Abstraction in Predictive State Representations

Vishal Soni and Satinder Singh
Computer Science and Engineering
University of Michigan, Ann Arbor
{soniv,baveja}@umich.edu

Abstract

Most work on Predictive Representations of State (PSRs) focuses on learning a complete model of the system that can be used to answer *any* question about the future. However, we may be interested only in answering certain kinds of abstract questions. For instance, we may only care about the presence of objects in an image rather than pixel level details. In such cases, we may be able to learn substantially smaller models that answer only such abstract questions. We present the framework of PSR homomorphisms for model abstraction in PSRs. A homomorphism transforms a given PSR into a smaller PSR that provides exact answers to abstract questions in the original PSR. As we shall show, this transformation captures structural and temporal abstractions in the original PSR.

Introduction

Predictive representations of state replace the traditional notion of latent state with a state composed of a set of predictions of the future. These predictions answer *questions* of the form, "What is the probability of seeing an observation sequence $o_1 o_2 \cdots o_n$ if actions $a_1 a_2 \cdots a_n$ are executed from the current history?". Littman, Sutton, & Singh (2001) showed that there can exist a finite set of questions whose predictions perfectly capture state and can be used to make *any* prediction about the system.

Learning a model to answer every possible question about the future of a system can be an arduous if not impossible task. We, as humans, certainly do not have such a detailed and complete model of the real world. Rather, we usually seek to abstract out details and answer a limited set of abstract questions about the world. For instance, we rarely try to answer questions about every pixel in a future image, often being satisfied with answers to abstract questions such as the presence or absence of objects of interest (say a door or a staircase) in the image. In many cases, our questions even generalize over time; we may want to know if we would encounter a coffee shop in the next half hour with the precise time of the encounter being unimportant. Given that we are interested in answering such structurally and temporally abstract questions, is there an appropriately reduced, and thereby simpler to learn, PSR model of the world?

There is more than one way to arrive at such a reduced model. One approach is to build an approximate model that provides approximate answers to every possible detailed question about the world. Another approach is to build a model that provides exact answers to a limited number of abstract questions of interest. This work is motivated by the latter approach. Note that determining the abstract questions of interest is itself an important research question that we ignore in this paper. We focus instead on defining models whose complexity scales with the complexity of the given abstract questions that the agent needs to answer.

As a simple motivating example, consider the deterministic world in Figure 1a in which an agent can move north, south, east, and west, and observes the color of the square it is on. The start state of the agent is randomly picked. To answer every question about this world, one would have to learn the model depicted in Figure 1a. However, assume that the questions of interest are of the form, "What is the probability of seeing a shaded square (as opposed to white) if I move in some direction?". In such a case, the reduced model of Figure 1b answers such questions exactly. In this domain the agent can move left, right, or stay in place. The states in Figure 1b denoted 'C' lead to the agent observing a shaded square. Intuitively, Figure 1b collapses the columns of Figure 1a into a single state. Figure 1c is a further reduction that exploits action symmetries. The model in Figure 1d can be used to answer the question, "What is the probability that I will see black within the next two time steps if I behave randomly?". This question is semi-oblivious to time; it doesn't care whether black is observed in 1 time-step, or in 2 time-steps.

This paper develops the mathematical formalism based on the notion of homomorphisms needed to transform detailed PSR models into reduced abstract PSR models that can answer certain kinds of abstract questions in the original model exactly (as in the above example). We do this in two steps, first addressing structural abstractions and then temporal abstractions.

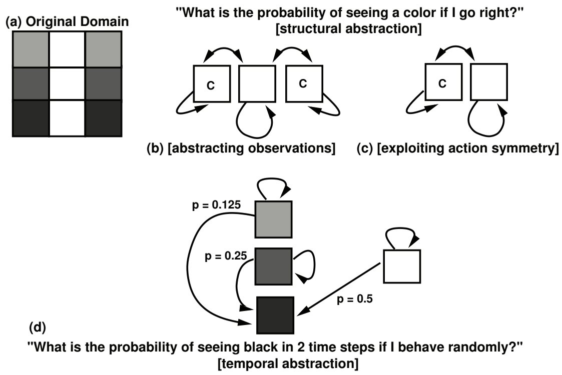

Figure 1: Illustration of abstraction.

