

Planning Interventions in Biological Networks

DANIEL BRYCE

Utah State University

MICHAEL VERDICCHIO, and SEUNGCHAN KIM

Arizona State University

Modeling the dynamics of biological processes has recently become an important research topic in computational biology and systems engineering. One of the most important reasons to model a biological process is to enable high-throughput *in-silico* experiments that attempt to predict or intervene in the process. These experiments can help accelerate the design of therapies through their rapid and cheap replication and alteration. While some techniques exist for reasoning with biological processes, few take advantage of the flexible and scalable algorithms popular in AI research. In reasoning about interventions in biological processes, where scalability is crucial for feasible application, we apply AI planning-based search techniques and demonstrate their advantage over existing enumerative methods. We also present a novel formulation of intervention planning that relies on models that characterize and attempt to change the phenotype a system. We study three biological systems: the yeast cell cycle, a model of the human aging process, and the Wnt5a network governing the metastasis of melanoma in humans. The contribution of our investigation is in demonstrating that i) prior approaches, based on dynamic programming, cannot scale as well as heuristic search, and ii) the newly found scalability enables us to plan previously unknown sequences of interventions that reveal novel and biologically significant responses in the systems which are consistent with biological knowledge in the literature.

Categories and Subject Descriptors: J.3 [**Computer Applications**]: Life and Medical Sciences—*Biology and genetics*; I.2.8 [**Computing Methodologies**]: Problem Solving, Control Methods, and Search—*Plan execution, formation, and generation*

General Terms: Algorithms, Experimentation, Performance

Additional Key Words and Phrases: Planning, systems biology, and search

1. INTRODUCTION

Biological systems maintain their functions via various interconnections and regulatory controls among genes and proteins, and their interaction with other molecules. Therefore it is critical to unravel such interactions in order to understanding biological systems. Mathematical modeling and computational simulation of biological systems, especially gene regulatory systems, has been the crux of computational systems biology and biomedical science in general. Once a model is constructed, it can be used to predict the behavior of a system under unusual conditions, to identify how a disease might develop, and/or how to intervene in such development to prevent the system from reaching undesirable states. Such models can aid biologists to rule out or confirm hypotheses before expensive wet lab experiments. In this paper, we address the problem of planning to intervene in biological processes, focusing on gene regulatory networks (networks describing only genes) and organism-wide networks (including genes, proteins, and molecules). Specifically, we find plans that change the phenotype (stable behavior) of a biological process, as we describe below.

Biological Networks: A biological network describes a process by a set of genes, proteins, and other molecules, and their influences over each other. Models of biological networks are either constructed by hand or are learned from microarray data [Kim et al. 2000]. Microarray experiments measure the activity levels of thousands of genes from living tissue in terms of mRNA concentrations (the products of gene transcription used to code proteins). Correlations between observed gene activity levels, for example, help describe regulatory influences through predictor functions.

In all of the systems that we study, predictor functions characterize the regulatory influences and provide a dynamic model (e.g., when g_1 and g_2 are highly active, g_3 becomes inactive). The state models the activity levels of genes and other entities, and the predictor functions describe possible next states. Successive applications of the predictor functions leads to a state sequence that demonstrating repeated states (called an attractor); these repeated states capture the stable behavior consistent with a phenotype. Outside interventions (e.g., using RNA interference to suppress a gene's activity level) alter predictor functions to control the dynamics of the network, providing a natural planning problem – changing the phenotype is of primary interest. In a plan, each action is either a possible intervention or non-intervention that will change the state of the system. We study three systems: yeast cell cycle, human aging, and melanoma (referred to by the most important gene, Wnt5a), the first two are manually constructed, while the last is learned from microarray data. We provide a biological interpretation of the plans constructed within each network, but refer the reader to prior related work on the derivation details of each network.

Planning: Planning is the problem of synthesizing a sequence of actions to transform an initial state into a goal state. A number of assumptions (e.g., full state observability, atomic duration actions, deterministic actions, etc.) in the classical model, when relaxed, give rise to several alternative models and types of plans. There are many features of the intervention planning problem that affect our choice of AI planning model. First, intervention plans need only focus on horizons long enough to ensure the network will naturally transition to stable sets of states. Biological knowledge and computational simulation of networks tell us that biological processes, left to their own, transition to (or through) stable attractor states that represent phenotypes [Kim et al. 2002; Kauffman 1969b]. Phenotypes characterize common behavioral patterns of biological systems, such as the cell cycle, accelerated or normal aging, and disease, such as the metastasis of cancer. From abnormal-phenotype states, planning interventions provides a method to direct the evolution of the system toward other, favorable phenotypes. Second, the network has an inherently hidden state because full observations are prohibitively costly and we have limited accessibility [Datta et al. 2004]. Planning with partial observability is important because biologists analyzing biological processes cannot be expected to understand, nor obtain complete state information. Third, cellular processes are commonly viewed as stochastic [Elowitz et al. 2002]. Genes are typically regulated in many different ways, meaning networks must allow for the probabilistic selection of predictor function for each gene. *Because state transitions are stochastic, provide only partial observations, and are only relevant over a limited horizon, we formulate our model as decision theoretic planning in a finite-horizon partially observable*

Markov decision process (POMDP).

This work is not the first to address the problem of planning interventions for biological networks, but it is the first to i) apply AI planning techniques (aside from our preliminary work [Bryce and Kim 2007]), and ii) consider planning to change the phenotype of a system. Existing research [Datta et al. 2004] enumerates all possible plans and then uses a dynamic programming algorithm to find the optimal finite horizon plan, and considers problems where goal includes the change of a single feature (rather than reaching a new stable behavior, or phenotype). Our planner is based on the A0* algorithm [Nilsson 1980], which uses upper bounds on plan quality (reward) for pruning. We evaluate our planner on several problems with different reward functions to see the benefit of pruning.

Our planner performs significantly better than the enumeration algorithm of [Datta et al. 2004]. Our study formalizes this interesting planning domain as a finite horizon POMDP and shows the gains of applying relatively standard AI techniques. We also make predictions of interventions that can change each of the three systems studied; one prediction, in the yeast cell cycle, was independently verified [Ferrezuelo et al. 2010] after we completed our experiments.

We start by providing background on the methodology used to model biological networks, and then describe a simple example of planning interventions in a gene regulatory network (GRN). We then briefly define the components of a finite horizon POMDP and an intervention problem. With these definitions, we map the intervention problem to the POMDP. The solution to the POMDP is a conditional plan, for which we define the semantics and describe two solution algorithms – A0* and enumeration [Datta et al. 2004]. We compare both algorithms within several networks that are either hand-crafted or learned from actual microarray experiments. We end with related work, a conclusion, and discussion of future work.

2. METHODS

In this section we provide a brief overview of Boolean networks, which will serve as our formalism for describing the behavior of a biological process.

Boolean Networks: A Boolean Network (BN) [Xiao 2009; Wuensche 1998; Kauffman 1969b] is made up of a set of binary variables and a corresponding set of functions which determine the state of the variables at the next time step based on their values at the current time step. Connecting variables based on the inputs and outputs of these functions creates a wiring diagram structure, a graphical representation of the Boolean network.

The state transition system induced by a Boolean network with n nodes is a directed graph with 2^n vertices. Each of the vertices represents one possible configuration of the n variables in the network and each of the directed edges represents the transition between two states as Boolean functions are synchronously applied to all variables.

In the absence of interventions and beginning in any initial state, repeated application of the functions will bring the network to a finite set of states and cycle among them forever in fixed sequence. This set of states is known as an *attractor*, and corresponds to a phenotype. An attractor with just one state is called a point attractor and an attractor with more than one state is called a cyclic attractor.

The complete set of states from which a network will eventually reach an attractor is known as the *basin of attraction* (or basin) for that attractor. Boolean networks may have anywhere from one cyclic attractor comprised of 2^n states to 2^n point attractors, although a network will normally have just a handful of singleton or short-cycle attractors.

A Probabilistic Boolean Network (PBN) [Shmulevich et al. 2002; Xiao 2009] is a stochastic extension of a BN that induces a Markov chain [Kim et al. 2002]. PBNs are closely related to Bayesian networks, differing in how the parents of a node determine its value. Our formulation in Section 4 uses these robust networks to allow probabilistic selection between several predictor functions that determine the next value of a variable. The advantage of a PBN over a Bayesian network, is mainly seen from the model formulation perspective, because it admits fewer parameters (one number for each predictor function versus one number for each joint assignment of parents). We maintain consistency with prior work on biological systems [Kim et al. 2000; Datta et al. 2003; Datta et al. 2004; Shmulevich et al. 2002] by using (probabilistic) Boolean networks, but there is no technical reason prohibiting reliance on Bayesian network models.

We focus on Boolean networks because of their simplicity, forgoing alternative models such as those based on differential equations (which allow continuous states and time), in favor of larger-scale, albeit more granular, reasoning about change. Despite their simplicity, Boolean networks have been found to display many properties critical to biological processes, such as homeostasis and switch-like behavior [Kauffman 1969a]. Recently, Boolean networks have been used to simulate the yeast cell cycle [Li et al. 2004; Davidich and Bornholdt 2008], and the expression pattern of the segment polarity genes in the fruit fly, *Drosophila melanogaster* [Albert and Othmer 2003].

