Issues in Reasoning about Interaction Networks in Cells: Necessity of Event Ordering Knowledge

Nam Tran and Chitta Baral
Computer Science & Engineering
Arizona State University
Tempe, AZ 85281, USA
{namtran,chitta}@asu.edu

Carron Shankland
Computing Science & Mathematics
University of Stirling
Stirling FK9 4LA, UK
ces@cs.stir.ac.uk

Abstract

In this paper we discuss several representation issues that we came across while modelling molecular interactions in cells of living organisms. One of the issues was that the triggering of events inside cells, an important modelling component, are not necessarily immediate, leading to multiple evolution models in the absence of additional information. Second, often an action or a trigger at one level of granularity of representation can be elaborated and refined. We show the problem that existing representation and modelling formalisms have in dealing with the above issues. We then present an action language which builds up on a previous language, and has the ability to express event ordering knowledge. We show that our language is able to adequately address the above-mentioned issues.

Motivation and Introduction

In this paper our goal is to use the action language approach for reasoning about actions to specify knowledge about various interactions inside cells of living beings, and to reason about these interactions. Inside the cell various interactions take place between genes, proteins and other biochemical molecules. These interactions influence most visible properties of cells and tissues, such as cells dying, cells proliferating, and cells becoming cancerous.

As done in (Tran & Baral 2004; Baral *et al.* 2004), the interactions can be modelled to some extent as triggered actions. A theory of the cell specifies effects of actions and how actions are triggered or inhibited. Such a specification dictates how the cell evolves (i.e. changes through time). An evolution of the cell starts from a state which triggers certain actions; these actions change the state, which may trigger further actions; and so on.

In this paper we delve deeper into the interaction process. To start with, unlike controllers in artificial domains such as softbots, or robots, triggers inside the cell do not necessarily fire immediately. For example, if a cell specification has the trigger "f **triggers** a" and the trigger "f **triggers** b", it is not always the case that in a state where f is true both a and b immediately occur. One does not have the luxury of specifying exactly how long it takes before a trigger fires, as that information is usually not available and it may not even be a deterministic duration. In some cases, there is insufficient knowledge for one to say with certainty which of a and b occurs before the other. This is an example of a common phenomenon in interaction networks called *specificity*: a network can respond in different ways to the same input (Tan & Kim 1999). In other cases, some form of ordering between triggered action occurrences is known. Thus a representation and reasoning mechanism should model specificity as well as model event ordering information when appropriate.

A related issue in modelling cell behavior is that one can (and often needs to) model at different granularities. As more details are known, an action or trigger at one level of granularity can be shown to consist of triggers and actions at a lower level of granularity. In this case, it is important that the reasoning mechanism, which has been able to reason correctly by incorporating observations about the cell behavior, does not make mistakes when an action or a trigger is replaced by its finer decomposition.

The following hypothetical example illustrates some of the above points and shows the problem that one has without using event ordering information, as done in (Tran & Baral 2004) and in almost all other approaches (Reddy, Liebman, & Mavrovouniotis 1996; Peleg, Yeh, & Altman 2002; Regev, Silverman, & Shapiro 2001; Giordano, Martelli, & Schwind 2001; Khan *et al.* 2003; Talcott *et al.* 2004; Chabrier *et al.* 2004; Giunchiglia *et al.* 2004; Calvanese, de Giacomo, & Vardi. 2002; Fukuda & Takagi 2001; Reiter 1996).

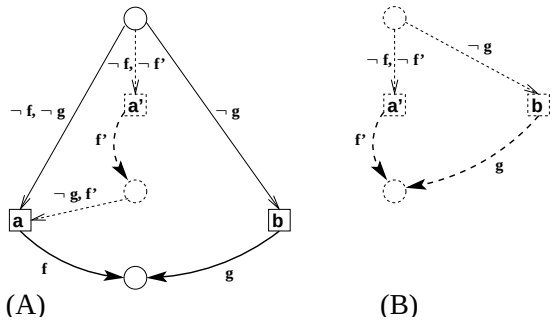

Figure 1: Different interpretations of the same process. (A) Original interpretation, regardless the details of $\neg f, \neg g$ **triggers** a in dashed lines. (B) New interpretation, considering the details.

Consider a cell specification in the language $\mathcal{A}_T^0$ (Tran & Baral 2004) given as follows.

$$\mathcal{D}_{afg} = \{\neg f, \neg g \textbf{ triggers } a; \quad a \textbf{ causes } f; \\ \neg g \textbf{ triggers } b; \quad b \textbf{ causes } g\}$$

The semantics of $\mathcal{A}_T^0$ dictates that if an action is triggered at time t then it has to occur at time t. Assume that f and g are false at time 0, then both a and b occur at time 0. The occurrence of a makes f *become true* at time 1. Now imagine that the trigger $\neg f, \neg g$ **triggers** a is refined by the following more elaborated knowledge:

$$\neg f\ , \neg f' \textbf{ triggers } a' \quad (1)$$

$$a' \textbf{ causes } f' \quad (2)$$

$$f', \neg g \textbf{ triggers } a \quad (3)$$

That is, the triggering of the action a by $\neg f \wedge \neg g$ is mediated by some new action a' and fluent f' (dashed lines in Fig. 1A). Given the same initial condition of f and g being false at time 0, a different conclusion about the final value of f would be drawn from the refined interaction specification (Fig. 1B). The argument goes as follows. If f' is true at time 0, then a' (hence, a) will not be triggered; thus f remains false. On the other hand, if f' is false at time 0 then the actions a' and b occur at time 0, which results in $\neg f$ and g at time 1. Then no additional action can be triggered from time 1 onwards, so f *is always false*. Thus the new conclusion (about f) contradicts the previous prediction from the old specification $\mathcal{D}_{afg}$. Intuitively, this non-monotonicity with respect to refinement is not acceptable. Since the detailed knowledge represented by the statements (1)-(3) is about what happens "inside" the trigger of a, the consequence of the trigger itself should not be affected, provided everything "outside" remains intact.