Related Work: Model Minimization in Dynamical Systems

Building reduced models in PSRs is related to model minimization methods studied in the context of dynamical systems. Dean & Givan (1997) have explored model minimization in the context of MDPs and base their definition of state equivalence on the concept of homogeneous partition of the state space. Ravindran & Barto (2002) also explore model minimization in their work on MDP homomorphisms, which are based on the notion of homomorphisms of finite state machines. Ravindran & Barto (2003) extend MDP homomorphisms to SMDPs. Van der Schaft (2004) explores model minimization in continuous linear dynamical systems by extending bisimulations for concurrent processes. Our framework of PSR homomorphisms extends the above ideas to model minimization in PSRs.

Predictive State Representations

Discrete controlled dynamical systems are characterized by a finite set of actions, A, and observations, O. At every time step i, the agent takes an action $a^i \in A$ and receives some observation $o^i \in O$. A *history* is a sequence of actions and observations $a^1 o^1 \cdots a^n o^n$ through time-step n. Similarly, a *test* is a possible sequence of future actions and observations $a_1 o_1 \cdots a_n o_n$ (note the use of subscripts and superscripts). A test is said to succeed if the observations of the test are received given that the actions are taken. A *prediction* of any n-length test t from history h is the probability that the test will succeed from that history h and is denoted $P(t|h) = Pr(o_1 \cdots o_n | h a_1 \cdots a_n)$.

A state ψ in a PSR is represented as a vector of predictions of a set of tests $\{q_1 \cdots q_k\}$ called *core tests* and is thus expressed completely in terms of grounded terms, i.e. actions and observations. This prediction vector $\psi_h = \{Pr(q_1|h), \cdots, Pr(q_k|h)\}$ is a sufficient statistic for history in that it can be used to make predictions of any test. As Littman et. al. (2001) show, for every test t there exists a $1 \times k$ projection vector m_t such that $P(t|h) = \psi_h \cdot m_t$ for all histories h, where importantly m_t is independent of h. Hereafter, we drop the history-subscript on ψ if the dependence on history is unambiguous. Thus if we know the predictions of core tests, then we know everything there is to know about the history of the system. In the rest of this paper, we use the symbol Ψ to denote the entire predictive state space of a PSR.

A linear PSR is thus composed of the set of core test predictions maintained from the null history (the 'start' state) and the update parameters (the m vectors) used to update the state as the agent takes new actions. Given that an action a is taken in history h and observation o is obtained, the prediction of each core test q is updated as:

$$Pr(q|hao) = \frac{Pr(aoq|h)}{Pr(ao|h)} = \frac{\psi_h \cdot m_{aoq}}{\psi_h \cdot m_{ao}}.$$

Structural Abstraction in PSRs

Going back to the example in Figure 1, we sought to use the reduced domain in Figure 1c to answer abstract questions about the future in the original domain. These questions are abstract in the sense that they aggregate several futures (i.e. tests). So a question of the form, "What is the probability that I will see a shaded square if I take a particular action?" is an aggregate of three tests, i.e. the probability of seeing *light-gray*, *dark-gray*, or *black* if that action is taken. We refer to these aggregates of tests as *block-tests*. A PSR homomorphism is essentially a function that groups tests into such blocks and maps each block-test to a unique test in the reduced domain. The reduced domain can then be used to make predictions of block-tests in the original domain. So, in Figure 1, the above block-test about seeing a shaded square maps to the test, "What is the probability that I will observe 'C' if I take a particular action?" in the domain shown in Figure 1c.

The fundamental intuition leading to homomorphisms is that if we are only going to use the PSR model to answer questions about aggregate or block futures, perhaps we can aggregate or block pasts/histories (equivalently states). However, the dynamics of the system would restrict the kind of aggregations that are permissible, and it is not clear what these constraints are. In the next section, we will define PSR homomorphisms as functions that partition the PSR predictive state space Ψ into blocks of states, and map each such block to a unique state in the reduced model. We then examine how this leads to an aggregation of tests into block-tests and how such block-tests abstract out details of the domain.

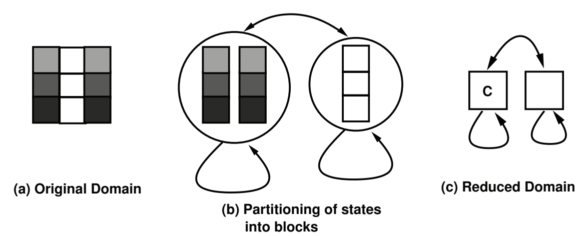

Figure 2: We can define a homomorphism that maps the original domain (a) to a reduced domain (c). In (b) we see the partitioning of states caused by this mapping.

