests of the form $N^k O_1$ are equivalent.

The double dual, shown in Fig. 8 has two states.

6   Related Work

Rivest and Schapire [33] present an approach to inferring the structure of a finite-state automaton from its input-output behavior, by running "experiments" on the automaton. They rely on an "update graph", which is essentially the dual in our representation, and on e-tests, of the form $a_1 \ldots a_n o$. They also present experiments in which they infer an automaton based on Rubik's cube using this structure. Their work is limited to deterministic automata. Nevertheless, the fact that they could deal with a system with 10^{19} states is very impressive. Their work shows that a very large system can have a much more compact dual.

More recently, predictive state representations [24, 34] (PSR), introduced in the AI community, generalized the work of Rivest and Schapire to the case of stochastic automata. The representation is based on the prediction of s-tests, which are of the form $a_1 o_1 \ldots a_n o_n$. In the work of Littman et al. [24], each state in a POMDP is viewed as represented by an infinite set of predictions, for
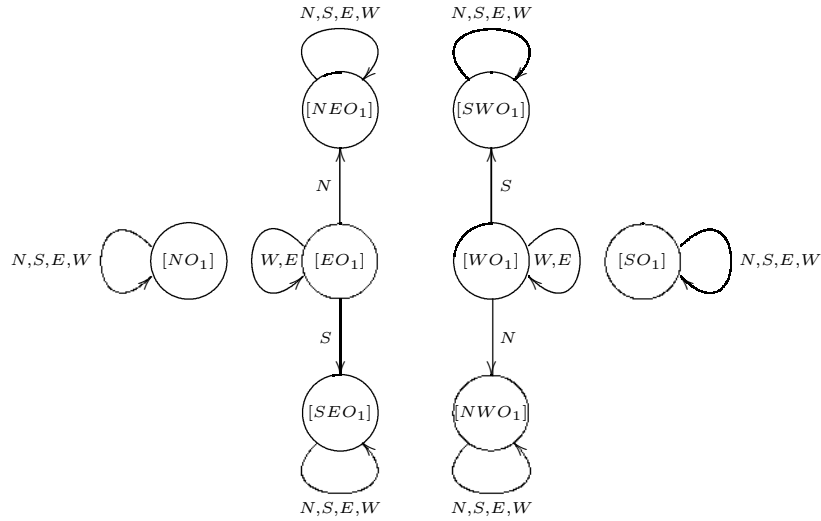
Fig. 6. Part of the dual of the POMDP in Fig. 5

State $[t]$	$\gamma'([t], s_1)$	$\gamma'([t], s_2)$	$\gamma'([t], s_3)$	$\gamma'([t], s_4)$
$[NO_1]$	0.2	0.2	0.2	0.2
$[SO_1]$	0.6	0.6	0.6	0.6
$[EO_1]$	0.3	0.3	0.4	0.4
$[WO_1]$	0.5	0.5	0.7	0.7
$[NEO_1]$	0.3	0.3	0.3	0.3
$[SWO_1]$	0.7	0.7	0.7	0.7
$[SEO_1]$	0.4	0.4	0.4	0.4
$[NWO_1]$	0.5	0.5	0.5	0.5
$[NO_1NO_1]$	0.04	0.04	0.04	0.04
$[SO_1SO_1]$	0.36	0.36	0.36	0.36
$[SO_1NO_1]$	0.12	0.12	0.12	0.12
$[NO_1EO_1]$	0.06	0.06	0.06	0.06
$[NO_1WO_1]$	0.1	0.1	0.1	0.1
$[EO_1EO_1]$	0.09	0.09	0.16	0.16
$[WO_1WO_1]$	0.25	0.25	0.49	0.49
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$