The above example illustrates a challenging elaboration tolerance issue in representing interaction networks, as often new details about various intracellular mechanisms emerge, and the modelling and reasoning mechanism should be able to accommodate the refinement without breaking down.

The rest of the paper is organized as follows. In the next section we will discuss a real example of molecular interactions that motivated our work in this paper. We then propose a language $\mathcal{A}_T^\infty$ that can gracefully address the modelling problems posed by this real example. Finally, we present formal results on elaboration tolerance of $\mathcal{A}_T^\infty$, which generalizes all the special cases discussed in the paper.

Challenges in Modelling the p53 Network

We came across this example when we were modelling interactions of the p53 protein network. The example poses the following modelling challenges: (1) elaboration of more refined knowledge that is correct with respect to reasoning; (2) representing and reasoning about non-deterministic cellular behaviors called *specificity*; and (3) incorporation of event ordering knowledge into representing and reasoning about interactions in cells.

First, let us review the special case of $\mathcal{A}_T^0$ language that does not involve exogenous actions (Tran & Baral 2004).

Brief overview of $\mathcal{A}_T^0$

The alphabet of the language $\mathcal{A}_T^0$ (without exogenous actions) consists of two nonempty disjoint sets of symbols: a set of *actions* and a set of *fluents*. A fluent is a propositional symbol. A fluent literal is a fluent or its negation. A set of fluent literals is *consistent* if it does not contain both a fluent and its negation. A state is an interpretation of the fluents (i.e. a maximal consistent set of fluent literals).

A *domain description* $\mathcal{D}$ in $\mathcal{A}_T^0$ is a set of statements of the following forms:

$$a \textbf{ causes } f \textbf{ if } f_1, \ldots, f_n \quad (4)$$

$$g_1, \ldots, g_m \textbf{ triggers } b \quad (5)$$

$$h_1, \ldots, h_l \textbf{ inhibits } c \quad (6)$$

where f_i, g_j, h_k and f and g are fluent literals and a, b, c are individual actions. (4) is a *dynamic causal rule*, which says that f is to be true in the state succeeding a state where a occurs and all the $f_1, \ldots, f_n$ hold. (5) is a *trigger* rule, which says that action b is to occur if it is not inhibited and if all the literals $g_1, \ldots, g_m$ hold. (6) is an *inhibition* rule, which says that action c can not happen (i.e. is *inhibited*) if all the literals $h_1, \ldots, h_l$ hold.

A *state transition* is a change of one state to another state due to effects of some actions. The effect of an action a in a state s is the set $E(a, s) = \{\ f \mid a \textbf{ causes } f \textbf{ if } f_1, \ldots f_n \in \mathcal{D}$ and $\{f_1, \ldots, f_n\} \subseteq s\ \}$. The effect of a set A of actions in a state s is the set $E(A, s) = \bigcup_{a \in A} E(a, s)$. Let $\neg\neg g = g$ and $\neg E(A, s) = \{\neg g | g \in E(A, s)\}$. A set A of actions transforms a state s to the state $\Phi(A, s)$ defined as:

- $\Phi(\emptyset, s) = s$;
- if $A \neq \emptyset$ and $E(A, s)$ is consistent, then
  $\Phi(A, s) = (s \setminus \neg E(A, s)) \cup E(A, s)$;
- otherwise $\Phi(A, s)$ is undefined.

The function Φ is also called the transition function of $\mathcal{D}$.

A *transition sequence* τ is a sequence of the form $\tau = \langle s_0, A_0, s_1, A_1, \ldots\rangle$; where s_i's are states and A_j's are sets of actions in $\mathcal{D}$, such that $s_{i+1} = \Phi(A_i, s_i)$ for all i. Throughout the rest of the paper we will use the notation that $\mathsf{s}_i(\tau) = s_i$ and $\mathsf{A}_i(\tau) = A_i$. Moreover, if τ is such that $A_i \neq \emptyset, A_j = \emptyset$ for all $j > i$, then we write $\tau = \langle s_0, A_0, s_1, A_1 \ldots A_i, s_i\rangle$.

A *trajectory* is a transition sequence τ where $\mathsf{A}_i(\tau)$ is the set of all the actions that are triggered but not inhibited in state $\mathsf{s}_i(\tau)$. *Observations* are statements of the form " f **at** i " or of the form " a **occurs_at** j ", where i and j are non-negative integers. The former statement means that the fluent literal f is observed to be true at time i. The latter means that the action a is observed to occur at time j. A trajectory τ satisfies " f **at** i " iff $f \in \mathsf{s}_i(\tau)$; and τ satisfies " a **occurs_at** j " iff $a \in \mathsf{A}_j(\tau)$.

A *theory* is a pair $(\mathcal{D}, \mathcal{O})$ where $\mathcal{D}$ is a domain description and $\mathcal{O}$ is a set of observations. A *model* of a theory $(\mathcal{D}, \mathcal{O})$ is a trajectory of $\mathcal{D}$ that satisfies all the observations of $\mathcal{O}$ and is minimal with respect to the following partial order $\leq_0$.