PSR homomorphisms

Throughout this exposition, we will refer to homomorphisms as relating two domains - an original domain Σ and its reduced image Σ'. We use Ψ to denote the predictive state space in Σ and use Ψ' to denote the same in Σ'. Similarly, A and O denote the action and observation sets in Σ, and A' and O' denote the same quantities in Σ'.

A *partitioning* of Σ's state space Ψ is a collection of disjoint subsets, or *blocks*, $B_1, \cdots, B_m$ of states such that each $B_i \subseteq \Psi$ and $\bigcup_i B_i = \Psi$. We can define the block transition probability as:

$$Pr(B_i|\psi, a) = \sum_{\hat{\psi} \in B_i} Pr(\hat{\psi}|\psi, a). \quad (1)$$

That is, when taking action a in state ψ, $Pr(B_i|\psi, a)$ is the probability that the next state is in block B_i.

Building on the concept of homomorphisms for MDPs, we now define PSR homomorphisms. Informally, a homomorphism is a tuple comprised of two surjections (many to one functions). The first is a surjection $f : \Psi \rightarrow \Psi'$ that maps predictive state from Σ to Σ'. We will use $f^{-1}(\psi')$ to refer to the block of states in Σ that map to ψ' in Σ'. The other is a surjection $v_\psi : A \rightarrow A'$ that is mapping between actions conditioned on state, i.e. specifies a separate action mapping for every state. A homomorphism partitions Ψ into blocks such that each block maps to unique state in Ψ'.

Definition 1. *A PSR homomorphism from a PSR Σ to PSR Σ' is defined by a tuple of surjections $\langle f, v_\psi(a)\rangle$ where $f : \Psi \rightarrow \Psi'$ and $v_\psi : A \rightarrow A'$ for all prediction vectors $\psi \in \Psi$ such that:*

$$Pr(\psi'|f(\psi), v_\psi(a)) = Pr(f^{-1}(\psi')|\psi, a) \quad (2)$$

for all $\psi' \in \Psi', \psi \in \Psi, a \in A$.

In words, Equation 2 requires that the transition probability from $f(\psi)$ to ψ' in Σ' equal the probability of transition from ψ to the block $f^{-1}(\psi')$ in Σ as defined in Equation 1.

We now show that a PSR homomorphism can be naturally thought of and expressed in grounded terms, i.e. in terms of actions and observations. Let $u_{h,a} : O \rightarrow O'$ be a surjection from observations in O to O' in history $\{h, a\}$, and let $v_h : A \rightarrow A'$ be a surjection from actions in A to A' in history h. Note that $u_{h,a}$ and v_h are conditioned on history and define a separate mapping for every history. Let $\Theta = \langle u, v\rangle$ so that in any history h, $\Theta_h(a, o) = \{v_h(a), u_{h,a}(o)\}$. Implicitly, Θ also denotes a mapping between histories in the two domains. Consider a history $h = \{a^1o^1, \cdots, a^no^n\}$. The corresponding history in Σ' is $\Theta_\emptyset(h) = \{\Theta_\emptyset(a^1o^1), \cdots, \Theta_{h^{n-1}}(a^no^n)\}$, where $h^{n-1} = a^1o^1 \cdots a^{n-1}o^{n-1}$. Since predictive state is a sufficient statistic for history, Θ also implies a surjection (f) from Ψ to Ψ' and a surjection (v_ψ) from A to A'.

The following theorem shows that a homomorphism from Σ to Σ' can be expressed in terms of surjections on actions and observations.

Theorem 1. *Consider PSRs Σ and Σ'. Define the $\Theta = \langle u, v\rangle$ as above. The function Θ is a homomorphism between Σ and Σ' if and only if for any history h in Σ, any action a in A and any observation $o' \in O'$:*

$$Pr(o'|\Theta_\emptyset(h), v_h(a)) = \sum_{o \in u^{-1}_{h,a}(o')} Pr(o|h, a). \quad (3)$$