Each basin has a set of attractor states which correspond to the steady-state behavior of the system in that basin. The other states, while part of the basin, represent transitory anomalous states (e.g., in response to environmental stress); the existence of these states, and the fact that the system (of its own accord) will transition from these states to the attractor states, corresponds to the conventional knowledge that biological systems are self-regulating [Li et al. 2004]. Thus, transitions within a basin result from the natural dynamics of the system.

Interventions: Of interest to this work are unnatural responses in the biological system, introduced by what we call interventions, that can transition the system between phenotypes. Planning sequences of interventions is necessary, as we will show, because a single intervention is not always enough to transition between basins. *In the problems we study, we will assume that the biological system starts in one of the attractor states of a basin (however, we may not know which attractor state), and construct a plan to reach any state in another basin.* The intuition is that the initial state in the source basin will be one of the attractor states because these are the steady-states of the basin, and it is enough to reach any state of the destination basin because thereafter the natural dynamics of the system will transition the state to one of the attractor states of the basin.

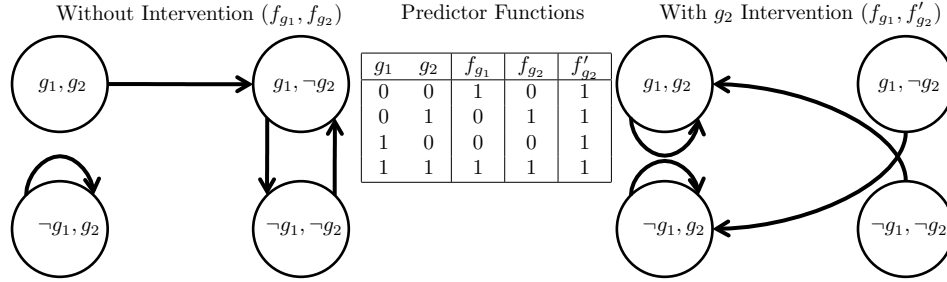


Fig. 1. Example process under differing predictor functions.

3. INTERVENTION PLANNING EXAMPLE

To clarify intervention planning, consider a small Boolean network with two genes, where each gene g_i is either active (g_i) or inactive ($\neg g_i$). There are two predictor functions $\{f_{g_1}, f_{g_2}\}$ (each determining the influences on a gene) that describe the system dynamics, and an intervention to activate g_2 , described by the predictor function f'_{g_2} . Figure 1 illustrates the biological process by the predictor function definitions (center), the state transition system under the natural dynamics (left), and the state transition system subject to g_2 intervention (right). The system, subject to the natural dynamics, has two phenotypes, each with a basin of attraction and attractor. The first basin of attraction includes states $\{g_1, g_2\}$, $\{g_1, \neg g_2\}$, $\{\neg g_1, \neg g_2\}$, where states $\{g_1, \neg g_2\}$, $\{\neg g_1, \neg g_2\}$ comprise the basin's cyclic attractor, and $\{g_1, g_2\}$ is an anomalous transitory state that leads to the basin. The second basin of attraction includes state $\{\neg g_1, g_2\}$, which is a point attractor.

A plan to change from the first to the second phenotype is depicted by Figure 2. The plan starts in a belief state that assigns equal probability to the attractor states of the first basin (because we know that the system exhibits the first phenotype, but we are unsure of which attractor state it is in). The belief state includes arcs for state transitions (attained from the predictor functions) corresponding to the action applied in the belief state. The initial belief state (the leftmost belief state) denotes that the g_2 intervention action is applied in it. We allow the practitioner to observe g_1 , meaning that after applying the first action, there is a branch in the plan for each outcome of the observation. In both of the subsequent belief states the plan applies the non-intervention action. The upper branch follows with the g_2 intervention action, and the lower with non-intervention because the goal (reaching the second phenotype) is attained. At a horizon of three, the plan guarantees reaching the second phenotype in all plan branches.

4. DECISION THEORETIC PLANNING

Unlike most work on POMDPs¹, which find a policy over the belief state space, we find a policy (conditional plan) rooted at an initial situation. The finite horizon POMDP problem P is defined as the tuple $\langle S, A, T, R, \Omega, O, h, b_I \rangle$, where S is a set

¹Exceptions include the recent work on improving upon point-based value iteration [Smith and Simmons 2004; Pineau et al. 2006]

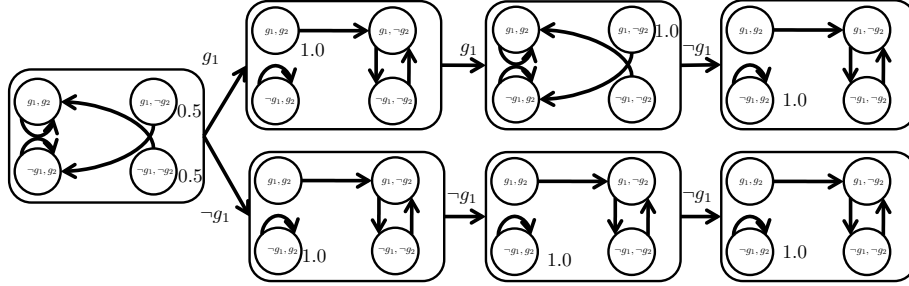


Fig. 2. Example plan to change phenotype.

of states, A is a set of actions, $T : S \times A \cup \{\perp\} \times S \cup \{\perp\} \rightarrow [0, 1]$ is a transition function, $R : S \times A \cup \{\perp\} \times S \cup \{\perp\} \rightarrow \mathbb{R}$ is the transition reward function, Ω is a set of observations, $O : S \times A \times \Omega \rightarrow [0, 1]$ is the observation function, h is the planning horizon, and $b_I : S \rightarrow [0, 1]$ is the initial belief state. We overload the symbol \perp to denote both the terminal action and state signifying the end of the plan.

Our rationale for selecting this POMDP formulation is two-fold: i) focussing on a plan for an initial belief state allows us to plan to transition from the states consistent with a particular phenotype (with which we are uncertain which is the current state), and ii) we can compare with the prior work of Datta et. al. [Datta et al. 2004]. While we could plan for all initial belief states using a more traditional POMDP value iteration algorithm [Kaelbling et al. 1995], many of the initial belief states are irrelevant to biologists, making the additional overhead unjustified. As we show in the following, the POMDP model is a natural choice of planning framework for reasoning with probabilistic Boolean network models of biological processes (incorporating probabilistic state transitions and partial observability of the initial phenotype states).

Intervention Planning Problem: In the following we show how an intervention planning problem maps to the POMDP. The intervention problem is defined by the tuple $\langle \mathcal{G}, Dom, F, X, W, Y, \mathcal{O}, h \rangle$, where \mathcal{G} is a set of variables (genes, proteins, and/or molecules), Dom is the set of activity levels for variables, F is a set of predictor functions, X is a set of interventions, W is an initial situation, Y is a goal description, \mathcal{O} is a set of observations, and h is the horizon. The variables and their activity levels describe states of the POMDP, the predictor functions and interventions describe actions, interventions and the goal description define the reward function, and the initial situation, observations, and horizon map directly to their POMDP counterparts.

States: Each variable $g \in \mathcal{G}$ has an activity level from the domain Dom of values, of which we will only illustrate boolean domains $\{g, \neg g\}$, active or inactive. A state $s : \mathcal{G} \rightarrow Dom$ of the network maps each variable to a value $d \in Dom$. The entire set of network states defines the POMDP states S . Our choice of Boolean valued variables is based on the, previously mentioned, prior work in Boolean networks

used to model the biological process.

Predictor Functions and Interventions: Given a state of the network, the predictor functions F describe states reachable after one step. Interventions rewrite predictor functions in F for specific variables to ensure the network transitions to specific states. Thus, each possible action in the POMDP is described by a set of predictor functions. A non intervention simply uses F to describe the action, but an intervention action corresponding to an $x \in X$ replaces predictor functions in F to attain a new set F_x . Each intervention $x \in X$ is a set of predictors $\{f_{g_1}, f_{g_2}, \dots\}$, allowing us to define

$$F_x = F \cup x \setminus \{f_g | f_g \in F, f_{g'} \in x, g = g'\}.$$

Each predictor function f_g is defined as the mapping $f_g : Dom^{|\mathcal{G}'|} \rightarrow Dom$ from activity levels of variables in $\mathcal{G}' \subseteq \mathcal{G}$ to the activity level of variable g .

The interventions we describe in this work contain a single predictor function f_g where $\mathcal{G}' = \emptyset$, meaning that irrespective of the state g has its activity level set deterministically. We select this class of interventions because they are the simple; however, our framework does not, in any way, prohibit alternative definitions of intervention predictor functions (e.g., we allow for multiple predictors to model a stochastic intervention, or conditional predictor functions). The potential drawback of assuming deterministic, unconditional interventions is that the intervention may not have the desired effect on the system (due to missing information about the conditions or stochastic nature of the intervention). This mismatch between the model and the system can be overcome when practitioners use the resulting intervention plans to guide their refinement of the model to match expectations put forth in the literature or their data.