Definition 1. Let τ and τ' be trajectories where $\mathsf{s}_0(\tau) = \mathsf{s}_0(\tau')$. Then $\tau \leq_0 \tau'$ iff there exists a sequence $0 \leq i_0 < i_1 < \ldots < i_n < \ldots$ such that for every $k \geq 0$, $\mathsf{A}_k(\tau) \subseteq \mathsf{A}_{i_k}(\tau')$. □

We are now ready to discuss the interaction network of the p53 protein. In the network description , "adjective" predicates denote fluent symbols and "verb" predicates denote action symbols.

Issues in modelling the p53 network

The p53 protein is a tumor suppressor that plays a key role in the regulation of the cell growth and cell death. It is estimated that about one half of human cancers contain mutant p53. It is also predicted that the p53 network is affected in the majority of the remaining tumors. Normally, a wild-type (i.e. not mutant) p53 functions to prevent cancer as follows (the solid lines in Fig. 2). Stimuli such as UV, ionizing radiation or chemical carcinogens can induce DNA damage. DNA damage will lead to genomic instability, which in turn triggers uncontrolled cell growth (i.e. tumor formation). However, the stimuli also upregulate the gene expression of wild-type p53. The upregulated gene expression produces high levels of p53 concentration, which suppresses abnormal cell growth thus preventing cancer (Frei 2003).

The initial knowledge base The p53 network can be represented in $\mathcal{A}_T^0$ as follows.

$\mathcal{D}_{p53} = \{\ high(UV), \neg unstable(cells)$ **triggers** $damage(DNA)$
$damage(DNA)$ **causes** $unstable(cells)$
$unstable(cells), \neg tumorous$ **triggers** *proliferate*$(cells)$
proliferate$(cells)$ **causes** $tumorous$
$high(UV), \neg high(p53)$ **triggers** $upregulate(p53)$
$upregulate(p53)$ **causes** $high(p53)$
$high(53)$ **inhibits** *proliferate*$(cells)\}$

Given that only $high(UV)$ is true at time 0, actions $damage(DNA)$ and $upregulate(p53)$ will occur at time 0. Then $unstable(cells)$ becomes true at time 1, which triggers *proliferate*$(cells)$. However, this action cannot happen, because it is inhibited by $high(p53)$ at the same time. As expected, the domain $\mathcal{D}_{p53}$ predicts the p53 prevention of cancer (Fig. 2).

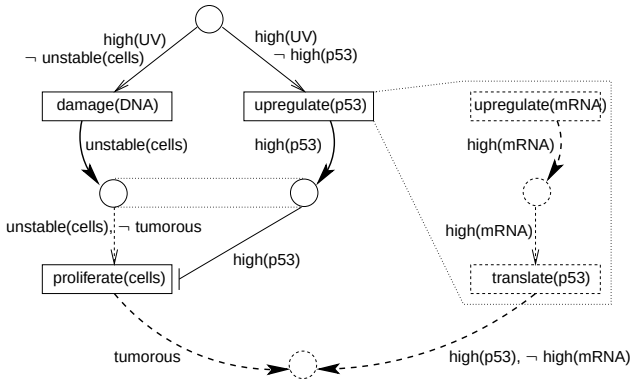

Figure 2: Detailed knowledge of p53 upregulation predicts cancer.

Elaboration of more detailed knowledge Each of the actions $damage(DNA)$, *proliferate*$(cells)$ and $upregulate(p53)$ represents a complex biological process. For example, let us update the network with the following refinement of the process of p53 upregulation (Frei 2003). The process first starts with the upregulation of the p53 mRNA, which causes a high level of mRNA. The high mRNA level induces translation of p53, which increases the quantity of p53 protein. Thus, to specify at a finer level, we replace the trigger

$high(UV), \neg high(p53)$ **triggers** $upregulate(p53)$

with the trigger

$high(UV), \neg high(p53)$ **triggers** $upregulate(mRNA)$

We then replace the causal rule

$upregulate(p53)$ **causes** $high(p53)$

with the following:

$upregulate(mRNA)$ **causes** $high(mRNA)$
$high(mRNA)$ **triggers** $translate(p53)$
$translate(p53)$ **causes** $high(p53), \neg high(mRNA)$

(The refinement is drawn in dashed lines in Fig. 2).

But unlike $\mathcal{D}_{p53}$, the updated domain description, which we refer to as $\mathcal{D}^+_{p53}$, does not predict that p53 prevents cancer. Assuming $high(UV)$ is true at 0, the actions occurring at 0 are $damage(DNA)$ and $upregulate(mRNA)$. Then $high(p53)$ remains false while $unstable(cells)$ becomes true at time 1. Thus *proliferate*$(cells)$ is triggered at time 1, and it does occur at time 1 because it is not inhibited by $high(p53)$ as in the previous case before the refinement. When $high(p53)$ becomes true at time 2, it is "too late" to block the occurrence of *proliferate*$(cells)$. This illustrates that the existing formalism does not gracefully deal with elaboration of a causal rule describing effects of $upregulate(p53)$.

Reasoning about specificity of cellular interactions With doubts sown in our mind, we wondered whether the original knowledge base and its conclusion is itself intuitively correct or not. Our biologist colleagues pointed us to the phenomenon in interaction networks called *specificity* (Tan & Kim 1999), and suggested to us that with the available information, one should not make a definite conclusion about cancer with respect to the original knowledge base. In particular, with high level UV, both the outcomes (cancer and non-cancer) are possible with respect to the p53 network. Thus we need a representation that allows for both the cancer and non-cancer scenarios from both (original and updated) knowledge bases.