Fig. 7. Table of observation probabilities in the dual

$$\gamma''([NO_1]) = 0.2 \qquad\qquad \gamma''([NO_1]) = 0.2$$
$$\gamma''([SO_1]) = 0.6 \qquad\qquad \gamma''([SO_1]) = 0.6$$
$$\gamma''([EO_1]) = 0.3 \qquad\qquad \gamma''([EO_1]) = 0.4$$
$$\gamma''([WO_1]) = 0.5 \qquad\qquad \gamma''([WO_1]) = 0.7$$
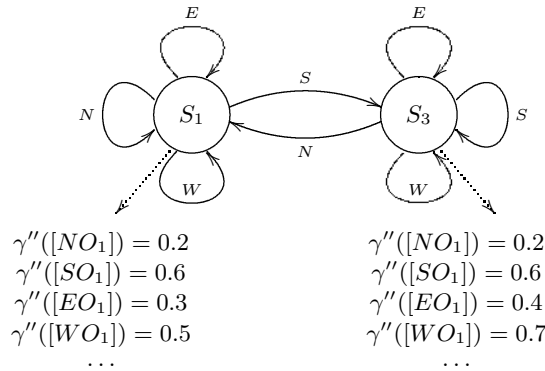$$\cdots \qquad\qquad\qquad \cdots$$

Fig. 8. The double dual

each possible test. However, there exists a finite number of linearly independent tests whose predictions are sufficient to compute the prediction for any other tests. Moreover, the number of such "core tests" is at most equal to the number of states in the POMDP. Singh et al. [34] then generalized this approach by considering predictions of s-tests based on histories. They also showed how different frameworks can be represented with a finite number of "core tests." From the point of view presented in this paper, we can view the PSRs as a way of representing the dual or the double dual of a system using the set of linearly independent columns. This view has the advantage of a finite representation, e.g. if the original system is a POMDP. However, this representation does not lend itself easily to approximations. It is our hope that by working directly with the dual or the double-dual, one can develop more easily a theory of approximation for such system.

It seems likely that the best setting to understand the categorical context is that of Chu spaces [5], at least for the deterministic automaton case. Our quotient construction on automata is closely related to the process of forming separated extensional Chu spaces [6]. There is also some similarity to the work of Pratt [31] though there the duality is between states and trajectories (which he calls "schedules").

There are several discussions in the literature on duality in systems theory; see, for example, the excellent paper by Bainbridge [4] and the several very interesting chapters in the book by Kalman, Farb and Arbib [21]. In systems theory one is concerned with *controlling* a system to obtain a desired behaviour. As with our POMDPs, the systems are partially observable; one does not see the state. One only has a *readout map* that maps the states into observables. The "fundamental duality" in this subject is the duality between controllability: the ability to steer a system into a known state, and observability: the ability to determine the state after a series of observations. There has been a rich categorical treatment of this subject: for example, there are several papers by Arbib and Manes on this topic [1–3]. These are largely concerned with the nonprobabilistic situation;

the papers by Arbib and Manes hint at the probabilistic case but we have never seen this spelled out satisfactorily.

There is a whole plethora of dualities in mathematics. The Stone-type dualities establish a duality between logics and transition systems. This has appeared in denotational semantics and is due to Plotkin and Smyth [36]. They establish a relationship between "forward" state-transformer semantics and "backwards" predicate-transformer semantics. Kozen [22] established similar results in the case of probabilistic programs and probabilistic dynamic logic. More recently, there has been very interesting work by Mislove, Ouaknine, Pavlovic and Worrell [28] and by Pavlovic, Mislove and Worrell [13] on duality for labelled Markov processes (LMPs). These are like POMDPs but they do not have the notion of observation. Other research on duality for logics and for transitions systems include the work by Bidoit, Hennicker and Kurz [7] and Bonsangue and Kurz [9].

7   Conclusions

We have shown that, in some informal sense, there is a duality between state and observation, or, more precisely, between state and experiment. In this view a state is an equivalence class of experimental data. It frees us from having to work with arbitrary preconceived notions of state. This view, we hope, will unify many ideas that are currently being investigated for representing systems with hidden state.