Since each variable may be predicted by several predictor functions, where each state transition selects one probabilistically; we assign a weight $w(f_g)$ to each. While we assume the predictor functions are given, we can also learn them from microarray experiments. In the empirical analysis we evaluate networks where the predictor functions and weights are taken from the literature or from microarray data.

While we do not explore the option in this paper, some alternative models are based on context-sensitive (probabilistic) Boolean networks [Pal et al. 2005; Dougherty et al. 2009]. A context-sensitive PBN or BN defines alternative sets of predictor functions for different states, in a similar manner to which intervention actions define alternate predictor functions. Our framework can be extended in a straight-forward manner to incorporate context-sensitive networks.

Each action a_F defined by the set of predictor functions F (similarly F_x) describes the transition function probability

$$T(s, a_F, s') = Pr(s' | F, s) = \prod_{g \in \mathcal{G}} \sum_{\substack{f_g \in F: \\ f_g(s) = s'(g)}} w(f_g) / w_g,$$

where $w_g = \sum_{f_g \in F} w(f_g)$.

Observations: After each step, whether by intervention or non intervention, we can observe the state of the network. The set $\mathcal{O} \subseteq \mathcal{G}$ defines which variables are

observable (by genetic markers, physiology, etc.). The set of observations $\Omega = \{o | o \in \text{Dom}^{|\mathcal{O}|}\}$ is defined by all joint activity levels of variables in \mathcal{O} . In this work, we assume that observations are perfect and the same for each action, meaning that if a state s and observation o agree on the activity level of each variable, then the probability of the observation is one (i.e., $O(s, a, o) = 1$), otherwise zero (i.e., $O(s, a, o) = 0$). Observations can be noisy (i.e., $0 \leq O(s, a, o) \leq 1$) in general. We assume perfect observations because if an observation proves useful in our model, biologists can typically design effective measurements to minimize any observation noise.

Rewards: The goal Y is a function describing desirable states. We assume that the goal maps states to real values $Y : \text{Dom}^{|\mathcal{G}|} \rightarrow \mathbb{R}$. The reward function for terminal actions and goal states is defined by the goal $R(s, \perp, \perp) = Y(s)$. We assume that the reward associated with actions is negative one for intervention actions (i.e., $R(s, a_{F_x}, s') = -1$) and zero for non intervention (i.e., $R(s, a_F, s') = 0$). The goal reward for each problem studied is related to the states of a desired target phenotype.

Initial Situation: The initial situation W is a distribution over states $W : \text{Dom}^{|\mathcal{G}|} \rightarrow [0, 1]$. This mapping to the POMDP initial situation is straight-forward, $b_I(s) = W(s)$. The initial belief state in each problem is a distribution over the states of a source phenotype’s attractor – it is not clear which is the current state of the phenotype (basin of attraction), but we are certain of the currently possible set of states within the attractor. The reasoning is that we assume the system starts in a stable situation (after an sufficiently large number of applications of F), allowing us to rule out any of the transitory states in the basin of attraction; this means the system is in one of the attractor states. In the goal formulation, above, we assume that any state in the target phenotype’s basin of attraction is sufficient because after the plan terminates the system will continue to apply F indefinitely and eventually reach the target phenotype’s attractor states.

5. SEARCH

This section describes our approach to solving a finite horizon POMDP. We start by defining the semantics of conditional plans and the search space. We follow with two search algorithms to find plans. The first algorithm is A0* [Nilsson 1980], and the second **Datta** is based on a competing approach [Datta et al. 2004] from the computational biology literature; the difference between the algorithms relating to how the belief state space is expanded. We employ A0* because it is a natural choice when the POMDP model requires an initial belief state and a finite horizon (limiting search to those belief states reached from the initial belief state within an acyclic search graph).

Conditional Plans: A solution to the problem P is a conditional plan π of horizon h , described by a partial function $\pi : V \rightarrow A \cup \{\perp\}$ over the belief state space graph $G = (V, E)$. A subset of the vertices $b \in V$ (which are belief states) are mapped to a “best” action a , denoted $\pi(b) = a$. Each edge $e \in E$ directed from b to b_a^o is mapped to an action $a \in A$ and an observation $o \in \Omega$, and denoted $e(b, b_a^o) = (a, o)$. If $\pi(b) = a$ and $e(b, b_a^o) = (a, o)$, then $\pi(b_a^o)$ is the action to execute after executing

a in b and receiving observation o . Throughout our discussion, we assume that the horizon is a feature of every state to ensure that the graph G is acyclic. Belief states where the horizon is equal to h have a single available action \perp to signify the end of the plan, leading to a terminal \perp .

If $\pi(b) = a$, and there exists an edge $e(b, b_a^o) = (a, o)$ then the successor b_a^o is defined

$$\begin{aligned} b_a(s') &= \sum_{s \in S} b(s)T(s, a, s') \\ b_a^o(s') &= \alpha b_a(s')O(s', a, o), \end{aligned}$$

where α is a normalization constant. If for all $s \in S$, $b_a^o(s) = 0$ because no observation is consistent with the belief state b_a , then the belief state is not added to the graph. Also, if all child belief states are identical to the parent, then we do not add the children to the graph (which can happen when reaching a point attractor state, i.e., applying a non-intervention action and following the biological process predictor functions F leads to the same belief state indefinitely).

The expected reward $q(b, a)$ of a plan that starts with action a at belief state b is the sum of current and future rewards:

$$q(b, a) = \sum_{s \in S} \sum_{s' \in S} b(s)T(s, a, s')R(s, a, s') + \sum_{o \in \Omega} b_a(s')O(s', a, o)\mathcal{V}(b_a^o),$$

where the expected reward for a belief state is $\mathcal{V}(b)$. Terminal vertices are assigned the expected goal reward

$$q(b, \perp) = \sum_{s \in S} b(s)R(s, \perp, \perp).$$

A0* **Algorithm:** We solve the finite horizon POMDP problem with A0* search [Nilsson 1980] in the space of belief states. The A0* algorithm, listed in Figure 3, takes the planning problem as input, and iteratively constructs the belief space graph G rooted at b_I . The algorithm involves three iterated steps: expand the current plan with the **ExpandPlan** routine (line 3), collect the ancestors Z' of new vertices Z (line 4), and compute the current best partial plan (line 5). The algorithm ends when it expands no new vertices. In the following we briefly describe the sub-routines used by A0*.

The **ExpandPlan** routine recursively walks the current plan to find unexpanded vertices (lines 2-5). Upon finding a vertex to expand, it generates all successors of the vertex (lines 7-18). Generating successors involves assigning the \perp action if i) the vertex is at the max horizon or ii) all non-zero probability states in the belief state are goal states (line 10). Otherwise, the search constructs the vertices reached by all action and observation combinations (lines 12-17). Notice that each vertex has its value initialized with an upper bound,

$$R^{max} = \max_{s, s' \in S, a \in A} R(s, a, s')h + \max(0, R(s, \perp, \perp)),$$

on its expected reward. The upper bound plays a role in pruning vertices from consideration in the search.

After expanding the current plan, **ExpandPlan** returns the set of expanded vertices Z . In order for **Update** (Figure 4) to find the best plan, given the new vertices, **AddAncestors** adds to Z' every ancestor vertex of a vertex in Z . The resulting set

```

AO*( $P$ )
1:  $\text{expanded}(b_I) = \text{FALSE}$ 
2: repeat
3:    $Z = \text{ExpandPlan}(b_I, 0)$ 
4:    $Z' = \text{AddAncestors}(Z)$ 
5:    $\text{Update}(Z')$ 
6: until  $|Z| = 0$ 

ExpandPlan( $b, \text{hzn}$ )
1: if  $\text{expanded}(b)$  then
2:   for  $e(b, b_{\pi(b)}^o) \in E$  do
3:      $Z' = \text{ExpandPlan}(b_{\pi(b)}^o, \text{hzn}+1)$ 
4:      $Z = Z \cup Z'$ 
5:   end for
6: else
7:    $\text{expanded}(b) = \text{TRUE}$ 
8:    $Z = Z \cup \{b\}$ 
9:   if  $\text{hzn} == h$  or
      $\forall s. b(s) > 0 \implies b(s)R(s, \perp, \perp) \neq 0$  then
10:     $E = E \cup \{e(b, \perp) = (\perp, \perp)\}$ 
11:   else
12:     for  $a \in A, o \in \Omega$  do
13:        $V = V \cup \{b_a^o\}$ 
14:        $E = E \cup \{e(b, b_a^o) = (a, o)\}$ 
15:        $\text{expanded}(b_a^o) = \text{FALSE}$ 
16:        $\mathcal{V}(b_a^o) = R^{\text{max}}$ 
17:     end for
18:   end if
19: end if
20: return  $Z$ 