Necessity of event ordering knowledge Biologically, there are many possible mechanisms that provide specificity of cellular interactions, which includes: (1) tissue or cell-line specific responses; (2) kinetics that alter the duration of interacting components; (3) integration of multiple interaction networks (Tan & Kim 1999). Since we want to explain about specificity with respect to the same precondition, we do not consider alternative (1). We noticed that event ordering underlies both alternatives (2) and (3). Event ordering can be represented implicitly by duration of events. However, exact knowledge about physical durations are normally not available for biological systems. Hence, we decided to allow for explicit event ordering information which is more commonly available.

As hinted in the earlier examples, the main problem with

$\mathcal{A}_T^0$ semantics is that the "temporal distance" between a triggering precondition and its associated triggered action is bounded (i.e. constant 0). In the following, we propose a new language $\mathcal{A}_T^\infty$ in which the temporal distance is not bounded (hence the superscript ∞).

Action Language $\mathcal{A}_T^\infty$

The *only* features that $\mathcal{A}_T^\infty$ inherits from $\mathcal{A}_T^0$ are: the definition of fluents, actions and states; the *syntax* of domain descriptions (rules (4)-(6)); the definition of state transition (i.e. the transition function Φ) and transition sequences.

Syntax

A theory in $\mathcal{A}_T^\infty$ is a triple $(\mathcal{D}, \mathcal{E}, \mathcal{I})$, where $\mathcal{D}$ is an *domain description*, $\mathcal{E}$ is an *event ordering specification*, and $\mathcal{I}$ is an initial state observation. A domain description in $\mathcal{A}_T^\infty$ is a set of statements of the forms (4)-(6). An initial state observation is a set of statements of the form

$$\textbf{initially } f_1, \ldots, f_n$$

where each f_i is a fluent literal.

An *event* is a set of fluent literals (i.e. *state event*), or a set of actions (i.e. *action event*). A state event is said to happen when its elements hold. An event ordering specification $\mathcal{E}$ is a set of *event orderings* of the form

$$E \textbf{ restricts } E_1 \textbf{ op } E_2$$

where E, E_1 and E_2 are events, and $\textbf{op} \in \{ \prec, \parallel, \preceq \}$. The statement encodes that if E happens, then the earliest happening of E_1 and E_2 after E must obey the ordering **op**. Here $\prec$ means *earlier*, $\parallel$ means *at the same time* and $\preceq$ means either $\prec$ or $\parallel$.

Queries in $\mathcal{A}_T^\infty$ are propositional linear temporal logic (LTL) formulas of fluents without operator "next". [1]

Semantics

We define the semantics of $\mathcal{A}_T^\infty$ in the following steps. First, we define *trajectories* of a domain description in $\mathcal{A}_T^\infty$. Next, we define *interpretations* of a theory $(\mathcal{D}, \mathcal{E}, \mathcal{I})$: an interpretation is a trajectory of $\mathcal{D}$ that satisfies certain properties with respect to $\mathcal{E}$ and $\mathcal{I}$. Finally, we show how a *model* of a theory is chosen among its interpretations.

Trajectories in a domain Informally, a trajectory of a domain $\mathcal{D}$ is a transition sequence of $\mathcal{D}$ such that: *(a)* actions must happen if dictated by the rules in the domain; *(b)* no action happens without being supported by some rules; and *(c)* the set of actions in any state transition is minimal.

Formally, a trajectory is a transition sequence τ that satisfies all the following conditions:

- If $\mathsf{A}_j(\tau) = \emptyset$, then $\mathsf{A}_k(\tau) = \emptyset$ for all $k \geq j$.
- For all rules $f_1, \ldots, f_m$ **triggers** a of $\mathcal{D}$, for all i, if $f_1, \ldots, f_m$ hold in $s_i(\tau)$, and a is not inhibited in $\mathsf{s}_i(\tau)$, then there exists $j \geq i$ such that $a \in \mathsf{A}_j(\tau)$.
- For all j and for all $a \in \mathsf{A}_j(\tau)$, there exists $i \leq j$ and a rule $f_1, \ldots, f_m$ **triggers** a of $\mathcal{D}$ such that $f_1, \ldots, f_m$ hold in state $\mathsf{s}_i(\tau)$ and a is not inhibited in $\mathsf{s}_i(\tau)$.
- There exists no transition sequence $\tau' \neq \tau$ such that τ' also satisfies all the above conditions, τ' starts from the same initial state as τ (i.e. $\mathsf{s}_0(\tau) = \mathsf{s}_0(\tau')$), and $\mathsf{A}_i(\tau') \subseteq \mathsf{A}_i(\tau)$ for all i.

Now let us consider an example of trajectories.

Example 1. Let a domain description D consist of rules: f **triggers** a; f **triggers** b; and b **causes** $\neg f$. There is an infinite number of trajectories of D, such as:

$$\begin{array}{rcl} \tau & = & \langle \{f\}, \{b\}, \{\neg f\}, \{a\}, \{\neg f\} \rangle \\ \tau_0 & = & \langle \{f\}, \{a, b\}, \{\neg f\} \rangle \\ \tau_1 & = & \langle \{f\}, \{a\}, \{f\}, \{a, b\}, \{\neg f\} \rangle \\ \ldots & = & \ldots\ldots \\ \tau_n & = & \langle \{f\}, \{a\}, \ldots, \{f\}, \{a\}, \{f\}, \{a, b\}, \{\neg f\} \rangle \\ \ldots & = & \ldots\ldots \end{array}$$