Proof. We first prove the forward direction. Let h_1 be a history in Σ and let $h'_1 = \Theta_\emptyset(h_1)$ be the corresponding history in Σ'. Consider an action $a \in A$ and action $a' = v_h(a)$. The probability that we take action a' and arrive in history $h'_2 = h'_1a'o'$ is

$$\begin{aligned} Pr(h'_2|h'_1, a') &= Pr(h'_1a'o'|h'_1a') \\ &= Pr(o'|h'_1a') \\ &= \sum_{o \in u^{-1}_{h_1,a}(o')} Pr(o|h_1a) \\ &= \sum_{o \in u^{-1}_{h_1,a}(o')} Pr(h_1ao|h_1a) \\ &= \sum_{h_2 \in \Theta^{-1}_\emptyset(h'_2)} Pr(h_2|h_1a). \end{aligned}$$

Since predictive state is a sufficient statistic for history, the above result implies

$$Pr(\psi'|h'_1, a') = \sum_{\psi \in \Theta^{-1}_\emptyset(\psi')} Pr(\psi|h_1a)$$

which satisfies the condition of Definition 1.

Since all of the above equations are equalities, the reverse direction automatically holds. Thus, Θ is a homomorphism from Σ to Σ' expressed in grounded terms. □

Consider Figure 2 (a redrawing of Figure 1); we now have the machinery to define a homomorphism that maps the example domain in Figure 2a to the domain in Figure 2c. The mapping between observations is straightforward: In every history, map all shades in Figure 2a to the observation 'C' in Figure 2c, and map the white color in 2a to the white color in Figure 2c. The mapping between actions is slightly more complicated (it exploits action symmetries):

- In all histories that end in the first column (from the left) in Figure 2a, map *Move-East* to *Move-Right* in Figure 2c. Map all the other actions to *stay in place*.
- In all histories that end in the middle column in Figure 2a, map both *Move-West* and *Move-East* to *Move-Left* in Figure 2c. Map all the other actions to *stay in place*
- In all histories that end in the third column in Figure 2a, map *Move-West* to *Move-Right* in 2c. Map all the other actions to *stay in place*

It is easy to see that this homomorphism effectively partitions the states in Figure 2a into the blocks in Figure 2b.

We now examine how a PSR homomorphism (that maps histories) leads to a surjection between tests, and how the image PSR Σ' can be used to answer questions about block-tests in the original PSR Σ.

Answering Abstract (Block) Questions

As is the case for histories, a PSR homomorphism Θ maps tests in Σ to tests in Σ' in the following manner. Let $t = a_1o_1, \cdots, a_no_n$ be a test in Σ. In any given history h, the corresponding test in Σ' is

$\Theta_h(t) = \{\Theta_h(a_1 o_1), \cdots, \Theta_{ht_{n-1}}(a_n o_n)\}$, where $t_{n-1} = a_1 o_1 \cdots a_{n-1} o_{n-1}$. Note that this mapping between tests is history dependent. Since Θ is a surjection, the mapping between tests will also be many-to-one. This implies that in any given history, Θ induces a partitioning of tests in Σ into disjoint blocks of tests such that all tests in a specific block map to a unique test in Σ'. The block-test in Σ that maps to test t' in Σ' from history h is denoted $\Theta_h^{-1}(t')$. By definition, all tests that belong to the same block will be of the same length. However, tests in the same block do not have to specify the same observation or action sequence. We use $\mathbf{a_B}$ to denote all the action sequences specified by tests in the block B, and $\mathbf{o_B}$ to denote all the observation sequences specified by tests in B.

We say that a block-test B 'succeeds' in history h if some observation sequence in $\mathbf{o_B}$ is obtained given that a particular action sequence in $\mathbf{a_B}$ is executed in h. As shorthand, in a slight abuse of notation, we denote this as $Pr(B|h)$. So for all action sequences $\{a_1, \cdots, a_n\} \in \mathbf{a_B}$,

$$Pr(B|h) = \sum_{\{o_1,\cdots,o_n\} \in \mathbf{o_B}} Pr(o_1, \cdots, o_n | h a_1 \cdots a_n). \quad (4)$$

Equation 4 requires that $Pr(B|h)$ be the probability of seeing $\mathbf{o_B}$ for every action sequence in $\mathbf{a_B}$.

The following theorem establishes the algebraic relationship between tests related by Θ, essentially stating that the probability of B succeeding in history h is the probability that the corresponding test t' succeeds in $\Theta_\emptyset(h)$.