```

Fig. 3. AO* Search Algorithm For Belief State Space.

```

AddAncestors( $Z$ )
1:  $Z' = Z$ 
2: while  $\exists b$  s.t.  $e(b, b_{\pi(b)}^o) \in E$  and  $b_{\pi(b)}^o \in Z$ 
   do
3:    $Z' = Z' \cup b$ 
4: end while
5: return  $Z'$ 

Update( $Z$ )
1: while  $|Z| > 0$  do
2:   Remove  $b \in Z$  s.t.  $\neg \exists b' \in Z$  where
      $e(b, b') \in E$ 
3:   Backup( $b$ )
4: end while

Backup( $b$ )
1: for all  $a \in A \cup \{\perp\}$  do
2:   Compute  $q(b, a)$ 
3: end for
4:  $\mathcal{V}(b) = \max_{a \in A \cup \{\perp\}} q(b, a)$ 
5:  $\pi(b) = \operatorname{argmax}_{a \in A \cup \{\perp\}} q(b, a)$ 

```

Fig. 4. AO* Subroutines for Belief Space MDP.

```

Datta( $P$ )
1:  $\text{ExpandPlanD}(b_I, 0)$ 
2:  $\text{Update}(V)$ 

ExpandPlanD( $b, \text{hzn}$ )
1: if  $\text{hzn} == h$  or  $\forall s. b(s) > 0 \implies b(s)R(s, \perp, \perp) \neq 0$  then
2:    $E = E \cup \{e(b, \perp) = (\perp, \perp)\}$ 
3: else
4:   for  $a \in A, o \in \Omega$  do
5:      $V = V \cup \{b_a^o\}$ 
6:      $E = E \cup \{e(b, b_a^o) = (a, o)\}$ 
7:      $\text{ExpandPlanD}(b_a^o, \text{hzn}+1)$ 
8:   end for
9: end if

```

Fig. 5. Datta Enumeration Algorithm.

of vertices consists of every vertex whose value (and best action) can change after **Update**. The **Update** routine iteratively removes vertices from Z that have no descendent in Z and calls **Backup** until no vertices remain in Z . The **Backup** routine computes the value of a vertex and sets its best action. The reason **Update** chooses vertices with no descendent in Z is to ensure each vertex has its value updated with the updated values of its children.

AO* can often avoid computing the entire belief space graph G , leading to significant savings in problems with large horizons. By initializing vertices with an upper bound on their value it is possible to ignore vertices that have a consistently lowest upper bound. For example in **Backup**, if there exists an action whose q-value

is always greater than the alternative actions, then the best action will never be set to one of the alternatives. Further, because the alternative actions are never considered best, `ExpandPlan` will never expand them. As we will explore in the empirical evaluation, the reward function has a significant effect on the number of vertex expansions. In the worst case, it is possible to expand the entire graph G , as would `Datta`.

Enumeration Algorithm: In order to compare our planner to the work of [Datta et al. 2004], we provide a description of their algorithm, we call `Datta`, in Figure 5. Unlike the iterative AO^* , `Datta` consists of two steps: expand G with `ExpandPlanD`, and then update each vertex $v \in V$ with `Update`. The `ExpandPlanD` routine recursively expands G by either reaching a terminal vertex at the maximum horizon (line 2), or generating and recursing on each child of a vertex (lines 4-8). Following `ExpandPlanD`, `Update` computes the best action for each vertex in V .

Unlike AO^* , `Datta` is unable to prune vertices from expansion, making it insensitive to the reward function. While the result of the two algorithms is identical, the time and space required can be very different. We implemented both algorithms within our planner and demonstrate their effectiveness on the intervention planning problems.

6. BIOLOGICAL NETWORKS

Each biological network represents a different system, has different properties, and has several intervention problems of interest. In the following, we describe each network in detail after first defining some simple metrics used in evaluation.

Network Metrics: The first metric that we use to characterize the variables is their *popularity* [Verdicchio and Kim 2010]. A popular variable, in a particular phenotype (basin of attraction), holds the same value in many of the phenotype states. Variables with high popularity are good candidates for intervention targets because they can often determine the difference between phenotypes.

Another metric, the coefficient of determination (CoD) [Kim et al. 2002], is used to quantify the strength of a predictor function learned from the microarray data. As described below, the CoD metric was used, in part, to construct the Wnt5A network. The CoD provides a method to interpret the interventions used in a network based on their predictive influence over other network variables.

Yeast Cell Cycle Network: Li et al. [Li et al. 2004] manually construct a Boolean network modeling the yeast cell cycle using eleven of the most important genes out of the approximately 800 genes known to play a role in the process. The cell cycle is the process by which cells replicate and involves several phases, such as DNA replication and Mitosis. This network, when simulated, results in seven basins of attraction, one of which is by far the largest and was studied exclusively in prior work. In this basin of attraction, which included 1,764 states, Li et al. were able to trace the trajectory of the yeast cell cycle from one of the fringe, or "Garden of Eden" [Wuensche 1998], states down to the eventual point attractor.

The yeast cell cycle network is comprised of the variables summarized, in Table I. The table indicates the common name for each variable, whether it can be used as an intervention to set the variable to true \uparrow or false \downarrow , whether the variable is observ-

Variable	Int.	Obsv.	Description
Cln1,2	↓↑		Cyclin
Cln3	↓↑	✓	Cyclin
Clb1,2	↓↑	✓	Cyclin
Clb5,6	↓↑	✓	Cyclin
Cdh1		✓	Inhibits Cyclin Complexes
Cdc20			Inhibits Cyclin Complexes
Swi5	↓↑		Transcription Factor
Sic1			Inhibits Cyclin Complexes
Mcm1	↓↑	✓	Transcription Factor
MBF	↑	✓	Transcription Factor
SBF	↑	✓	Transcription Factor

Table I. Yeast Cell Cycle Network Variables

Beneficial Variable	Abbrv.	Int.	Obsv.	Description
Mitochondria	mito			Health of Mitochondria
Mitochondrial Metabolism	mtmet			Produces ROS as by-product
Mitochondria proteins	mtprot	↓↑		Healthy or Oxidized
Mitochondria DNA	mtdna	↑		Health of DNA influences mtmet
Inner mitochondrial membrane	imm			Subject to ROS damage
Lysosome	lyso			Processes spent proteins and mitos
Lysosome Membrane	lymem		✓	Contains digestive enzymes
Lon Protease	lonp			Needed to combat ROS stress
Nucleus DNA	ndna	↑		Health of nucleus DNA
Nucleus DNA repair	ndnarep	↑		Presence of proteins for DNA repair
Cellular Antioxidants	cellantx			Antioxidants produced by the body
Dietary Antioxidants	dietantx	↓↑	✓	Antioxidants introduced by diet
Harmful Variable	Abbrv.	Int.	Obsv.	Description
Apoptosis	apop	↓		(Premature) Programmed cell death.
Lipofuscin	lipofuscin	↓	✓	Residue of lysomal digestion
Ions	ions	↓	✓	Metal ions which react to form ROS
Reactive Oxygen Species	ros	↓↑	✓	Causes cellular damage
Hydrogen Peroxide	h2o2		✓	Subtype of ROS
Junk	junk		✓	Includes Damaged proteins and giant mitos

Table II. Aging Network Variables

able, and a brief description. The predictor functions can be obtained from our test instances: <http://www.cs.usu.edu/~danbryce/supplemental/intervention2010.tgz>. The yeast network is deterministic, using a single predictor function per variable. Each intervention planning problem involves the first basin, as either the source or destination. The rationale is that the first basin is the most biologically significant (representing the complete cell cycle) and accounts for nearly eighty-six percent of the states in the network. By finding intervention plans either to or from the first basin, we determine how to restore or disrupt the network’s nominal behavior. The goal assigns ten reward for reaching any state in the destination basin.

Aging Network: The systems biology of human aging is a complex, quantitative process. Many theories regarding senescence involve the roles of cellular components, such as mitochondria and lysosomes, as well the transportation and accumulation of various entities. In recent years, work by [Furber 2010] has

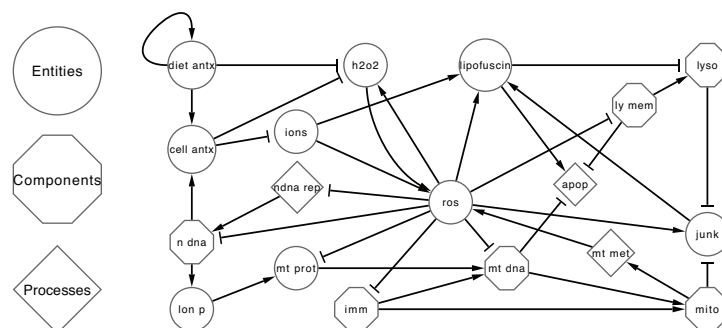


Fig. 6. Aging Boolean Network

Gene	Int.	Obsv.	Description
Wnt5a	↓ (WP, WP')	√ (PW, PW', MPW, MPW')	Tumor Suppressor
Pirin	↓ (PW, PW', MPW, MPW')	√ (WP, WP')	Nuclear Protein
Hadhb			Protein Coding Gene
Stc2			Regulates Cell Proliferation
Ret1			Protein Coding Gene
S100p			Codes Calcium Binding Protein
Mart-1	↓ (PW, PW', MPW, MPW')		Melanoma Gene

Table III. Wnt5a Network Variables

amalgamated the research of many prominent biogerontologists into a large chart illustrating many of the leading theories on human aging. The chart is organized into cellular components and describes many intricate, quantitative processes, along with their input and output entities. The chart can be viewed online at <http://legendarypharma.com/chartbg.html>. Prior work [Verdicchio and Kim 2010] converted a significant portion of Furber’s qualitative network into a formalized Boolean network, which is shown schematically in Figure 6. In the figure, the pointed edges indicate activating regulatory influence and those with blunted edges indicate inhibitory regulatory influence. For example, *ndna* is activated if *ndnarep* is active and *ros* is inactive.