The sequence $\langle \{f\}, \{b\}, \{\neg f\} \rangle$ is not a trajectory, because a does not occur despite the rule f **triggers** a. The sequence $\langle \{\neg f\}, \{b\}, \{\neg f\} \rangle$ is not a trajectory, because (with respect to this sequence) no trigger rule supports the occurrence of b. The sequence $\tau'' = \langle \{f\}, \{a, b\}, \{\neg f\}, \{a\}, \{\neg f\} \rangle$ is not a trajectory, since $\mathsf{s}_0(\tau_0) = \mathsf{s}_0(\tau'')$ and $\mathsf{A}_i(\tau_0) \subseteq \mathsf{A}_i(\tau'')$. □

Interpretations of a theory Let τ be a trajectory. An event E is said to happen at i with respect to τ if $E \subseteq \mathsf{s}_i(\tau) \cup \mathsf{A}_i(\tau)$. The smallest $i > j$ at which an event E happens with respect to τ is denoted $first(E, \tau, j)$. If E does not happen after j then $first(E, \tau, j) = \infty$. Also, let us define that $i < \infty, \forall i$.

Intuitively, an interpretation of a theory $(\mathcal{D}, \mathcal{E}, \mathcal{I})$ is a trajectory τ that satisfies the specification $\mathcal{E}$ and starts from an initial state described by $\mathcal{I}$. Formally, an interpretation of a theory $(\mathcal{D}, \mathcal{E}, \mathcal{I})$ is a trajectory τ such that:

- For all **initially** $f_1, \ldots, f_n$ of $\mathcal{I}$, for all i: $f_i \in \mathsf{s}_0(\tau)$.
- For all event ordering E **restricts** $E_1 \prec E_2$ in the specification $\mathcal{E}$, if E happens at i with respect to τ, then $first(E_1, \tau, i) < first(E_2, \tau, i)$.
- For all event ordering E **restricts** $E_1 \parallel E_2$ in the specification $\mathcal{E}$, if E happens at i with respect to τ, then $first(E_1, \tau, i) = first(E_2, \tau, i)$.
- For all event ordering E **restricts** $E_1 \preceq E_2$ in the specification $\mathcal{E}$, if E happens at i with respect to τ, then $first(E_1, \tau, i) \leq first(E_2, \tau, i)$.

Example 2. Let us continue with Example 1. Let $I = \{$ **initially** $f\}$. It can be verified that there exists no interpretation of $(D, \{f$ **restricts** $a \prec b\}, I)$. The trajectories τ and τ_0 are interpretations of $(D, \{f$ **restricts** $b \prec a\}, I)$ and $(D, \{f$ **restricts** $a \parallel b\}, I)$ respectively. □

Models of a theory An interpretation τ of $(\mathcal{D}, \mathcal{E}, \mathcal{I})$ is a model of $(\mathcal{D}, \mathcal{E}, \mathcal{I})$ iff there exists no interpretation $\tau' \neq \tau$ such that $\tau' \leq_\infty \tau$, where $\leq_\infty$ is the following partial order.

Definition 2. Let τ and τ' be trajectories where $\mathsf{s}_0(\tau) = \mathsf{s}_0(\tau')$. Then $\tau \leq_\infty \tau'$ iff there exists a sequence $0 = i_0 < i_1 < i_2 < \ldots < i_n < \ldots$ such that for every $0 \leq k$, $\mathsf{A}_k(\tau) \subseteq \bigcup_{j=i_k}^{i_{k+1}-1} \mathsf{A}_j(\tau')$. □

In Example 2, the trajectories τ and τ_0 are the unique models of the theories $(D, \{f$ **restricts** $b \prec a\}, I)$ and

[1] The notion of "next" events is not modelled in $\mathcal{A}_T^\infty$.

$(D, \{f$ **restricts** $a \parallel b\}, I)$ respectively. Note that τ_0 is also the unique model of $(D, \emptyset, I)$.

The semantics based on $\leq_\infty$ captures the intuition that if an action a is triggered by a state event E then: (i) the action a happens at most once after E has happened, unless being triggered by events different from E; and (ii) the action a happens as early as possible after the happening of E.

Query entailment Let τ be a trajectory model of a theory $(\mathcal{D}, \mathcal{E}, \mathcal{I})$. A query Q is entailed by τ iff Q is entailed by the sequence of states $\langle \mathsf{s}_0(\tau), \mathsf{s}_1(\tau), \ldots, \mathsf{s}_i(\tau), \ldots \rangle$ by the standard LTL semantics (Emerson 1990).

A theory $(\mathcal{D}, \mathcal{E}, \mathcal{I})$ is called *consistent* if it has at least one model. A query Q is *weakly entailed* by a consistent theory $(\mathcal{D}, \mathcal{E}, \mathcal{I})$ if it is entailed by a model of $(\mathcal{D}, \mathcal{E}, \mathcal{I})$. The weak entailment is denoted $(\mathcal{D}, \mathcal{E}, \mathcal{I}) \models_{\mathcal{W}} Q$. The query Q is *strongly entailed* by a consistent theory $(\mathcal{D}, \mathcal{E}, \mathcal{I})$ if it is entailed by all the models of $(\mathcal{D}, \mathcal{E}, \mathcal{I})$. The strong entailment is denoted $(\mathcal{D}, \mathcal{E}, \mathcal{I}) \models Q$.