Theorem 2. *Let Θ be a homomorphism from Σ to Σ'. Let $t' = \{a'_1 o'_1, \cdots, a'_n o'_n\}$ be a test in Σ'. Define B to be the block $\Theta_h^{-1}(t')$ for any history h. Then,*

$$P(B|h) = Pr(t'|\Theta_\emptyset(h)).$$

Proof. Consider any action sequence $\{a_1, \cdots, a_n\} \in \mathbf{a_B}$. Let t_{n-1} be shorthand for the sequence $a_1 o_1, \cdots, a_{n-1} o_{n-1}$. Then $P(B|h)$ is:

$$\begin{aligned} &= \sum_{\{o_1,\cdots,o_n\} \in \mathbf{o_B}} Pr(o_1, \cdots, o_n | h a_1 \cdots a_n) \\ &= \sum_{o_1 \cdots o_n \in \mathbf{o_B}} Pr(o_1 | h, a_1) \cdots \; Pr(o_n | h, t_{n-1}, a_n) \\ &= \sum_{o_1 \in u_{h,a_1}^{-1}(o'_1)} Pr(o_1 | h a_1) \cdots \sum_{o_n \in u_{h,t_{n-1},a_n}^{-1}(o'_n)} Pr(o_n | h, t_{n-1}, a_n). \end{aligned}$$

Let t'_{n-1} be shorthand for $a'_1 o'_1, \cdots, a'_{n-1} o'_{n-1}$. Invoking Theorem 1, the above equation is:

$$\begin{aligned} &= Pr(o'_1 | \Theta_\emptyset(h) a'_1) \cdots Pr(o'_n | \Theta_\emptyset(h), t'_{n-1}, a'_n) \\ &= Pr(o'_1 \cdots o'_n | \Theta_\emptyset(h) a'_1 \cdots a'_n) \\ &= Pr(t' | \Theta_\emptyset(h)). \end{aligned}$$

□

Thus a prediction of a test in the image Σ' gives us a prediction of a block of tests in Σ. But what kinds of questions does the prediction of a block-test answer? Since a block-test can contain several different observation sequences, the prediction of a block-test is the probability that one of these sequences will occur for some sequence of actions. These predictions of block-tests thus generalize over observation sequences contained in that block-test.

Let us go back to in the example of Figure 2. In the previous section we described the PSR homomorphism that maps the domain in Figure 2a to the domain in the image PSR Figure 2c. Now consider the following test in Figure 2c: move left, and observe 'C'. This translates to a block-test comprised of 6 tests in the original domain: (a) move east and see *light gray*, (b) move east and see *dark gray*, (c) move east and see *black*, (d) move west and see *light gray*, (e) move west and see *dark gray*, and (f) move west and see *black*.

A prediction of this block would be a question of the form, "what is the probability that I will move east and see *light gray* OR *dark gray* OR *black*?", and will not be able to disambiguate between the observations contained in the block. Moreover this block-test can make the same prediction for moving west. In this manner, it also captures the symmetric nature of the *move-east* and *move-west* actions.

This example shows that predictions of block-tests abstract over observations and capture symmetry in actions. Thus, these predictions provide answers to structurally abstract questions about the world.

Temporal Abstractions with Homomorphisms

In the formalism of homomorphisms developed thus far, a homomorphic image of a domain cannot be used to answer temporally abstract questions about the domain. A question (and its associated prediction) is said to be temporally abstract if it generalizes over action and observation sequences of different lengths. Questions of the form, "If I walk down this hallway, what is the probability that I will see a door?" abstract out the length of time it takes to walk down the hallway and anything else that we might observe in the process. As has been our motivating theme, we would like to scale the complexity of the domain with the degree of temporal coarseness of our questions. To get a more concrete idea of the types of temporal abstractions we wish to model, consider again the domain of Figure 1a. The following is a sample of the type of temporally abstract questions one can ask:

- "If I behave in a particular way (for instance, take random actions at every time-step), what is the probability I will see a black square in two time steps?". Questions such as these are only concerned with the final observation. Not only do they ignore the actions and observations that might occur along the way, they also ignore the length of any such sequences.
- "If I behave in a particular way, what is the probability that I will see three white squares in succession?". Questions such as these care about more than just the last observation. However, they still generalize over the length of observation sequences and over all the possible action sequences that eventually result in three white squares.
- "If I behave in a particular way, what is the probability that I will see a black square after I turn east in a white square?". This question abstracts over the length of time,

and over any actions and observations that might have occurred before “white, east, black” is observed.