This network, when simulated (repeatedly applying F from different initial states), partitioned the 2^{18} unique states into three basins of attraction. The first and largest basin led to what is classified as the “unhealthy” attractor, with all variables being fixed in undesirable configurations in the single attractor state. The second-largest basin led to what was classified as the “healthy” attractor, comprised of a 9-state cycle. The smallest basin led to what was classified as the “middle-of-the-road” attractor, which could not be called “healthy”, but fairly represented a state of “average health” in its 3-state cycle.

The aging network is comprised of the eighteen state variables summarized in Table II. The network is deterministic, meaning that there is a single predictor per variable, and the number of variables per predictor varies from one to three. The goal assigns ten reward to reaching the destination basin, as in the yeast network.

Wnt5a Network: The Wnt5a network centers around the Wnt5a protein, a prod-

uct of gene WNT5A, which is a member of the Wnt family of proteins. The WNT gene family consists of structurally related genes which encode secreted signaling proteins. These proteins have been implicated in oncogenesis (malignant transformation leading to the formation of a tumor) and in several developmental processes, including regulation of cell fate and patterning during embryogenesis (embryo development) [Weeraratna et al. 2002]. Specifically, the expression level of Wnt5a is closely related with metastatic status (spread from one organ to another) of melanoma [Bittner et al. 2000]. It was proven experimentally that increasing the level of Wnt5a protein can directly change the cell’s metastatic competence [Weeraratna et al. 2002; Sosman et al. 2004]. It has been also suggested that controlling the influence of Wnt5a in the regulation can reduce the chance of melanoma metastasizing [Weeraratna et al. 2002].

We will focus on a model constructed via various previous studies [Bittner et al. 2000; Kim et al. 2002; Weeraratna et al. 2002]. The focus is the network of elements of the Wnt5a signaling pathway that deal with the phenotypic change from a less motile, less aggressive cell to a much more motile, more invasive cell. Early attempts at constructing mathematical models of the network of genes showing shared regulatory information with Wnt5a resulted in the identification of an interesting relationship between Wnt5a and Mart-1 (MLANA) expression [Kim et al. 2002].

The Wnt5a network is comprised of seven state variables corresponding to genes, summarized in Table III. Each gene is predicted by two predictor functions, each constructed with the pair of genes with the respective highest and second highest CoD for the gene, as given by the microarray data. Each initial state for the network is given with the probability that it is afforded by the data (microarrays measure the activity levels of genes exhibiting a particular phenotype). The goal is to ensure that Wnt5a is inactive, so as to prevent metastasis. This formulation is slightly different from the yeast and aging networks because we do not plan to transition between basins, and is used to recreate the problems studied by prior work [Datta et al. 2004], which does not rely on basins and attractors. Without an explicit basin as the goal, we are not guaranteed that Wnt5a will remain suppressed, rather we simply reach a state where it is suppressed. However, a suppressed or activated Wnt5a is consistent with the major phenotypes of the network.

We reproduce two intervention problems studied by [Datta et al. 2004], and add a third. The first (denoted WP) directly intervenes to suppress Wnt5a (which happens to be the goal) and observes the Pirin gene. The second (denoted PW) attempts to indirectly control Wnt5a by Pirin (a predictor gene of Wnt5a) intervention. The third (denoted MPW) adds a possible interevent for Mart-1 to the PW problem; the predictor with the highest CoD for Wnt5a involves both Pirin and Mart-1.

7. EMPIRICAL EVALUATION

To test the feasibility of using our planner to solve intervention problems, we experiment with the three networks described in the previous section. There are two primary research questions that we address in this experimental evaluation:

- Do AI planning and search techniques improve upon prior work that solves intervention planning problems by dynamic programming?

—Do the intervention plans exhibit biological significance?

We continue by briefly describing the test environment and planner implementation, and discuss the two research questions with respect to each network.

Environment and Implementation: Our planner, POND² [Bryce et al. 2006], ran on a 2.0GHz Xeon Processor in the Linux OS. Due to the volume of experiments, each problem was given a twenty minute time limit and a 2GB memory limit; these limits were selected from preliminary data collection runs that showed no variation from the results presented below when using more generous limits. Results denoted by “-” indicate a failed attempt due to exhausted time or memory; in all cases time was the most limiting constraint. For each problem we report the total time in seconds to synthesize a plan and the expected reward of the plan.

POND is written in C++ and uses the following technologies to improve its efficiency over a naive implementation of the A0* and Datta algorithms. The actions are represented by dynamic Bayesian networks (DBNs) [Murphy 2002] and both the DBN conditional probability tables and belief states are represented by algebraic decision diagrams [Bryant 1986]. The problem instances are encoded in a modified version of the probabilistic planning domain description language (PPDDL) [Younes et al. 2005] to allow for probabilistic initial belief states and observations.

7.1 Yeast Cell Cycle Network

The results in the yeast cell cycle network demonstrate that A0* improves upon the Datta algorithm, and that the interventions chosen in the problems are supported by prior biological knowledge of the yeast network.

Scaling: Table IV demonstrates the runtime results of A0* and Datta algorithm (denoted by “D”) and the quality of the optimal intervention plan, in terms of the initial belief state’s value $\mathcal{V}(b_I)$ (expected reward). The table groups results for each problem into sets of three rows, and the columns denote the algorithms or value, the problem, and the result for a specific horizon. As expected, the enumerative algorithm does not scale well with the horizon, whereas A0* scales much better. There is one unsolvable problem, given the interventions provided (it is not possible to reach basin seven from attractor one).

The reason that A0* scales much better in this set of problems is that there are many belief states where all possible states satisfy the goal (acquire a non-zero reward) and allow early (prior to the maximum horizon) termination of plan expansion under the belief state (c.f. line 9, of the A0* algorithm). The enumerative algorithm has the same benefit of this truncated expansion, but will continue to expand other belief states that are reached by other, lower reward plans. This is the benefit of A0*: it always expands the belief states corresponding to the current highest reward partial plan, and high reward plans reach the goal (allowing such early termination). For example, the first problem (transitioning from attractor one to basin two) reveals a one intervention plan that achieves the goal in one step.

Biological Significance: As previously mentioned, with over eighty-six percent of the yeast cell cycle network state space covered by basin 1, planning interventions in

²The Partially-Observable Non-Deterministic planner.

	A→B	h = 1	h = 2	h = 3	h = 4	h = 5	h = 6	h = 7	h = 8	h = 9	h = 10
D	A1→B2	1.90	1.94	1.93	2.44	2.50	27.17	-	-	-	-
A0*	A1→B2	1.87	1.93	1.92	1.92	1.91	1.88	1.85	1.89	1.89	1.94
$\mathcal{V}(b_I)$	A1→B2	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00
D	A1→B3	1.93	1.93	1.98	2.55	2.52	40.34	-	-	-	-
A0*	A1→B3	1.91	1.86	1.92	1.95	1.93	1.89	1.93	1.94	1.93	1.94
$\mathcal{V}(b_I)$	A1→B3	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00
D	A1→B4	1.95	1.93	1.94	2.41	2.45	44.98	-	-	-	-
A0*	A1→B4	1.95	1.91	1.95	1.98	1.97	2.01	2.09	2.12	2.19	2.19
$\mathcal{V}(b_I)$	A1→B4	0.00	8.00	8.00	8.00	8.00	8.00	8.00	8.00	8.00	8.00
D	A1→B5	1.84	1.93	1.94	2.30	2.39	44.70	-	-	-	-
A0*	A1→B5	1.94	1.94	2.01	2.17	2.37	2.49	2.81	3.01	3.20	3.37
$\mathcal{V}(b_I)$	A1→B5	0.00	0.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00
D	A1→B6	1.87	1.85	1.97	2.50	2.47	40.34	-	-	-	-
A0*	A1→B6	1.85	1.86	1.87	1.93	1.96	1.93	1.91	1.94	1.90	1.90
$\mathcal{V}(b_I)$	A1→B6	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00
D	A1→B7	1.88	1.90	1.91	2.48	2.49	40.77	-	-	-	-
A0*	A1→B7	1.85	1.95	2.01	2.44	3.40	4.95	11.07	57.61	241.82	285.59
$\mathcal{V}(b_I)$	A1→B7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D	A2→B1	1.88	1.89	1.89	2.20	2.25	102.66	-	-	-	-
A0*	A2→B1	1.92	1.84	1.84	1.86	1.86	1.93	1.91	1.87	1.92	1.84
$\mathcal{V}(b_I)$	A2→B1	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00
D	A3→B1	1.90	1.92	1.92	2.25	2.22	182.66	-	-	-	-
A0*	A3→B1	1.93	1.89	1.92	1.93	1.94	1.93	1.92	1.91	1.92	1.90
$\mathcal{V}(b_I)$	A3→B1	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00
D	A4→B1	1.95	1.86	1.87	2.32	2.27	186.88	-	-	-	-
A0*	A4→B1	1.94	1.87	1.87	1.86	1.87	1.89	1.87	1.97	1.88	1.87
$\mathcal{V}(b_I)$	A4→B1	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00
D	A5→B1	1.88	1.87	1.93	2.49	2.46	-	-	-	-	-
A0*	A5→B1	1.89	1.88	1.91	1.87	1.88	1.93	1.93	1.90	1.96	1.86
$\mathcal{V}(b_I)$	A5→B1	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00
D	A6→B1	1.92	1.90	1.94	2.66	3.81	72.76	-	-	-	-
A0*	A6→B1	1.94	1.92	1.94	1.91	1.91	1.89	1.90	1.88	1.88	1.89
$\mathcal{V}(b_I)$	A6→B1	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00
D	A7→B1	1.93	1.93	1.93	2.66	3.34	98.02	-	-	-	-
A0*	A7→B1	1.95	1.88	1.96	1.88	1.84	1.92	1.95	1.95	1.89	1.86
$\mathcal{V}(b_I)$	A7→B1	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00

