Completing causal networks by meta-level abduction

Katsumi Inoue, Andrei Doncescu, Hidetomo Nabeshima

Machine Learning manuscript No.
(will be inserted by the editor)

Completing causal networks by meta-level abduction

Katsumi Inoue · Andrei Doncescu · Hidetomo Nabeshima

Received: date / Accepted: date

Abstract Meta-level abduction is a method to abduce missing rules in explaining observations. By representing rule structures of a problem in a form of causal networks, meta-level abduction infers missing links and unknown nodes from incomplete networks to complete paths for observations. We examine applicability of meta-level abduction on networks containing both positive and negative causal effects. Such networks appear in many domains including biology, in which inhibitory effects are important in several biological pathways. Reasoning in networks with inhibition involves nonmonotonic inference, which can be realized by making default assumptions in abduction. We show that meta-level abduction can consistently produce both positive and negative causal relations as well as invented nodes. Case studies of meta-level abduction are presented in p53 signaling networks, in which causal relations are abduced to suppress a tumor with a new protein and to stop DNA synthesis when damage has occurred. Effects of our method are also analyzed through experiments of completing networks randomly generated with both positive and negative links.

Keywords abduction · theory completion · meta-reasoning · causal network · inhibition · predicate invention · default reasoning · Inductive Logic Programming · Systems Biology

This research was supported in part by the 2008-2011 JSPS Grant-in-Aid for Scientific Research (A) No. 20240016.

K. Inoue
National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan
E-mail: inoue@nii.ac.jp

A. Doncescu
LAAS-CNRS UPR 8001
Avenue du Colonel Roche, 31007 Toulouse, France
E-mail: andrei.doncescu@laas.fr

H. Nabeshima
Division of Medicine and Engineering Science, University of Yamanashi
4-3-11 Takeda, Kofu, Yamanashi 400-8511, Japan
E-mail: nabesima@yamanashi.ac.jp

1 Introduction

Abduction and induction are both ampliative reasoning, and play essential roles in knowledge discovery and development in science and technology. Integration of abduction and induction has been discussed in such diverse aspects as implementation of inductive logic programming (ILP) systems using abductive methods [67, 14, 52, 54, 27, 3] and "closing the loop" methodologies in scientific discovery [10, 28, 55]. The use of prior or *background knowledge* in scientific applications has directed our attention to *theory completion* [40] rather than classical learning tasks such as concept learning and classification. There, *abduction* is mainly used to complete proofs of observations from *incomplete* background knowledge, while *induction* refers to generalization of abduced cases.

In scientific domains, background knowledge is often structured in a form of *networks*. In biology, a sequence of signalings or biochemical reactions constitutes a network called a *pathway*, which specifies a mechanism to explain how genes or cells carry out their functions. However, information of biological networks in public-domain databases is generally incomplete in that some details of reactions, intermediary genes/proteins or kinetic information are either omitted or undiscovered. To deal with incompleteness of pathways, we need to predict the status of relations which is consistent with the status of nodes [63, 68, 51, 21, 55], or insert missing arcs between nodes to explain observations [69, 28, 65, 1]. These goals are characterized by abduction as theory completion, in which status of nodes or missing arcs are added to account for observations.

If a network is represented in a logical theory, inference on the network can be realized on a language in which each network element such as a node and a link is itself an entity of the language. Such a theory (or program) refers a language of the network, and is hence regarded as a *meta-theory* (or *meta-program*), while a theory (or program) representing networks is referred to as an *objective theory* (or *objective program*). Then, to perform abduction on networks, we need abduction on meta-theories, which is referred to as *meta-level abduction*. Note that *meta-reasoning* has been intensively investigated in logic programming, e.g., [29, 12, 4], yet main inference considered there has been deduction.

Meta-level abduction was introduced in [18] as a method to discover unknown relations from incomplete networks. Given a network representing causal relations, called a *causal network*, missing links and nodes are abduced in the network to account for observations. The main objective in [18] is to provide a logical foundation and knowledge representation for *abducing rules*, which is an important abductive problem in ILP. It is notable that other abductive methods [49, 3] need a predetermined set of candidate rules (called *abducible rules* or *abducibles*) and then select consistent combinations of abducibles to form explanations. In contrast, meta-level abduction does not need any such abducibles in advance. Meta-level abduction is implemented in SOLAR [44, 45], an automated deduction system for consequence finding, using a first-order representation for algebraic properties of causality and a full-clausal form of network information and constraints. Meta-level abduction by SOLAR is powerful enough to infer missing rules, missing facts, and unknown causes involving *predicate invention* [41] in the form of existentially quantified hypotheses. Note that predicate invention had been intensively investigated in the initial stage of ILP research and has recently been revisited [42], since it should play an important role in discovery.