This is obviously not an exhaustive list, but it gives an idea of the kind of temporal abstractions that are possible. Rather than being conditioned on action sequences, these questions are conditioned on ‘behaviors’ that prescribe some way of acting in the world. We model these behaviors using the framework of *options*. Options were first defined for MDPs by Sutton, Precup, & Singh (1999) and can be viewed as temporally extended actions that prescribe a closed loop behavior until some termination condition is met.

We extend PSR homomorphisms to capture this type of temporal abstraction. The basic idea is the same as before - a homomorphism will partition the space of tests in the original PSR into block-tests, and then map each of these blocks to a unique test in the image PSR. In the previous section, structural abstraction was obtained by creating block-tests that contained different observation sequences. It makes sense then that temporal abstraction can be obtained by creating block-tests that contain tests of different lengths. We will equip PSR homomorphisms to achieve such groupings.

Next, we formalize options for PSRs and define the concepts of *option histories* and *option tests* (as opposed to primitive histories and tests). We show that the definition of PSR homomorphisms implies a mapping from option tests in the original PSR to primitive tests in the image PSR. This extension will give us the power and flexibility we need to capture the kind of temporal abstractions outlined above.

Options

An option ω can be thought of as a temporally extended action that is defined by: a) a closed loop *policy* π^ω that gives a probability distribution over actions for any history; b) an *initiation set* of histories in which the option can be executed; and c) a set of *termination conditions*, $\beta_h \in [0,1]$ that specify the probability of the option terminating in any given history h. Coupled with the system dynamics, an option policy also maps histories to a probability distribution over action sequences, which we denote $\pi^\omega(a_1 \cdots a_n|h)$.

Option tests are defined as a sequence of options and primitive tests, i.e. $\bar{t} = \omega_1 t_1 \cdots \omega_n t_n$ where $t_1 \cdots t_n$ are primitive tests. Similarly, an option history is a sequence of options and primitive tests from the beginning of time, $\bar{h} = \omega^1 t^1 \cdots \omega^n t^n$. We can trivially represent a primitive action by wrapping it in an option that deterministically takes the action and then terminates. Similarly, we can trivially define a primitive test $t = ao$ as an option-test $\bar{t} = \omega t$ where ω is a wrapper for action a.

An option-test $\bar{t} = \omega_1 t_1 \cdots \omega_n t_n$ is said to succeed if all the primitive tests $t_1 \cdots t_n$ succeed in sequence given that the option sequence $\omega_1 \cdots \omega_n$ is executed. The probability of success is given by:

$$\begin{aligned} Pr(\bar{t}|\bar{h}) &= Pr(t_1 \cdots t_n|\bar{h}\omega_1 \cdots \omega_n) \\ &= Pr(t_1|\bar{h},\omega_1)\prod_{i=2}^{n} Pr(t_i|\bar{h},\{\omega t\}_{i-1},\omega_i) \end{aligned}$$

where $\{\omega t\}_{i-1}$ denotes the sequence $\omega_1 t_1 \cdots \omega_{i-1} t_{i-1}$. The probability that primitive test t succeeds when option ω is executed in history $\bar{h}$ is:

$$Pr(t|\bar{h}\omega) = Pr(t|\bar{h})\pi^\omega(a_1 \cdots a_n|h).$$

A primitive test t is said to be a possible *outcome* of an option ω in history $\bar{h}$ if $Pr(t|\bar{h}\omega) > 0$. Note that we can condition the prediction of a primitive test on an option history since since a history represents actions and observations that have already occurred. Thus an option history $\bar{h} = \omega^1 t^1 \cdots \omega^n t^n$ can always be converted into a unique equivalent primitive history $h = t^1 \cdots t^n$.

Temporally Abstract Homomorphisms

While options are temporally abstract actions, option-tests are not temporally abstract tests - the length of any option-test $\omega_1 t_1 \cdots \omega_n t_n$ is just the sum of the lengths of the component tests $t_1 \cdots t_n$. We now extend the framework of homomorphisms to capture temporal abstraction over tests. Consider a PSR Σ and it's image PSR Σ'. Let Ω refer to the set of options in Σ and let T refer to all tests that are possible outcomes of options in Ω in any history. Let A' and O' denote the action and observation sets in Σ'.