Table IV. Yeast Network: Total time and expected reward for A0* and Datta algorithms with increasing horizon.

this network was less a question of healthy vs. unhealthy, as in the aging network, but more a question of normal v. abnormal. As [Li et al. 2004] show, the entire cell cycle can be found within basin 1, and we aim to find out how to transition to and from basin 1 in our intervention plans. As shown in Table V, the variables selected for intervention include SBF, MBF, MCM1, Cln1,2, and Swi5. In our previous study [Verdicchio and Kim 2010], SBF and MBF were found to be variables with high popularity in basin 1. As such their selection is quite warranted. Furthermore, the selection of MCM1 makes great sense as MCM1 has relatively high popularity in basin 1. However, the selection of Cln1,2 and Swi5 are slightly more puzzling as

A→B	Interventions	Description
A1→B2	SBF	SBF has high popularity in basin 1
A1→B3	MBF	MBF has high popularity in basin 1
A1→B4	MCM1, Cln12	MCM1 has high popularity in basin 1, Cln12 has low popularity
A1→B5	MCM1, MBF, Cln12	See above
A1→B6	Cln12	See above
A1→B7		No solution through available interventions
A2→B1	MCM1	See above
A3→B1	SBF	See above
A4→B1	MCM1	See above
A5→B1	SBF	See above
A6→B1	MCM1	See above
A7→B1	Swi5	Swi5 has low scores

Table V. Summary of Yeast Intervention Plans

they had low popularity reported in the same study.

We see a trend in the selection of unpopular variables for intervention plans involving rare basins (i.e. basins 4-7, when combined total just over one percent of the total state space). This is because these unpopular variables separate the basins of such small size. In one case, namely intervening on basin 1 to get to basin 7, no solution was able to be found, even with the use of unpopular variables. For the converse case, from basin 7 to basin 1, we see the use of the low popularity Swi5 variable, as basin 7 is only one single state, to make the difference to basin 1. Finally, we see the widespread use of the high-scoring variables for intervention plans to and from the larger basins as those variables define the biggest differences between the basins.

As shown in Table V, the variables selected for intervention include SBF, MBF, MCM1, Cln1,2, and Swi5. In our previous study [Verdicchio and Kim 2010] SBF and MBF were found to be variables with high popularity in basin 1. As such their selection is warranted mathematically. Nonetheless, a recent report states that SFB and MBF are driving the transition of yeast cell cycle [Ferrezuelo et al. 2010], as predicted by our intervention plan. We note that our prediction was made prior to Ferrezuelo et al.’s publication.

7.2 Aging Network

The experiments in the aging network demonstrate the scalability of AO^* compared to the enumerative algorithm and have intervention plans of biological significance.

Scaling: The algorithms demonstrate scaling similar to the yeast network in the aging network. The aging network problems are somewhat more difficult to solve because the initial belief state is consistent with multiple states in all problems, but the network is also deterministic. Thus, the intervention plans include branching over observations.

An interest aspect of the plans in the aging network, is that as the horizon increases, it is possible to find higher reward plans. For example, in transitioning from attractor 1 to basin 3, only at horizon ten is it possible to increase the reward by using fewer interventions; our interpretation of this result is that waiting, and

	A→B	h = 1	h = 2	h = 3	h = 4	h = 5	h = 6	h = 7	h = 8	h = 9	h = 10
D	A1→B2	0.32	0.36	0.38	2.95	3.02	-	-	-	-	-
A0*	A1→B2	0.31	0.35	0.36	0.38	0.38	0.36	0.37	0.36	0.35	0.35
$\mathcal{V}(b_I)$	A1→B2	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00
D	A1→B3	0.31	0.33	0.38	2.96	2.91	-	-	-	-	-
A0*	A1→B3	0.31	0.37	0.68	1.38	2.06	2.83	4.00	4.91	5.58	1.93
$\mathcal{V}(b_I)$	A1→B3	0.00	0.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	8.00
D	A2→B1	0.34	0.74	2.96	286.54	-	-	-	-	-	-
A0*	A2→B1	0.31	0.62	0.60	0.69	0.66	0.75	0.77	0.72	0.82	0.81
$\mathcal{V}(b_I)$	A2→B1	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00
D	A2→B3	0.31	0.68	2.99	280.31	-	-	-	-	-	-
A0*	A2→B3	0.33	0.77	3.40	9.82	26.71	97.01	-	-	-	-
$\mathcal{V}(b_I)$	A2→B3	0.11	1.88	5.22	7.22	7.77	8.22	-	-	-	-
D	A3→B1	0.29	0.47	1.81	-	-	-	-	-	-	-
A0*	A3→B1	0.33	0.39	0.39	0.39	0.41	0.39	0.41	0.45	0.44	0.43
$\mathcal{V}(b_I)$	A3→B1	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00
D	A3→B2	0.33	0.47	1.82	-	-	-	-	-	-	-
A0*	A3→B2	0.33	0.43	0.70	0.51	0.47	0.53	0.51	0.50	0.53	0.53
$\mathcal{V}(b_I)$	A3→B2	5.66	8.00	8.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00

Table VI. Aging Network: Total time and expected reward for A0* and Datta algorithms with increasing horizon.

A→B	Interventions Used	Description
A1→B2	dietantx	The lack of dietary antioxidants has complete popularity in Basin 1.
A1→B3	dietantx, ¬ros	See A1→B2 regarding dietantx. ROS has complete popularity in Basin 3.
A2→B1	¬dietantx	See A1→B2, involves opposite intervention.
A2→B3	ros, dietantx, mtdna, mtprot	See text.
A3→B1	¬dietantx	See A2→B1.
A3→B2	¬ros or ros	See A1→B3 and A2→B3.

Table VII. Description of Aging Intervention Plans for plan with greatest horizon.

allowing the network to change without intervention, provides the opportunity to strategically plan the interventions. In transitioning from attractor 2 to basin 3, the increasing horizon allows more of the possible initial states (source attractor states) to reach the destination basin.

Biological Significance: In investigating all pairings of attractors and basins for planning interventions, we are able to create intervention plans which correspond to natural aging and deterioration. As mentioned, basin 1 is a collection of states which all lead to an attractor representing the worst case scenario health-wise. However, one characteristic of this basin was that every one of the states had the variable for dietary antioxidants set to false; this fact was capitalized upon in all problems involving basin 1. While our previous study [Verdicchio and Kim 2010] showed that the importance of dietary antioxidants was elevated by certain network topology characteristics, its importance was nonetheless recognized and utilized in the formulation of the intervention plans. In addition to dietary antioxidants, ROS was also suppressed in the intervention plan to get from basin 1 to basin 3. Since

	Inst.	Y	h = 1	h = 2	h = 3	h = 4	h = 5	h = 6	h = 7	h = 8	h = 9	h = 10
D	MPW ⁷	-10	2.01	2.65	4.84	15.23	66.09	311.32	-	-	-	-
AO*	MPW ⁷	-10	1.95	2.21	2.67	3.02	3.82	4.95	6.96	11.16	17.44	26.54
V(b _I)	MPW ⁷	-10	-2.56	-0.45	-0.84	-0.17	-0.06	-0.06	-0.01	0.00	0.00	0.00
D	MPW	10	1.97	2.53	4.69	14.13	62.45	317.96	-	-	-	-
AO*	MPW	10	1.98	2.53	4.80	14.34	59.98	315.31	-	-	-	-
V(b _I)	MPW	10	7.43	9.59	9.28	9.88	9.85	9.93	-	-	-	-
D	PW ⁷	-10	2.07	2.33	3.06	5.03	10.32	26.91	76.18	249.91	784.79	-
AO*	PW ⁷	-10	2.12	2.41	3.12	5.30	6.48	30.33	86.32	261.72	845.69	-
V(b _I)	PW ⁷	-10	-7.43	-9.59	-9.22	-9.53	-2.00	-9.76	-9.82	-9.84	-9.83	-
D	PW	10	2.44	2.76	3.67	6.54	14.05	37.68	114.00	337.31	1168.81	-
AO*	PW	10	2.48	2.88	3.86	6.38	14.16	37.27	111.91	349.05	1180.82	-
V(b _I)	PW	10	2.56	0.40	0.84	0.65	0.87	0.33	0.32	0.19	0.20	-
D	WP ⁷	-10	1.65	2.02	3.00	5.12	12.06	33.13	109.23	410.75	-	-
AO*	WP ⁷	-10	1.67	2.02	2.65	3.91	5.77	9.80	17.69	31.67	59.27	107.48
V(b _I)	WP ⁷	-10	-7.43	-3.29	-1.36	-1.71	-1.11	-1.15	-1.05	-1.03	-1.01	-1.01
D	WP	10	2.23	2.57	3.53	6.08	11.85	31.04	83.61	211.27	533.67	-
AO*	WP	10	2.06	2.42	3.66	6.54	12.17	28.45	56.43	336.96	998.48	-
V(b _I)	WP	10	2.56	6.71	8.59	8.11	7.00	7.00	7.00	8.93	8.94	-