Meta-level abduction has been applied to discover physical skills in terms of *hidden rules* to explain given *empirical rules* in cello playing examples in [18], and a thorough experimental analysis with a variety of problem instances has been presented in [45]. However, all those examples of meta-level abduction in [18, 45] contain only one kind of causal effects, which are positive, and it was left open how to deal with *both positive and negative effects*.

Then, we shall examine applicability of meta-level abduction to deal with networks expressing both positive and negative causal effects. Such networks are often used in biology, where *inhibitory effects* are essential in gene regulatory, signaling and metabolic networks.

In this paper, we present axioms for meta-level abduction to produce both positive and negative causal relations as well as newly invented nodes. We show two axiomatizations for such meta-level abduction. One is a set of *alternating axioms* which define relations of positive and negative causal effects in a double inductive manner. This axiom set reduces to the axioms for ordinary meta-level abduction defined in [18] when there is no negative causal link. The other axiomatization is a variant of the alternating axioms, but prefers negative links to positive ones if both are connected to the same node. In this case, reasoning in causal networks becomes *nonmonotonic*, and involves default assumptions in abduction.

Then, applications to p53 signal networks [50, 36] are presented as case studies of our framework, in which meta-level abduction reproduces theories explaining how tumor suppressors work [65] and how DNA synthesis stops [62]. Analysis of such abstract signaling networks, although simple, provides one of the most fundamental inference problems in computer-aided scientific research including Systems Biology: *Given an incomplete causal network, infer possible connections and functions of network entities to reach the target entities from the sources.* Meta-level abduction in this paper is crucial in this task for the following reasons. First, suggestion of possible additions in prior networks enables scientists to conduct hypothesis-driven experiments with those focused cases. If a suggested hypothesis is justified through a through set of experiments, the corresponding new links and/or nodes are considered to be discovered. In network completion, however, the larger the network becomes, the more abductive inference steps are required to get a hypothesis and the more candidate hypotheses are inferred. Then, it is hard for human scientists to consider all possibilities without losing any important ones. Therefore, automation of hypothesis enumeration is very important. Second, abduction in such network domains often involves a *goal* rather than an *observation*: A hypothesis is inferred to achieve the goal that has not been observed yet. For example, in drug design and pharmacology as well as therapeutic research, the effect of introduction of new entities and links to a known network is goal-oriented and the same hypotheses cannot be applied to other goals in general. This feature of goal-oriented abduction also exists in completing causal networks for improvement of physical techniques in musical performance [18], in which specific skills are required for requested tasks.

Finally, scalability of meta-level abduction is analyzed through experiments of completing networks randomly generated with both positive and negative links. By varying the average degree of nodes to edges, from sparse to dense networks are generated with several sizes. We will see that it is not easy to generate a large and dense network by keeping the consistency, since there are more chances to become inconsistent in such networks. Then the growth rate of hypotheses in the size of networks rather decreases in dense networks.

This is an extended version of the paper [17], and contains several technical details such as an abductive procedure based on consequence finding, theoretical correctness for the proposed formalization and their proofs, detailed analysis of experiments on p53 pathways, and scalability issues. The rest of this paper is organized as follows. Section 2 offers the essential of meta-level abduction and its use for rule abduction. Section 3 then extends meta-level abduction to allow for two types of causal effects, in which positive and negative rules are called *triggers* and *inhibitors*, respectively, and investigate properties of two axiomatizations. Section 4 presents two case studies of meta-level abduction applied to completion of sub-networks in p53 signal networks. Section 5 shows experiments on causal networks, and analyzes scalability of our method in completing networks. Section 6 discusses related work, and Section 7 gives a summary and future work.

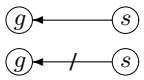
2 Meta-level Abduction

This section revisits the framework for *meta-level abduction* [18], and provides the correctness of rule abduction.