Example 3. Let $\mathcal{D}^+_{afg}$ be the elaborated version of the domain $\mathcal{D}_{afg}$ in Section 1 (based on the rules (1)-(3)).

$$\mathcal{D}^+_{afg} = \{\ \neg f, \neg f' \ \textbf{triggers}\ a'\ ;\quad a' \ \textbf{causes}\ f'\ ;$$
$$f', \neg g \ \textbf{triggers}\ a\ ;\quad a \ \textbf{causes}\ f\ ;$$
$$\neg g \ \textbf{triggers}\ b\ ;\quad b \ \textbf{causes}\ g\ \ \}$$

Let $I^+ = \{$ **initially** $\neg f, \neg f', \neg g\}$. Then the theory $\mathcal{H}_1 = (\mathcal{D}^+_{afg}, \emptyset, I^+)$ has the trajectory models

$\tau_1 = \langle \{\neg f, \neg f', \neg g\}, \{a', b\}, \{\neg f, f', g\} \rangle$

$\tau_2 = \langle \{\neg f, \neg f', \neg g\}, \{a'\}, \{\neg f, f', \neg g\}, \{a, b\}, \{f, f', g\} \rangle$

It follows that $\mathcal{H}_1 \models_{\mathcal{W}} \Diamond \Box\ f$ and $\mathcal{H}_1 \models_{\mathcal{W}} \Box \Diamond \neg f$. Now let $\mathcal{H}_2 = (\mathcal{D}^+_{afg}, \{\{\neg f, \neg g\}$ **restricts** $a' \prec b\}, I^+)$. Then $\mathcal{H}_2$ has the unique model τ_2. It follows that $\mathcal{H}_2 \models \Diamond \Box\ f$.

Reasoning in $\mathcal{A}^\infty_T$

Elaboration tolerance We have shown how $\mathcal{A}^\infty_T$ is motivated by elaboration tolerance issues with specific examples. We now present our general results.

Definition 3 (Elaboration of triggers). Let r be the rule $g_1, \ldots, g_m$ **triggers** b. An elaboration of r is any theory $(R, \emptyset, I)$ such that: *(i)* the theory $H = (R, \emptyset, I \cup \{$ **initially** $g_1, \ldots, g_m\})$ is consistent; *(ii)* any model of H has the form $\langle s_0, A_0, \ldots, A_{i-1}, s_i, \{b\}, s_{i+1} \rangle$, where $b \notin A_j$ for all $0 \leq j \leq i-1$. □

For example, let R consist of the rules (1)-(3) and I be $\{$ **initially** $\neg f'\}$, then $(R, \emptyset, I)$ is an elaboration of the rule $\neg f, \neg g$ **triggers** a. This elaboration is "tolerated" by $\mathcal{A}^\infty_T$, which is a corollary of the following theorem.

Theorem 1. *Let $(\mathcal{D}, \mathcal{E}, \mathcal{I})$ be an $\mathcal{A}^\infty_T$ theory. Let $r \in \mathcal{D}$ be the rule $g_1, \ldots, g_m$* **triggers** *b. Assume that $\mathcal{D}$ contains only one causal rule for a fluent f_b, which is b* **causes** *f_b; and that $\mathcal{D}$ contains only one inhibition rule of b, which is f_b* **inhibits** *b. Let $(R, \emptyset, I)$ be an elaboration of r such that:*

- *all the common symbols (fluents and actions) of R and $\mathcal{D}$ are the symbols found in r; and*
- *the domain R does not contain inhibition rules; and*
- *no action in R affects any fluent in $\{g_1, \ldots, g_m\}$.*

Let $\mathcal{D}^+$ be $\mathcal{D} \setminus \{r\} \cup R$, and let **initially** *$\neg f_b \in \mathcal{I}$. Then:*

(a) If a query Q is weakly entailed by $(\mathcal{D}, \mathcal{E}, \mathcal{I})$ then Q is weakly entailed *by $(\mathcal{D}^+, \mathcal{E}, \mathcal{I} \cup I)$.*

(b) If a query Q is strongly entailed by $(\mathcal{D}, \mathcal{E}, \mathcal{I})$ then Q is strongly entailed *by the theory $(\mathcal{D}^+, \mathcal{E} \cup E, \mathcal{I} \cup I)$; where E is the ordering $\{\{\neg f_b, g_1, \ldots, g_m\}$* **restricts** *$(A_R \setminus b) \prec A_{\mathcal{D}}\}$ with A_R, $A_{\mathcal{D}}$ being the set of the actions in R and $\mathcal{D}$ respectively.*

Proof (sketch). Denote $G = \{g_1, \ldots, g_m\}$; $\mathcal{H} = (\mathcal{D}, \mathcal{E}, \mathcal{I})$; $\mathcal{I}^+ = \mathcal{I} \cup I$ and $\mathcal{H}^+ = (\mathcal{D}^+, \mathcal{E}, \mathcal{I}^+)$. Let τ be a model of $\mathcal{H}$, and let τ_r be a model of $(R, \emptyset, I \cup \{$ **initially** $g_1, \ldots, g_m\})$. It can be verified that b occurs at most once in τ. Intuitively, we will construct a transition sequence τ^+ of $\mathcal{D}^+$ by "merging" τ and τ_r. The construction goes as follows. Denote $s_0 = \mathsf{s}_0(\tau_r) \setminus G$. If b is not triggered by G in τ, then let $\mathsf{s}_i(\tau^+) = \mathsf{s}_i(\tau) \cup s_0$ and $\mathsf{A}_j(\tau^+) = \mathsf{A}_j(\tau)$, for all i, j. Now assume that $G \subseteq \mathsf{s}_N(\tau)$, $f_b \notin \mathsf{s}_N(\tau)$ and $b \in \mathsf{A}_M(\tau)$ (where $N \leq M$). By Definition 3, there exists L such that $\mathsf{A}_{L-1}(\tau_r) = \{b\}$. Let $\mathsf{s}_i(\tau^+) = \mathsf{s}_i(\tau) \cup s_0$ for all $0 \leq i < N$, and let $\mathsf{A}_j(\tau^+) = \mathsf{A}_j(\tau)$ for all $0 \leq j < N$. Let $\mathsf{s}_{N+i}(\tau^+) = \mathsf{s}_i(\tau_r) \cup \mathsf{s}_N(\tau)$ and $\mathsf{A}_{N+j}(\tau^+) = \mathsf{A}_j(\tau_r)$, for all $0 \leq i, j < L-1$. Then let $\mathsf{s}_{L+i-1}(\tau^+) = \mathsf{s}_i(\tau_r) \cup \mathsf{s}_{L-1}(\tau_r)$ for $N \leq i \leq M$, and let $\mathsf{s}_{L+j-1}(\tau^+) = \mathsf{s}_j(\tau) \cup \mathsf{s}_L(\tau_r)$ for all $j > M$. Finally, let $\mathsf{A}_{L+i-1}(\tau^+) = \mathsf{A}_i(\tau)$ for all $i \geq N$.