Table VIII. Wnt5a Network: Total time and expected reward for AO* and Datta algorithms with increasing horizon.

this is an intervention from the larger unhealthy basin to a smaller, healthier one, a more exact plan was required; that plan utilized the variable with the highest popularity in the target basin, namely, ROS. Biologically this also makes sense, as ROS and antioxidants are found to be at cross purposes within the cell: as ROS are damaging cellular entities which accumulate with age and deterioration, dietary antioxidants work to neutralize their effects. Thus, it is quite biologically relevant to see these two variables used in various combinations to intervene both as treatment and as aging effects.

To transition between basins 2 and 3, which represented differing degrees of health, more creative intervention plans needed to be devised. First, consider the transition from basin 3 (less healthy) to basin 2 (more healthy). In basin 3, ROS was the variable recognized again as a critical intervention target; it had complete variable popularity in basin 3, and was suppressed or activated to make those specific adjustments needed to get to basin 2, depending on the starting state. Second, we discuss the intervention to get from basin 2 to basin 3. Because this intervention moved from a larger basin to a smaller one, and because it moved to a mid-range health status, the interventions were more complex. A total of five plan branches were generated, and each involved the aforementioned ROS, but also mitochondrial DNA (mtdna) in both suppressing and activating interventions, and mitochondrial proteins (mtprot) in suppressing interventions. Both of these new variables carried proportional popularity in both basins 2 and 3, so their selection is to be expected mathematically, but biologically they also control the health of the mitochondria and its DNA. By controlling either oxidative stress or DNA health in different branches, the plan transitioned between basins 2 and 3.

7.3 Wnt5a Network

The Wnt5a network is markedly different from the other networks in that it is a probabilistic Boolean network (there is more than one predictor per gene), and does not benefit as much from early truncation due to achieving the goal in all possible states. We experiment with six variations of the problem of making Wnt5a inactive by using different interventions and observable variables, and by using different reward functions. There are three configurations of the interventions and observable variables, either intervene by suppressing Wnt5a and observe Pirin (WP), intervene by suppressing Pirin and observing Wnt5a (PW), or intervene by suppressing Mart-1 or Pirin and observing Wnt5a (MPW). There are two possible goals, either a positive ten reward for suppressing Wnt5a (WP, PW, and MPW), or a negative ten reward for activating Wnt5a (WP', PW', and MPW'). While the two goals are related, they affect planner performance by changing the upper bound on the reward function. The significance of the two goal formulations is that both place bias on plans that suppress Wnt5a, but the latter leads to more favorable pruning in the $A0^*$ search (as discussed in Section 6) – knowing which formulation is better can aid practitioners during problem formulation.

Scaling: Table VIII lists results similar to the aging and yeast networks, but with an additional column to indicate the goal reward. When the goal reward is ten, $A0^*$ performs worse than the enumerative algorithm because the upper bound on the reward is very loose (compared to the optimal value found), forcing the search to evaluate nearly the entire search space searching for a better solution. The repeated belief state evaluation method used by $A0^*$ leads to more belief state evaluations than the enumerative method (which evaluates each belief state once), and hence larger search times. When the reward is negative ten (and the goal is to activate Wnt5a) $A0^*$ performs better in all problems except PW' because in both PW and PW' it is difficult to control Wnt5a by intervening with Pirin. However, by intervening to suppress both Mart-1 and Pirin it is much easier to control Wnt5a (leading to higher value plans).

Biological Significance: Like the yeast cell cycle set of variables, the set of variables in the Wnt5a network have been selected carefully from a much larger list [Kim et al. 2002], and so the inclusion of any one in a particular plan is biologically sound. Even as such, as the allowable interventions and observations are not only based upon existing biological knowledge of the Wnt5a signaling pathway, but also behave in accordance with it, we are able to justify each plan as biologically significant. We saw that allowing both Pirin and Mart-1 interventions improved the chance of suppressing Wnt5a over the use of Pirin alone, which fits with how Pirin and Mart-1 involved in a high CoD predictor for Wnt5a.

8. RELATED WORK

Prior work related to our approach can be categorized into techniques for reasoning about and simulating biological systems as Boolean networks, approaches for controlling such biological systems, and algorithms within Artificial Intelligence.

Boolean Networks: Studies of genetic regulatory networks can take many forms; some attempt to merely determine associative or predictive relations between genes,

others seek to model network dynamics to the smallest biological nuance. A number of models have been proposed to study gene regulatory networks [de Jong 2002]. The complexity ranges from simplistic models such as Boolean networks [Kauffman 1969b; Kauffman 1993] to more complete and intricate models based on differential equations [Goutsias and Kim 2004; Goutsias and Kim 2006]. Evaluation of these models is done, as in this work, by examining how accurately the model reflects actual biology and the difficulty of inference in a specific instance of the model [de Jong 2002]. In general, the more low-level detail that is added, the more difficult to reconstruct the network from data.

Studies have shown that Boolean network models exhibit a number of biologically interesting properties [Kauffman 1969b]. Boolean networks primarily focus on determining gene-gene interactions at a qualitative level, instead of quantitative aspects. Additionally, Boolean networks can provide insight into cellular states. Both the steady-state and switch-like behavior of cells can be captured and studied with a Boolean model [Kim et al. 2000]. The ability to model both of these behaviors allows the analysis of common functions of the cell such as cell growth and cell cycle, as well as the response to external stimuli.

Beyond Boolean models, models can add a stochastic component, such as in PBNs, and Bayesian networks [Murphy and Mian 1999; Friedman et al. 2000; Hartemink et al. 2001]. The addition of the stochastic component attempts to model both intrinsic and extrinsic noise in gene regulatory networks [Elowitz et al. 2002]. In probabilistic Boolean networks, as in traditional Boolean networks, Boolean functions are used to determine the next state of the network. Unlike traditional Boolean networks, however, there is not a single function corresponding to a network state. The next state of the network is determined via selection of Boolean function from a set of valid Boolean functions based on the current state. PBNs have been studied using Markov chains and shown to demonstrate both homeostasis and the switch-like properties exhibited by actual biological systems [Kim et al. 2002].

Controlling Biological Systems: Planning interventions in biological systems has been previously studied within the context of control theory, specifically, controlling Markov chains. AI planning and control theory have rich connections, and these similarities provide a common ground on which the empirical evaluations described herein are based. The primary approach of previous work [Datta et al. 2003; Datta et al. 2004] on formulating intervention planning as a control problem was to characterize a dynamic programming operator that could identify the intervention to make in any state of the biological process. In this formalism, Datta et al. [Datta et al. 2003] formulate the problem as a fully observable Markov decision process and explore finite and infinite horizon control. Datta et al. [Datta et al. 2004] also explore an extension to partial observability with finite horizon control. In both works, the underlying dynamics of the networks is based on probabilistic Boolean networks and the techniques are evaluated within the Wnt5a network.

Artificial Intelligence Approaches: Some recent works in the AI community have focused on simply representing biological processes. Khan et al. [Khan et al. 2003] seek to discover signal transduction pathways with a deterministic classical planner. The actions in the plan represent various chemical reactions, and the goal of the plan is to establish that there is a sequence of actions leading to an event, such as

transcription. This particular planning problem has received considerable attention due to its inclusion in a recent International Planning Competition. The most significant differences from the work described in this paper, are that the model is assumed deterministic, is at a finer level of granularity (modeling many cellular products), and is primarily concerned with modeling the problem (versus exploring appropriate solution techniques).

Further along this vein of improved models of change in biological processes, Tran and Baral [Tran and Baral 2005] model change in biological processes as exogenous actions, termed triggers. In the planning considered in this work, the plans contain actions that represent both intervention and non-intervention, where the actions respectively model how the GRN changes under outside influence or naturally. Tran and Baral factor out the model of natural change in the biological process, representing it as triggers, which plans can only indirectly affect. Actions become much simpler, in that they describe only the change due to the intervention.

9. CONCLUSION & FUTURE WORK

We have presented a formulation of biological network intervention as decision theoretic planning. Our planner relies on several AI planning tools to perform efficient reasoning. As a result, we have improved the scalability of planning interventions over previous work.