2.1 Causal Networks

We suppose a background theory represented in a network structure called a *causal graph* or a *causal network*. A causal graph is a directed graph representing causal relations, which consists of a set of *nodes* and a set of *(directed) arcs* (or *links*).[1] Each node in a causal graph represents some event, fact or proposition. A *direct causal relation* corresponds to a directed arc, and a *causal chain* is represented by the reachability between two nodes. The interpretation of a "cause" here is kept rather informal, and just represents the connectivity, which may refer to a mathematical, physical, chemical, conceptual, epidemiological, structural, or statistical dependency [48]. Similarly, a "direct cause" here simply represents the adjacent connectivity, while its effect is direct only relative to a certain level of abstraction.

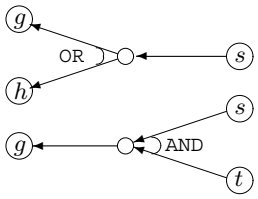
We then consider a first-order language to express causal networks. Each node is represented as a *proposition* or a (ground) atom in the language. When there is a direct causal relation from a node s to a node g, we define that $connected(g, s)$ is true as in (1). Note that $connected(g, s)$ only shows that s is one of possible causes of g, and thus the existence of another $connected(g, t)$ ($s \neq t$) means that s and t are alternative causes for g. Here, expression of causal relations is done at the *meta level* using the *meta-predicate* $connected$, while the *object level* refers to nodes in a causal network. An atom $connected(s, t)$ at the meta level corresponds to a rule $(s \leftarrow t)$ at the object level. The fact that a direct causal link cannot exist from s to g is represented in an *(integrity) constraint* of the form (2).



$$connected(g, s) \tag{1}$$

$$\neg connected(g, s) \tag{2}$$

A direct causal relation from s which has *nondeterministic effects* g and h, written as $(g \vee h \leftarrow s)$ at the object level, is represented in a disjunction of the form (3) at the meta level. On the other hand, the relation that "g is *jointly* caused by s and t", written as $(g \leftarrow s \wedge t)$ at the object level, is expressed in a disjunction of the form (4) at the meta level, viz., $(g \leftarrow s \wedge t) \equiv (g \leftarrow s) \vee (g \leftarrow t)$.



$$connected(g, s) \vee connected(h, s) \tag{3}$$

$$connected(g, s) \vee connected(g, t) \tag{4}$$

There can be more than two atoms in a disjunction of the form (3) or (4). For example, $(g \leftarrow s \wedge t \wedge u)$ at the object level can be expressed as $connected(g, s) \vee connected(g, t) \vee connected(g, u)$. A complex relation of the form $(g \vee h \leftarrow s \wedge t)$ that has more than one node in both the left-hand and right-hand sides of the rule can be decomposed into two relations,

[1] Precisely speaking, our causal networks bring us more information than directed graphs, since negation, disjunctive effects and joint causes are all represented in a network.

$(s\text{-}t \leftarrow s \wedge t)$ and $(g \vee h \leftarrow s\text{-}t)$, where s-t represents the intermediate complex. Then, any direct causal relation in a causal network can be represented by at most two disjunctions of atoms of the forms (3) and (4) at the meta level, using intermediate complexes.

In the above expression, (i) each atom at the object level is represented as a *term* at the meta level, and (ii) each rule at the object level is represented as a *fact* or a disjunction of facts at the meta level. The point (ii) can not only hold for rules given in the axioms, but can also be applied to express *inferred rules* at the meta level. Now, to express inferred rules, we introduce another meta-predicate $caused$. For object-level propositions g and s, we define that $caused(g, s)$ is true if there is a *causal chain* from s to g. Then, the causal chains are generally defined transitively in terms of $connected$ as the axioms with variables:

$$caused(X, Y) \leftarrow connected(X, Y). \tag{5}$$

$$caused(X, Y) \leftarrow connected(X, Z) \wedge caused(Z, Y). \tag{6}$$

Here, the $caused/2$ relation is recursively defined with the $connected/2$ facts. The first-order expression with variables is thus used to represent that these axioms hold for all instances of them. Other algebraic properties as well as some particular constraints (e.g., $\neg caused(a, b)$) can also be defined if necessary. Variables in object-level expressions like $g(T)$ and $s(T)$ can be allowed in the meta-level expression like $connected(g(T), s(T))$, where the predicates g and s are treated as function symbols in the same way that Prolog can allow higher-order expressions. Here, an expression $g(T)$ can represent a set of nodes with the same property g with different values of the argument T such as time.