It is important to clarify what we mean by minimal. Of course the "minimal" gadget that we construct is larger than the POMDP that one starts with; the latter is finite and the former is infinite. But, among the class of deterministic systems with stochastic observations that represent the same behaviour it is minimal in the following sense. Minimal means that it is the "most highly quotiented version possible." More precisely, suppose that we have a system S of some kind. One can "reduce its size" by defining an appropriate equivalence relation $\sim$ and constructing $S/\sim$. We can say that a system S' is a *minimal realization* of S if it is behaviourally equivalent and if it is the quotient of S by some equivalence relation *and if there is any other system X* which is also a quotient of S then X can be *further quotiented* to yield S'. In the companion categorical paper this is formalized as a couniversal property and comes out naturally. For finite-state systems this is indeed the same as having as few states as possible.

There is much to be investigated from the algorithmic point of view. Perhaps the most pressing issue is a pleasant approximation theory. The dual and the double-dual are both based on exact equivalences. This raises the possibility of working with metric notions [15] and constructing more compact representations based on identifying "nearby" experiments.

Finally, everything has been worked out here for discrete systems. Clearly, for realistic applications one would need to extend the theory to continuous state spaces, and, even more importantly, to continuous observations.

References

1. Arbib, M.A., Manes, E.G.: Machines in a category: An expository introduction. SIAM Review 16, 163–192 (1974)
2. Arbib, M.A., Manes, E.G.: Foundations of system theory: decomposable systems. Automatica 10, 285–302 (1974)
3. Arbib, M.A., Manes, E.G.: Adjoint machines, state behavior machines and duality. J. Pure Appl. Algebra 6, 313–343 (1975)
4. Bainbridge, E.S.: The fundamental duality of system theory. In: Hartnett, W.E. (ed.) Systems: Approaches, Theories, Applications, pp. 45–61. Reidel (1977)
5. Barr, M.: The chu construction. Theory Appl. Categories 2, 17–35 (1996)
6. Barr, M.: The separated extensional Chu category. Theory Appl. Categories 4(6), 137–147 (1998)
7. Bidoit, M., Hennicker, R., Kurz, A.: Observational logic, constructor-based logic, and their duality. Theor. Comput. Sci. 3(298), 471–510 (2003)
8. Bonchi, F., Bonsangue, M.M., Rutten, J.J.M.M., Silva, A.: Brzozowski's Algorithm (Co)Algebraically. In: Constable, R.L., Silva, A. (eds.) Logic and Program Semantics, Kozen Festschrift. LNCS, vol. 7230, pp. 12–23. Springer, Heidelberg (2012)
9. Bonsangue, M.M., Kurz, A.: Duality for Logics of Transition Systems. In: Sassone, V. (ed.) FOSSACS 2005. LNCS, vol. 3441, pp. 455–469. Springer, Heidelberg (2005)
10. Brzozowski, J.A.: Canonical regular expressions and minimal state graphs for definite events. In: Fox, J. (ed.) Proceedings of the Symposium on Mathematical Theory of Automata. MRI Symposia Series, vol. 12, pp. 529–561. Polytechnic Press of the Polytechnic Institute of Brooklyn (April 1962), book appeared in 1963
11. Chrisman, L.: Reinforcement learning with perceptual aliasing: The perceptual distinctions approach. In: Proceedings of the Tenth National Conference on Artificial Intelligence, pp. 183–188 (1992)
12. Clarke, E.M., Grumberg, O., Peled, D.A.: Model Checking. The MIT Press, Cambridge (1999)
13. Pavlovic, D., Mislove, M., Worrell, J.B.: Testing Semantics: Connecting Processes and Process Logics. In: Johnson, M., Vene, V. (eds.) AMAST 2006. LNCS, vol. 4019, pp. 308–322. Springer, Heidelberg (2006)
14. Desharnais, J., Edalat, A., Panangaden, P.: Bisimulation for labeled Markov processes. Information and Computation 179(2), 163–193 (2002)
15. Desharnais, J., Gupta, V., Jagadeesan, R., Panangaden, P.: A metric for labelled Markov processes. Theoretical Computer Science 318(3), 323–354 (2004)