In the case *(a)*, the constructed τ^+ is a model of $\mathcal{H}^+$ which entails Q. In *(b)*, all sequences τ^+ constructed from all the pairs of (τ, τ_r) are all the models of $(\mathcal{D}^+, \mathcal{E} \cup E, \mathcal{I}^+)$. □

We have a similar result on elaboration of causal rules.

Proposition 1. *Let $(\mathcal{D}, \mathcal{E}, \mathcal{I})$ be an $\mathcal{A}^\infty_T$ theory. Let $r \in \mathcal{D}$ be the rule: a* **causes** *f* **if** *f_1. Let g, g_1, g_2 be fluents and b be action that are not in $\mathcal{D}$. Let R consist of the trigger g, g_2* **triggers** *b and the causal rules a* **causes** *g* **if** *f_1, g_1; and b* **causes** *f, $\neg g$. Let $I = \{$* **initially** *$\neg g, g_1, g_2\}$. Then:*

(a) If a query Q is weakly entailed by $(\mathcal{D}, \mathcal{E}, \mathcal{I})$ then Q is weakly entailed *by $(\mathcal{D} \setminus \{r\} \cup R, \mathcal{E}, \mathcal{I} \cup I)$.*

(b) If a query Q is strongly entailed by $(\mathcal{D}, \mathcal{E}, \mathcal{I})$ then Q is strongly entailed *by $(\mathcal{D} \setminus \{r\} \cup R, \mathcal{E} \cup E, \mathcal{I} \cup I)$; where E is the ordering $\{\{a, f_1\}$* **restricts** *$b \prec A_{\mathcal{D}}\}$ with $A_{\mathcal{D}}$ being the set of the actions in $\mathcal{D}$.*

Reasoning about exogenous interventions It is important to be able to reason about effects of external agent actions on triggered evolutions of a system; for example, we want to know how to intervene to alter the cell behavior in desirable ways. Since $\mathcal{A}^\infty_T$ does not represent explicit time, we cannot specify an execution of an exogenous action by explicitly setting a time for it (e.g. exogenous actions with time steps (Tran & Baral 2004)), or implicitly setting the execution time using ordering (e.g. a plan as a sequence of actions). However, we can combine the modelled system (e.g. an interaction network) together with the agent control (e.g. a scientist) into one system and reason about the exogenous actions of the agent as triggers in the combined system.

Formally, we consider external interventions as conditional actions of the form:

$$\alpha = \ \textbf{do}\ a\ \textbf{if}\ f_1, f_2, \ldots, f_n,$$

where a is an action, and $f_1, f_2, \ldots, f_n$ are fluent literals. Given a conditional action α of the above form, we denote $trig(\alpha) = f_1, f_2, \ldots, f_n$ **triggers** a and $act(\alpha) = a$.

Given a set C of conditional actions, we denote $trig(C) = \{trig(\alpha) \mid \alpha \in C\}$ and $act(C) = \{act(\alpha) \mid \alpha \in C\}$.

An *agent control* can be represented as a pair (C, L), where C is a set of conditional actions and L is a set of dynamic causal laws describing effects of the actions in $act(C)$. Let $(\mathcal{D}, \mathcal{E}, \mathcal{I})$ be a theory in $\mathcal{A}_T^\infty$ and (C, L) be an agent control. We can assume that the actions of $act(C)$ are not in $\mathcal{D}$. Then we say that $(\mathcal{D}, \mathcal{E}, \mathcal{I})$ *entails* a query Q given (C, L) iff:

$$(\mathcal{D} \cup L \cup trig(C), \mathcal{E}, \mathcal{I}) \models Q \ .$$

Planning for a goal Q is to find a subset $P \subseteq C$ such that:

$$(\mathcal{D} \cup L \cup trig(P), \mathcal{E}, \mathcal{I}) \models Q \ .$$

Application to the p53 network Let $\mathcal{D}^+_{p53}$ be the updated version of $\mathcal{D}_{p53}$ in Section 2. Let $\mathcal{I}$ be an initial state observation that only $high(UV)$ is true.