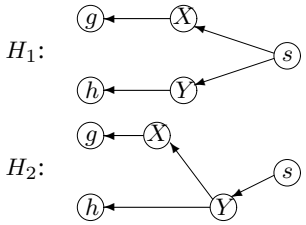
2.2 Rule Abduction

Reasoning about causal networks is realized by deduction and abduction from the meta-level expression of causal networks together with the axioms for causal relations including (5) and (6). In *rule deduction*, we will later prove in Proposition 1 that, if a meta-level expression of the form $caused(g, s)$ for some facts g and s can be derived, it means that the rule $(g \leftarrow s)$ can be derived at the object level.

Similarly, we can realize *rule abduction* in the meta-level representation as follows. Suppose that a fact g is somehow caused by a fact s, which can also be regarded as an *input-output relation* that an output g is obtained given an input s. Here, g and s are called a *goal (fact)* (or a *target (fact)*) and a *source (fact)*, respectively. Setting an *observation* O as the causal chain $caused(g, s)$, we want to explain why or how O is caused. At the object level, O corresponds to the rule $(g \leftarrow s)$, which can be given as either a real observation (called an *empirical rule*) or a *virtual goal* to be achieved. An abductive task is then to find *hidden rules* that establish a connection from the source s to the goal g by filling the gaps in causal networks. As we will later see in Theorem 1, such an empirical rule can have more than one antecedent. For example, an object-level observation $(g \leftarrow s \wedge t)$ can be expressed as the meta-level formula $(caused(g, s) \vee caused(g, t))$ in the same way as (4).

Logically speaking, a *background theory* B consists of the meta-level expression of a causal network N and the axioms for causal relations at the meta level containing (5) and (6). When B is *incomplete*, there may be no path between g and s in B, that is, $caused(g, s)$ cannot be derived from B. Then, abduction infers an *explanation* (or *hypothesis*) H consisting of missing relations (links) and missing facts (nodes). This is realized by setting the *abducibles* Γ, the set of candidate literals to be assumed, as the atoms with the predicate $connected$: $\Gamma = \{connected(\_, \_)\}$. It is sometimes declared that there cannot exist any direct causal relation between the source and the goal, i.e., $\neg connected(g, s)$.

Formally, given a set O of formulas, a set H of instances of elements of Γ is an (*abductive*) *explanation* of O (*with respect to* B) if $B \cup H \models O$ and $B \cup H$ is consistent. A set of formulas can be interpreted as the conjunction of them. An explanation H of O is *minimal* if it does not imply any explanation of O that is not logically equivalent to H. Minimal explanations in meta-level abduction correspond to minimal additions in causal graphs, and are reasonable according to the principle of Occam's razor. For example, suppose the observation $O = caused(g, s) \wedge caused(h, s)$, that is, the multiple causal chains between two goal facts g, h and the source fact s. Examples of minimal explanations of O containing two intermediate nodes are as follows.



$$\exists X \exists Y (connected(g, X) \wedge connected(h, Y) \wedge connected(X, s) \wedge connected(Y, s))$$

$$\exists X \exists Y (connected(g, X) \wedge connected(X, Y) \wedge connected(h, Y) \wedge connected(Y, s))$$

H_1 and H_2 represent different connectivities, and we may want to enumerate different types of network structures that are missing in the original causal network. Here, H_1 corresponds to the four rules $\{(g \leftarrow \chi), (\chi \leftarrow s), (h \leftarrow \psi), (\psi \leftarrow s)\}$, hence rule abduction, i.e., abduction of rules, is realized. Moreover, these hypotheses contain existentially quantified variables, where χ and ψ are newly invented here. Those new terms can be regarded as either some existing nodes or new unknown nodes. Since new formulas can be produced at the object level, *predicate invention* [41] is partially realized in meta-level abduction.[2]