Define $u_{\bar{h},\omega} : T \rightarrow O'$ as a surjection from tests in T to observations in O' in history $\{\bar{h},\omega\}$. Define $v_{\bar{h}} : \Omega \rightarrow A'$ as a surjection from options in Ω to actions in A' in history $\bar{h}$. Let $\Theta = \langle u, v\rangle$ so that in any history $\bar{h}$, $\Theta_{\bar{h}}(\omega, t) = \{v_{\bar{h}}(\omega), u_{\bar{h},\omega}(t)\}$. As before, Θ also induces a surjection from histories in Σ to Σ'. Let $\bar{h} = \omega^1 t^1 \cdots \omega^n t^n$ be a history in Σ. The corresponding history in Σ' is $\Theta_\varnothing(\bar{h}) = \{\Theta_\varnothing(\omega^1, t^1), \cdots, \Theta_{\bar{h}^{n-1}}(\omega^n t^n)\}$, where $\bar{h}^{n-1}$ is used as shorthand for $\omega^1 t^1 \cdots \omega^{n-1} t^{n-1}$.

The next theorem specifies the condition under which Θ will be a homomorphism from Σ to Σ'.

Theorem 3. *Let Σ be a PSR with options Ω and outcomes T, and Σ' be a PSR with actions A' and observations O'. Define the tuple $\Theta = \langle u, v\rangle$ as above. The function Θ is a homomorphism from Σ to Σ' if and only if for all option histories $\bar{h}$ in Σ, options $\omega \in \Omega$, and observations $o' \in O'$*

$$Pr(o'|\Theta_\varnothing(\bar{h}), v_{\bar{h}}(\omega)) = \sum_{t \in u^{-1}_{\bar{h},\omega}(o')} Pr(t|\bar{h},\omega). \quad (5)$$

Proof. We first prove the forward direction. That is, given that Equation 5 holds, we want to show that Θ is a homomorphism from Σ to Σ', i.e satisfies the condition of Definition 1. Let $\bar{h}_1$ be a history in Σ and let $h'_1 = \Theta_\varnothing(\bar{h}_1)$ be the corresponding history in Σ'. Consider an option $\omega \in \Omega$ and action $a' = v_{\bar{h}}(\omega)$. The probability that we take action a' and arrive in history $h'_2 = h'_1 a' o'$ is

$$\begin{aligned} Pr(h'_2|h'_1, a') &= Pr(h'_1 a' o'|h'_1 a') \\ &= Pr(o'|h'_1 a') \\ &= \sum_{t \in u^{-1}_{\bar{h}_1,\omega}(o')} Pr(t|\bar{h}_1\omega) \\ &= \sum_{t \in u^{-1}_{\bar{h}_1,\omega}(o')} Pr(\bar{h}_1\omega t|\bar{h}_1\omega) \\ &= \sum_{\bar{h}_2 \in \Theta^{-1}_\varnothing(h'_1)} Pr(\bar{h}_2|\bar{h}_1\omega). \end{aligned}$$

Since predictive state is a sufficient statistic for history, the above result implies

$$Pr(\psi'|h'_1, a') = \sum_{\psi \in \Theta_{\varnothing}^{-1}(\psi')} Pr(\psi|\bar{h}_1\omega)$$

Since all of the above equations are equalities, the reverse direction also holds. Thus, Θ induces a partitioning over the predictive state space in Σ such that the conditions of definition 1 are satisfied. □

Answering Temporally Abstract Questions

We now examine how PSR homomorphisms lead to a mapping from blocks of option tests in Σ to primitive tests in Σ'. Let $\bar{t} = \omega_1 t_1 \cdots \omega_n t_n$ be an option-test in Σ. In any history $\bar{h}$, the corresponding primitive test in Σ' is $\Theta_{\bar{h}}(\bar{t}) = \{\Theta_{\bar{h}}(\omega_1 t_1), \cdots, \Theta_{\bar{h}\bar{t}_{n-1}}(\omega_n t_n)\}$ where $\bar{t}_{n-1}$ is used as shorthand for the sequence $\omega_1 t_1 \cdots \omega_{n-1} t_{n-1}$.

Since Θ is a surjection, it groups option-tests in Σ into blocks such that all option-tests in a particular block map to the same primitive test in Σ'. Option tests in a particular block do not have specify the same option sequence or the same primitive-test sequence. Let $\omega_{\mathbf{B}}$ denote the option sequences contained in a block B of option-tests, and let $\mathbf{t_B}$ denote the sequences of primitive tests contained in B.