16. Hopcroft, J.E., Motwani, R., Ullman, J.D.: Introduction to automata theory, languages, and computation, 2nd edn. Addison-Wesley Series in Computer Science. Addison-Wesley-Longman (2001)

17. Hundt, C., Panangaden, P., Pineau, J., Precup, D.: Representing systems with hidden state. In: The Twenty-First National Conference on Artificial Intelligence, AAAI (2006)

18. James, M., Singh, S.: Learning and discovery of predictive state representations in dynamical systems with reset. In: International Conference on Machine Learning, vol. 21, pp. 417–424 (2004)

19. Johnstone, P.: Stone Spaces, Cambridge Studies in Advanced Mathematics, vol. 3. Cambridge University Press (1982)

20. Kaelbling, L.P., Littman, M.L., Cassandra, A.R.: Planning and acting in partially observable stochastic domains. Artificial Intelligence 101 (1998)

21. Kalman, R.E., Falb, P.L., Arbib, M.A.: Topics in Mathematical Systems Theory. McGraw Hill (1969)

22. Kozen, D.: A probabilistic PDL. Journal of Computer and Systems Sciences 30(2), 162–178 (1985)

23. Kozen, D.: Automata and computability. Undergraduate texts in computer science. Springer (1997)

24. Littman, M., Sutton, R., Singh, S.: Predictive representations of state. In: Advances in Neural Information Processing Systems 14 (NIPS), vol. 14, pp. 1555–1561 (2002)

25. McCallum, A.: Reinforcement Learning with Selective Perception and Hidden State. Ph.D. thesis, University of Rochester (1995)

26. Milner, R.: A Calculus of Communication Systems. LNCS, vol. 92. Springer, Heidelberg (1980)

27. Milner, R.: Communication and Concurrency. Prentice-Hall (1989)

28. Mislove, M., Ouaknine, J., Pavlovic, D., Worrell, J.: Duality for Labelled Markov Processes. In: Walukiewicz, I. (ed.) FOSSACS 2004. LNCS, vol. 2987, pp. 393–407. Springer, Heidelberg (2004)

29. Park, D.: Concurrency and Automata on Infinite Sequences. In: Deussen, P. (ed.) GI-TCS 1981. LNCS, vol. 104, pp. 167–183. Springer, Heidelberg (1981)

30. Park, D.: Title unknown (1981), slides for Bad Honnef Workshop on Semantics of Concurrency

31. Pratt, V.R.: The Duality of Time and Information. In: Cleaveland, W.R. (ed.) CONCUR 1992. LNCS, vol. 630, pp. 237–253. Springer, Heidelberg (1992)

32. Puterman, M.L.: Markov Decision Processes: Discrete Stochastic Dynamic Programming. Wiley (1994)

33. Rivest, R.L., Schapire, R.E.: Diversity-based inference of finite automata. Journal of the ACM 41(3), 555–589 (1994)

34. Singh, S., James, M.R., Rudary, M.R.: Predictive state representations: A new theory for modeling dynamical systems. In: Proceedings of the Twentieth Conference on Uncertainty in Artificial Intelligence (UAI), pp. 512–519 (2004)

35. Sipser, M.: Introduction to the theory of computation. PWS Publishing Company (1997)

36. Smyth, M.: Powerdomains and Predicate Transformers. In: Díaz, J. (ed.) ICALP 1983. LNCS, vol. 154, pp. 662–676. Springer, Heidelberg (1983)

37. Sutton, R.S., Tanner, B.: Temporal-difference networks. In: Advances in Neural Information Processing Systems, vol. 17 (2005)