A hypothesis with a joint cause of the form (4) can be obtained by taking a disjunction of explanations of the form $connected(g, \_)$ or by obtaining a *disjunctive answer* [53] for an observation containing a free variable X of the form $caused(g, X)$. Alternatively, this can be realized by adding a meta-level axiom:

$$connected(X, Y) \vee connected(X, Z) \leftarrow jointly\_connected(X, Y, Z). \tag{7}$$

to the background theory B and the literals of the form $jointly\_connected(\_, \_, \_)$ to the abducibles Γ. Causes consisting of more than two joint links can be represented in a similar way. Atoms of the form $jointly\_connected(\_, \_, \_)$ together with axiom (7) can also be used to represent conjunctive causes of the form (4) in a causal network. That is, to express $(g \leftarrow s \wedge t)$ at the object level, the atom $jointly\_connected(g, s, t)$ can be used instead of the formula $(connected(g, s) \vee connected(g, t))$ at the meta level.

The soundness and completeness of rule abduction in meta-level abduction can be derived as follows.[3] For any meta-level theory N such that the predicate of any formula appearing in N is $connected$ only, let $\lambda(N)$ be the object-level theory obtained by replacing every $connected(t_1, t_2)$ (t_1 and t_2 are terms) appearing in N with the formula $(t_1 \leftarrow t_2)$. We first show the correctness of meta-level deduction in the next proposition.

Proposition 1 *Suppose a meta-level theory N, which consists of disjunctions of facts of the form $connected(\_, \_)$. Let the background theory be $B = N \cup \{(5), (6)\}$. Then, $B \models (caused(g, s_1) \vee \cdots \vee caused(g, s_n))$ if and only if $\lambda(N) \models (g \leftarrow s_1 \wedge \cdots \wedge s_n)$.*

[2] Predicate invention in this form can also be regarded as a realization of *hidden object invention* discussed by Muggleton in [7].

[3] The properties presented here are generalizations of preliminary results in [18, Section 3.3], which proved the correctness when the observation is a fact, i.e., a single atom, of the form $caused(g, s)$.

Proof We prove the proposition by induction on the depth of proof trees.[4]

Induction basis. It holds by the meaning of causal networks that $caused(g, s)$ is derived in a proof having depth 1 iff $connected(g, s) \in B$ (by (5)) iff $(g \leftarrow s) \in \lambda(N)$. Then, $(caused(g, s_1) \vee \cdots \vee caused(g, s_n))$ is derived in a proof having depth 1 iff $(connected(g, s_1) \vee \cdots \vee connected(g, s_n)) \in N$ iff $((g \leftarrow s_i) \vee \cdots \vee (g \leftarrow s_n)) \in \lambda(N)$ iff $\lambda(N) \models (g \leftarrow s_1 \wedge \cdots \wedge s_n)$.

Induction hypothesis. Suppose that the proposition holds for any formula $(caused(g, s_1) \vee \cdots \vee caused(g, s_n))$ derived from B in a proof tree having depth d such that $d \leq k$.

Induction step. A formula $(caused(g, s_1) \vee \cdots \vee caused(g, s_n))$ is derived in a proof tree having depth $k+1$ iff $B \models ((connected(g, s_1) \vee \exists s_1'(connected(g, s_1') \wedge caused(s_1', s_1))) \vee \cdots \vee (connected(g, s_n) \vee \exists s_n'(connected(g, s_n') \wedge caused(s_n', s_n))))$ (by (5) and (6)) such that $caused(s_j', s_j)$ is derived in a proof tree having depth k for $j = 1, \ldots, n$ iff $\lambda(N) \models (((g \leftarrow s_1) \vee \exists s_1'((g \leftarrow s_1') \wedge (s_1' \leftarrow s_1))) \vee \cdots \vee ((g \leftarrow s_n) \vee \exists s_n'((g \leftarrow s_n') \wedge (s_n' \leftarrow s_1))))$ (by the induction hypothesis) iff $\lambda(N) \models ((g \leftarrow s_1) \vee \cdots \vee (g \leftarrow s_n))$ iff $\lambda(N) \models (g \leftarrow s_1 \wedge \cdots \wedge s_n)$. □

Note in Proposition 1 that we do not need the λ-counterparts of axioms (5) and (6) in the object level. This logic reflects the assumption that transitivity of cause by $\leftarrow$ holds for a chain in a causal network. Now we have the correctness of meta-level abduction.