Moreover, unlike in the previous section, option-tests in a particular block do not have to be of the same length either. For instance, if in some history h, as long as $Pr(t_1|\bar{h}\omega) = Pr(t_2|\bar{h}\omega)$, option-tests ωt_1 and ωt_2 can belong to the same block. No restriction is placed on their lengths.

Going back to our example in Figure 1, let ω be an option that selects a random action at every time step. Consider the test, "execute ω and see a black square in the next two timesteps". This is a block test that is comprised of several one-step option tests. Each of these option tests is of the form ωt where t is any length-1 or a length-2 primitive test ending in observation "black" (e.g. (a) *move-east* and see black; (b) *move-south* , see white, *move-east* , see black; etc.). So we see that a block-test can contain primitive tests of different lengths and different observation sequences.

As in the previous section, we say that a block test B 'succeeds' in history $\bar{h}$ if some primitive-test sequence in $\mathbf{t_B}$ is observed given that a particular option sequence in $\omega_{\mathbf{B}}$ is executed. So for all option sequences $\omega_1 \cdots \omega_n \in \omega_{\mathbf{B}}$, and any history h, the probability that B succeeds is

$$Pr(B|\bar{h}) = \sum_{t_1 \cdots t_n \in \mathbf{t_B}} Pr(t_1 \cdots t_n|\bar{h}\omega_1 \cdots \omega_n).$$

If $\mathbf{t_B}$ contains primitive tests of different lengths, the prediction of block-test B will generalize over time.

We now show how a question in the image PSR translates to a temporally abstract question in the original PSR. The analogue of Theorem 2 in this context is given below, and states that the prediction of any block-test B succeeding in history $\bar{h}$ is the probability that the corresponding primitive test t' succeeds in the $\Theta_{\varnothing}(h)$.

Theorem 4. *Let Θ be a homomorphism from Σ to Σ'. Let $t' = \{a'_1 o'_1, \cdots, a'_n o'_n\}$ be a test in Σ'. Define B to be the block $\Theta_{\bar{h}}^{-1}(t')$ for any history $\bar{h}$. Then,*

$$P(B|\bar{h}) = Pr(t'|\Theta_{\varnothing}(\bar{h})).$$

Thus, the prediction of any primitive test in the image PSR translates to the prediction of a block test in the original PSR. If this block test generalizes over lengths of tests, the translated prediction will generalize over time. If the block test also generalizes over observations, the translated prediction will also generalize over those observations. Thus, these translated predictions can be answers to structurally and temporally abstract questions in the original domain.

Conclusion

In this work, we are motivated by the desire to scale the complexity of a domain with the complexity of questions we seek to answer. We presented PSR homomorphisms as a theoretical framework to transform a given PSR into an equivalent smaller image PSR that answers certain questions in the original PSR. We have shown that these questions capture structural and temporal abstractions in the original PSR.

Since we were only interested in making predictions, we have ignored rewards in this paper. In future work, we will incorporate rewards into our framework for PSR homomorphisms and examine the implications this has on the problem of planning.

Acknowledgments

Vishal Soni and Satinder Singh were supported by the NSF Grant Number CCF-0432027 and by a grant from DARPA's IPTO program. Any opinions, findings, and conclusions or recommendations expressed in this material are solely of the authors and do not reflect the views of the NSF or DARPA.

References

Dean, T., and Givan, R. 1997. Model minimization in markov decision processes. In *American Association for Artificial Intelligence*, 106–111.

Littman, M. L.; Sutton, R. S.; and Singh, S. 2001. Predictive representations of state. In *Neural Information Processing Systems*.

Ravindran, B., and Barto, A. G. 2002. Model minimization in hierarchical reinforcement learning. In *Proceedings of the Fifth Symposium on Abstraction, Reformulation and Approximation*, 196–211.

Ravindran, B., and Barto, A. G. 2003. SMDP homomorphisms: An algebraic approach to abstraction in semi-Markov Decision Processes. In *Proceedings of the Eighteenth Internatinal Joint Conference on Artificial Intelligence (IJCAI 2003).* Menlo Park, CA: AAAI Press.

Sutton, R. S.; Precup, D.; and Singh, S. P. 1999. Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence* 112(1-2):181–211.

Van der Schaft, A. 2004. Bisimulation of Dynamical Systems. In *Hybrid Systems: Computation and Control*, 555–569.