We have shown that $A0^*$ is a viable approach to solving biological intervention problems. Where the Datta algorithm quickly exceeds our time limit as the horizon increases, $A0^*$ can scale to larger horizons. Using three biological networks of practical interest, we have seen that $A0^*$ performs well by pruning vertices based on upper bounds. Each of the networks studied led to biologically supported intervention plans that may have been otherwise challenging to construct manually, supporting our claim that automated planning can supplement a biologist's understanding of a system. The yeast network plans led to an interesting prediction of an intervention target that was independently verified after our experiments and prior to the publication of this work.

In the future, we would like to decouple the representation of cellular process dynamics from our intervention actions. We think the right approach is to use exogenous actions or processes. While recent work in deterministic planning [McDermott 2003] has discussed models for controlling exogenous processes and triggers [Tran and Baral 2005], less thought has been given to such processes in a probabilistic setting (with the exception of [Blythe 1994]). Alternatively, work on simulating cellular processes [Ramsey et al. 2005] has considered probabilistic exogenous processes, but not under the control of a planner.

Acknowledgements: This work was supported in part by NIH grants 3R21-LM009706-02W1 and 5R21-LM009706-02, and DARPA contract HR001-07-C-0060. We thank Ashish Choudhary and Edward Dougherty for their helpful suggestions while preparing this work.

REFERENCES

- ALBERT, R. AND OTHMER, H. G. 2003. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in drosophila melanogaster. *J Theor Biol* 223, 1, 1–18.

- BITTNER, M. ET AL. 2000. Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature* 406, 536–40.
- BLYTHE, J. 1994. Planning with external events. In *Proceedings of the Tenth Conference on Uncertainty in Artificial Intelligence*. 94–101.
- BRYANT, R. 1986. Graph-based algorithms for Boolean function manipulation. *IEEE Trans. Comput. C-35*, 8 (Aug.), 677–691.
- BRYCE, D., KAMBHAMPATI, S., AND SMITH, D. 2006. Planning graph heuristics for belief space search. *Journal of AI Research* 26, 35–99.
- BRYCE, D. AND KIM, S. 2007. Planning for gene regulatory network intervention. In *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence*. 1834–1839.
- DATTA, A., CHOUDHARY, A., BITTNER, M., AND DOUGHERTY, E. 2004. External control in Markovian genetic regulatory networks: the imperfect information case. *Bioinformatics* 20, 6, 924–930.
- DATTA, A., CHOUDHARY, A., AND DOUGHERTY, E. 2003. External control in Markovian genetic regulatory networks. *Machine Learning* 52, 52, 169–191.
- DAVIDICH, M. I. AND BORNHOLDT, S. 2008. Boolean network model predicts cell cycle sequence of fission yeast. *PLoS ONE* 3, 2, e1672.
- DE JONG, H. 2002. Modeling and simulation of genetic regulatory systems: A literature review. *Journal of Computational Biology* 9, 1, 67–103.
- DOUGHERTY, E. R., BRUN, M., TRENT, J. M., AND BITTNER, M. L. 2009. Conditioning-based modeling of contextual genomic regulation. *IEEE/ACM Trans. Comput. Biol. Bioinformatics* 6, 2, 310–320.
- ELOWITZ, M., LEVINE, A., SIGGIA, E., AND SWAIN, P. 2002. Stochastic gene expression in a single cell. *Science* 297, 1183–1186.
- FERREZUELO, F., COLOMINA, N., FUTCHER, B., AND ALDEA, M. 2010. The transcriptional network activated by cln3 cyclin at the g1-to-s transition of the yeast cell cycle. *Genome Biology* 11, 6, R67.
- FRIEDMAN, N., LINIAL, M., NACHMAN, I., AND PEÉR, D. 2000. Using Bayesian networks to analyze expression data. *Journal of Computational Biology* 7, 601–620.
- FURBER, J. D. 2010. Systems biology of human aging: Network model of biochemical and physiological interactions in human senescence.
- GOUTSIAS, J. AND KIM, S. 2004. A nonlinear discrete dynamical model for transcriptional regulation: Construction and properties. *Biophysics Journal* 86, 1922–1945.
- GOUTSIAS, J. AND KIM, S. 2006. Stochastic transcriptional regulatory systems with time delays: A mean-field approximation. *Journal of Computational Biology* 13, 1049–1076.
- HARTEMINK, A., GIFFORD, D., JAAKKOLA, T., AND YOUNG, R. 2001. Using graphical models and genomic expression data to statistically validate models of genetic regulatory networks. In *Pacific Symposium on Biocomputing*. 422–433.
- KAELBLING, L. P., LITTMAN, M. L., AND CASSANDRA, A. R. 1995. Planning and acting in partially observable stochastic domains. *Artificial Intelligence* 101, 99–134.
- KAUFFMAN, S. 1969a. Homeostasis and differentiation in random genetic control networks. *Nature* 224, 215, 177–8.
- KAUFFMAN, S. 1969b. Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology* 22, 437–467.
- KAUFFMAN, S. 1993. *The origins of order: Self-organization and selection in evolution*. New York: Oxford University Press.
- KHAN, S., DECKER, K., GILLIS, W., AND SCHMIDT, C. 2003. A multi-agent system-driven AI planning approach to biological pathway discovery. In *Proceedings of Thirteenth International Conference on Automated Planning and Scheduling*.
- KIM, S., DOUGHERTY, E., BITTNER, M., CHEN, Y., SIVAKUMAR, K., MELTZER, P., AND TRENT, J. 2000. Multivariate measurement of gene expression relationships. *Genomics* 67, 2, 201–209.
- KIM, S., LI, H., DOUGHERTY, E., CAO, N., CHEN, Y., BITTNER, M., AND SUH, E. 2002. Can Markov chain models mimic biological regulation? *J. of Biological Systems* 10, 4, 337–357.

- LI, F., LONG, T., LU, Y., OUYANG, Q., AND TANG, C. 2004. The yeast cell-cycle network is robustly designed. *Proceedings of the National Academy of Sciences of the United States of America* 101, 14 (April), 4781–4786.
- MCDERMOTT, D. 2003. The formal semantics of processes in PDDL. ICAPS Workshop on PDDL.
- MURPHY, K. AND MIAN, S. 1999. Modeling gene expression data using dynamic Bayesian networks. Tech. rep., University of California, Berkeley.
- MURPHY, K. P. 2002. Dynamic bayesian networks: Representation, inference and learning. Ph.D. thesis, UC Berkeley.
- NILSSON, N. 1980. *Principles of Artificial Intelligence*. Morgan Kaufmann.
- PAL, R., DATTA, A., BITTNER, M. L., AND DOUGHERTY, E. R. 2005. Intervention in context-sensitive probabilistic Boolean networks. *Bioinformatics* 21, 7, 1211–1218.
- PINEAU, J., GORDON, G. J., AND THRUN, S. 2006. Anytime point-based approximations for large POMDPs. *J. Artif. Intell. Res. (JAIR)* 27, 335–380.
- RAMSEY, S., ORRELL, D., AND BOLOURI, H. 2005. Dizzy: Stochastic simulation of large-scale genetic regulatory networks. *J. Bioinformatics Comput. Biol* 3, 2, 415–436.
- SHMULEVICH, I., DOUGHERTY, E., KIM, S., AND ZHANG, W. 2002. Probabilistic Boolean networks: a rule-based uncertainty model for gene regulatory networks. *Bioinformatics* 18, 2, 261–274.
- SHMULEVICH, I., DOUGHERTY, E., AND ZHANG, W. 2002. Gene perturbation and intervention in probabilistic boolean networks. *Bioinformatics* 18, 10, 1319–1331.
- SMITH, T. AND SIMMONS, R. G. 2004. Heuristic search value iteration for POMDPs. In *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence*. 520–527.
- SOSMAN, J., WEERARATNA, A. T., AND SONDAK, V. K. 2004. When will melanoma vaccines be proven effective? *Journal of Clinical Oncology* 22, 387–389.
- TRAN, N. AND BARAL, C. 2005. Issues in reasoning about interaction networks in cells: necessity of event ordering knowledge. In *Proceedings of the Twentieth National Conference on Artificial Intelligence*.
- VERDICCHIO, M. AND KIM, S. 2010. Reduction of Boolean network basins of attraction reveals intervention targets. Tech. Rep. TR-10-004, Arizona State University, Tempe, Arizona. April.
- WEERARATNA, A. T., JIANG, Y., HOSTETTER, G., ROSENBLATT, K., DURAY, P., BITTNER, M., AND TRENT, J. M. 2002. Wnt5a signaling directly affects cell motility and invasion of metastatic melanoma. *Cancer Cell* 1, 3, 279–88.
- WUENSCH, A. 1998. Genomic regulation modeled as a network with basins of attraction. In *Pacific Symposium on Biocomputing*. 89–102.
- XIAO, Y. 2009. A tutorial on analysis and simulation of boolean gene regulatory network models. *Current Genomics* 10, 7 (November), 511–525.
- YOUNES, H. L. S., LITTMAN, M. L., WEISSMAN, D., AND ASMUTH, J. 2005. The first probabilistic track of the international planning competition. *J. Artif. Intell. Res. (JAIR)* 24, 851–887.