Theorem 1 *Let N and B be the same theories as in Proposition 1. Suppose the observation $O = (caused(g, s_1) \vee \cdots \vee caused(g, s_n))$, and let $\Gamma = \{connected(\_, \_)\}$. Then, H is an abductive explanation of $caused(g, s)$ with respect to B and Γ if and only if $\lambda(H)$ is a hypothesis satisfying that*

$$\lambda(N) \cup \lambda(H) \models (g \leftarrow s_1 \wedge \cdots \wedge s_n), \text{ and} \tag{8}$$

$$\lambda(N) \cup \lambda(H) \text{ is consistent.} \tag{9}$$

Proof The equivalence between the relation that $B \cup H \models O$ and the abductive derivation (8) holds by Proposition 1. The equivalence between the consistency of $B \cup H$ and the consistency (9) is obvious: $B \cup H$ is consistent because it contains no integrity constraint, and so is $\lambda(N) \cup \lambda(H)$. □

2.3 Abduction of Rules and Facts

Besides the use in rule abduction, meta-level abduction can also be applied to *fact abduction* [18], which has been focused on almost exclusively in research of abduction in AI.[5] Abduction of facts at the object level can be formalized as *query answering* at the meta level. Given a goal of the form $caused(g, X)$, abduction of causes is computed by answer substitutions to the variable X. To this end, each abducible literal a at the object level is associated with the fact $caused(a, a)$ at the meta level. That is, an abducible can hold by assuming itself. Equivalently, the axiom for abducibles is expressed using the meta-predicate abd as:

$$caused(X, X) \leftarrow abd(X).$$

[4] For example, SOL tableaux [44,45] (Section 2.4) can be used as proof trees in deducing target consequences. Since SOL(AR) is complete for consequence-finding [13,45], for any minimal clause C derived from a consistent axiom set, there is an SOL tableau producing C with a certain depth.

[5] In a philosophical work [60], fact abduction and rule abduction are classified as *factual abduction* and *law-abduction*, respectively. Our meta-level abduction also gives a realization of *2nd order existential abduction*, which is most important to produce new theories with new concepts [60].

Then, each abducible a at the object level should be declared as $abd(a)$. Answer extraction for the query $\leftarrow caused(g, X)$ can be realized by giving the meta-level formula of the form:

$$ans(X) \leftarrow caused(g, X) \wedge abd(X). \tag{10}$$

Here, ans is the *answer predicate* [23], and the variable X is used to collect only abducibles which cause g. An *integrity constraint* that two facts p and q cannot hold at the same time ($\leftarrow p \wedge q$) can be represented as:

$$\leftarrow caused(p, X) \wedge caused(q, Y) \wedge abd(X) \wedge abd(Y).$$

This makes any combination of abducibles that causes p and q incompatible. Such a set of incompatible abducibles is called a *nogood*. Finally, by combining rule abduction and fact abduction in the form of *conditional query answering* [23], which extracts answers in a query with additional abduced conditions, meta-level abduction enables us to *abduce both rules and facts* [18].

2.4 Computation by Consequence Finding

All types of meta-level inferences in this section, including generation of existentially quantified hypotheses in meta-level abduction as well as conditional query answering to abduce rules and facts, can be realized by SOLAR [44,45]. SOLAR is a *consequence-finding* system based on SOL resolution [13] and the connection tableaux.

In SOLAR, the notion of production fields [13] is used to represent language biases for hypotheses. A *clause* is a disjunction of literals. A *production field* $\mathcal{P}$ is a pair $\langle \mathbf{L}, Cond \rangle$, where $\mathbf{L}$ is a set of literals and $Cond$ is a certain condition. If $Cond$ is $true$ (empty), $\mathcal{P}$ is denoted as $\langle \mathbf{L} \rangle$. A clause C *belongs to* $\mathcal{P} = \langle \mathbf{L}, Cond \rangle$ if every literal in C is an instance of a literal in $\mathbf{L}$ and C satisfies $Cond$. Let Σ be a clausal theory. The set of consequences of Σ belonging to $\mathcal{P}$ is denoted as $Th_{\mathcal{P}}(\Sigma)$. The *characteristic clauses* of Σ with respect to $\mathcal{P}$ are defined as $Carc(\Sigma, \mathcal{P}) = \mu\, Th_{\mathcal{P}}(\Sigma)$, where μT denotes the set of clauses in T that are minimal with respect to subsumption. The *new characteristic clauses* of a clause C with respect to Σ and $\mathcal{P}$ are defined as $Newcarc(\Sigma, C, \mathcal{P}) = \mu\,[\, Th_{\mathcal{P}}(\Sigma \cup \{C\}) \setminus Th_{\mathcal{P}}(\Sigma)\,]$.