First, we have that $(\mathcal{D}_{p53}, \emptyset, \mathcal{I}) \models \Box\ \neg tumorous$. This corresponds to the fact that the presence of high UV does not cause cancer. Next, let $\mathcal{H}_0 = (\mathcal{D}^+_{p53}, \emptyset, \mathcal{I})$. We have that $(\mathcal{D}^+_{p53}, \emptyset, \mathcal{I}) \models_{\mathcal{W}} \Diamond\ \neg tumorous$ and $(\mathcal{D}^+_{p53}, \emptyset, \mathcal{I}) \models_{\mathcal{W}} \Diamond\ tumorous$. Thus the theory $\mathcal{H}_0$ captures the specificity phenomenon of the p53 network.

Now let $\mathcal{E}$ be the event ordering

$high(UV), \neg high(p53)$ **restricts**

$$high(p53) \preceq unstable(cells).$$

This event ordering knowledge has a known biological counterpart. It is known that in a healthy cellular environment, the upregulation of p53 is triggered rapidly in the presence of high level UV and before the cell becomes unstable.

Now let $\mathcal{H}_1 = (\mathcal{D}^+_{p53}, \mathcal{E}, \mathcal{I})$, then $\mathcal{H}_1 \models \Box\ \neg tumorous$. Thus one can make more specific predictions by incorporating event orderings.

Conclusion

In this paper we describe several interesting issues that we encountered while modelling cellular interactions, and that are not appropriately modelled in existing languages and systems. We then build up on an existing language, and give the semantics of our language. We show that using our language one can express the specificity issue in cell signaling, can correctly reason when actions or triggers are elaborated, and can express and reason with event ordering knowledge.

In terms of related work, there have been several proposal for modelling cell behavior and cellular interaction, including proposals such as the use of Petri nets (Peleg, Yeh, & Altman 2002), process algebra (Regev, Silverman, & Shapiro 2001) and many others mentioned in (Baral *et al.* 2004). Most of these methods are more focussed towards simulation than about reasoning, and can not do explanation or reasoning. Even for prediction, they do not address the issues of specificity, elaboration and refinement, and do not allow for specification of event ordering. There exist action formalisms for representing triggered events and concurrent actions, such as situation calculus with natural actions (Reiter 1996), the language $\mathcal{E}$ (Kakas & Miller 1998), the language C+ (Giunchiglia *et al.* 2004), and (Thielscher 2000), but none of them address the issues studied in this paper, in particular the issue of elaboration of actions and triggers at one granularity via more detailed description at a finer granularity. The work of (Fukuda & Takagi 2001) addresses presentation of hierarchical knowledge, however it does not support reasoning about triggers and concurrent actions.

Acknowledgments

We acknowledge the support of NSF under grants numbered 0412000 and 0070463 and thank Dr. Vinay Nagaraj for stimulating and insightful discussion on the p53 network.

References

Baral, C.; Chancellor, K.; Tran, N.; Tran, N.; and Berens, M. 2004. A knowledge based approach for representing and reasoning about signaling networks. *Bioinformatics 20 (Suppl 1)* i15–i22.

Calvanese, D.; de Giacomo, G.; and Vardi., M. 2002. Reasoning about actions and planning in LTL action theories. In *Proc. of KR 2002*, 593–602.

Chabrier, N.; Chiaverini, M.; Danos, V.; Fages, F.; and Schachter, V. 2004. Modeling and querying biomolecular interaction networks. *Theoretical Computer Science* 325(1):25–44.

Emerson, E. 1990. Temporal and modal logic. In van Leeuwen, J., ed., *Handbook of Theoretical Computer Science, Volume B*, 997–1072.

Frei, H. 2003. *Cancer Medicine*.

Fukuda, K., and Takagi, T. 2001. Knowledge representation of signal transduction pathways. *Bioinformatics* 17(9):829–37.

Giordano, L.; Martelli, M.; and Schwind, C. 2001. Reasoning about actions in dynamic linear time temporal logic. *Journal of the IGPL* 9(2):289–303.

Giunchiglia, E.; Lee, J.; Lifschitz, V.; McCain, N.; and Turner, H. 2004. Nonmonotonic causal theories. *Artificial Intelligence* 153(1-2):49–104.

Kakas, A., and Miller, R. 1998. Reasoning about actions, narratives and ramifications. *ETAI* 1:39–72.

Khan, S.; Decker, K.; Gillis, W.; and Schmidt, C. 2003. A multi-agent system-driven AI planning approach to biological pathway discovery. In *Proc. of ICAPS 2003*.

Peleg, M.; Yeh, I.; and Altman, R. B. 2002. Modelling biological processes using workflow and petri net models. *Bioinformatics* 18(6):825–837.

Reddy, V. N.; Liebman, M. N.; and Mavrovouniotis, M. L. 1996. Qualitative analysis of biochemical reaction systems. *Computers in Biology and Medicine* 26:9–24.

Regev, A.; Silverman, W.; and Shapiro, E. 2001. Representation and simulation of biochemical processes using π-calculus process algebra. In *Proc. of PSB 2001*, 459–470.

Reiter, R. 1996. Natural actions, concurrency and continuous time in the situation calculus. In *Proc. of KR 1996*, 2–13.

Talcott, C.; Eker, S.; Knapp, M.; Lincoln, P.; and Laderoute, K. 2004. Pathway logic modeling of protein functional domains in signal transduction. In *Proc. of PSB 2004*, 568–580.

Tan, P. B., and Kim, S. K. 1999. Signaling specificity: the RTK/RAS/MAP kinase pathway in metazoans. *Trends in Genetics* 15(4):145–149.

Thielscher, M. 2000. Representing the knowledge of a robot. In *Proc. of KR 2000*, 109–120.

Tran, N., and Baral, C. 2004. Reasoning about triggered actions in AnsProlog and its application to molecular interactions in cells. In *Proc. of KR 2004*, 554–563.