Let B be a clausal theory (background theory) and O a set of literals (observations). Then, a set H of literals is obtained as an abductive explanation of O by *inverse entailment* [13]:

$$B \cup \{\neg O\} \models \neg H, \tag{11}$$

where both $\neg O = \bigvee_{L \in O} \neg L$ and $\neg H = \bigvee_{L \in H} \neg L$ are clauses (because O and H are sets of literals and are interpreted as conjunctions of them). Similarly, the condition that $B \cup H$ is consistent is equivalent to $B \not\models \neg H$. Hence, for any hypothesis H, its negated form $\neg H$ is deductively obtained as a "new" consequence of $B \cup \{\neg O\}$ which is not an "old" consequence of B alone. Given the abducibles Γ, any literal in $\neg H$ is an instance of a literal in $\overline{\Gamma} = \{\neg L \mid L \in \Gamma\}$. Hence, the set of minimal explanations of O with respect to B and Γ is characterized as $\{\, H \mid \neg H \in Newcarc(B, \neg O, \langle \overline{\Gamma} \rangle) \,\}$, while the set of minimal nogoods with respect to B and Γ is $\{\, H \mid \neg H \in Carc(B, \langle \overline{\Gamma} \rangle) \,\}$.

SOLAR is complete for finding (new) characteristic clauses with respect to a given production field. SOLAR can thus be used to implement a *complete* abductive system for finding and enumerating minimal explanations from *full clausal theories* containing non-Horn clauses. A simple way to compute $Newcarc(\Sigma, C, \mathcal{P})$ in SOLAR is: (1) enumerate $Carc(\Sigma, \mathcal{P})$, and then (2) enumerate the *SOL tableau deductions* from $\Sigma \cup \{C\}$ with the top clause C [45] by removing each produced clause subsumed by some clause in $Carc(\Sigma, \mathcal{P})$.

3 Reasoning about Positive and Negative Causal Effects

So far, meta-level abduction has been defined for causal networks without explicitly arguing the meaning of causes. Indeed, links in a causal network have been of one kind, and $connected(g, s)$ at the meta level, i.e., $(g \leftarrow s)$ at the object level, just represents that g directly depends on s somehow. However, mixing different types of causalities in one type of links often makes analysis of actual causes complicated [48]. For example, suppose that increase of the amount of p decreases the amount of q and that increase of q causes increase of r. In this case, we cannot say that increase of p causes increase of r because q cannot mediate between p and r. For this problem, it is not appropriate to represent the causalities as $(p \rightarrow q \rightarrow r)$ because transitivity does not hold. Instead, $(inc(p) \rightarrow dec(q))$ and $(inc(q) \rightarrow inc(r))$ would be more precise but we need more entities and relations between $inc(\_)$ and $dec(\_)$. In this section, we consider one of the most important problems of this kind: networks with two types of causalities, i.e., *positive* and *negative* causal effects. With this regard, from now on we can understand that each arc of the form $connected(g, s)$ in Section 2 only represents positive effects.

We extend applicability of meta-level abduction to deal with networks expressing both positive and negative causal effects. Such networks are seen in biological domains, where *inhibition* effects negatively in gene regulatory, signaling and metabolic pathways. Now we consider two types of direct causal relations: $triggered$ and $inhibited$. For two nodes g and t, the relation $triggered(g, t)$ represents a positive cause such that t is a *trigger* of g, written as $g \longleftarrow t$ in a causal network. On the other hand, the relation $inhibited(g, s)$ represents a negative cause such that s is an *inhibitor* of g, written as $g \longdashv s$ in a causal network.[6] The meaning of these links will be given in two ways in Sections 3.1 and 3.2.

As in Section 2.1, negation, disjunctive effects and conjunctive causes can be defined for $triggered$ and $inhibited$, cf., (2), (3) and (4), and complex causal relations can be represented using those combinations and intermediate complexes. For instance, g is jointly triggered by t_1 and t_2 can be expressed as $triggered(g, t_1) \vee triggered(g, t_2)$.

The notion of causal chains is also divided into two types: the positive one (written $promoted$) and the negative one (written $suppressed$), respectively corresponding to $triggered$ and $inhibited$. Now our task is to design the axioms for these two meta-predicates.

3.1 Alternating Axioms for Causality

Suppose first that there is no inhibitor in a causal network, that is, all links are positive. In this case, the axioms for $promoted$ should coincide with (5) and (6):

$$promoted(X, Y) \leftarrow triggered(X, Y). \tag{12}$$

$$promoted(X, Y) \leftarrow triggered(X, Z) \wedge promoted(Z, Y). \tag{13}$$

Next, let us interpret the meaning of an inhibitor as a toggle switch of signals flowed in the inhibitor, just as an inverter in a logic circuit [62]. Then, in the presence of inhibitors, we need one more axiom (14), which blocks an adjacent inhibitor for X in order to promote X:

$$promoted(X, Y) \leftarrow inhibited(X, Z) \wedge suppressed(Z, Y). \tag{14}$$

[6] A trigger and an inhibitor are often called an *activator* and a *repressor*, respectively.

As for the axioms of the negative causal chain $suppressed$, we can consider the following axioms (15), (16) and (17), which are the counterpart of positive ones (12), (13) and (14):

$$suppressed(X, Y) \leftarrow inhibited(X, Y). \tag{15}$$
$$suppressed(X, Y) \leftarrow inhibited(X, Z) \wedge promoted(Z, Y). \tag{16}$$
$$suppressed(X, Y) \leftarrow triggered(X, Z) \wedge suppressed(Z, Y). \tag{17}$$

That is, a negative causal chain to X can be established if negative influence is propagated to X either directly by an adjacent inhibitor of an active (15) or promoted (16) item or indirectly by a trigger of suppressed (17). By this way, we can establish the connection between X and Y by mixing both positive and negative links.

One nice property with the axiomatization by (12–17) is that all possible paths from a source to a goal, which is either positive or negative, can be obtained by meta-level abduction. We here prove this property step by step.[7] Firstly, the correctness of the axiomatization by $(12–17)$ with respect to deduction is shown as follows. Suppose a meta-level theory N, which represents a causal network with triggers and inhibitors, and a clause at the meta-level $O^+ = promoted(g, s)$ (resp. $O^- = suppressed(g, s)$), where g and s are nodes in N. When O^+ (resp. O^-) can be proved from $N \cup \{(12–17)\}$, a *proof* Π of O^+ (resp. O^-) is a sub-network of N such that Π consists of triggers and inhibitors that connect paths from s to g. A proof Π of O^+ (resp. O^-) is also called a *positive* (resp. *negative*) *causal chain* from s to g in N. Then, each proof of O^+ or O^- is a sequence of links, $L_1, L_2, \ldots, L_m$, where m is the length of the chain and each L_j is either a trigger or an inhibitor in N such that the endpoint of L_1 is g, the start point of L_m is s, and the start point of L_j is the same as the endpoint of L_{j+1} for $j = 1, \ldots, m - 1$.

Lemma 1 *Let N be a causal network, and s and g be nodes in N, which respectively represent a source and a goal. There exists a positive (resp. negative) causal chain from s to g in N if and only if there exists a path from s to g in N such that there are an even (resp. odd) number of occurrences of inhibitors.*

Proof We here prove the only-if direction of the lemma, but the if direction can be proved in a similar way. Let Π^+ (resp. Π^-) be a positive (resp. negative) causal chain from s to g in N. We prove the only-if direction of the lemma by a double induction on the number k of occurrences of inhibitors in Π^+ or Π^-.

Firstly, consider the case of $k = 0$. In this case, Π^+ consists of atoms with the predicate $triggered$ only, each of which appears in either (12) or (13). Construction of Π^+ is indeed possible by Proposition 1 when we replace $triggered$ with $connected$ and axioms (12) and (13) with (5) and (6), respectively. On the other hand, if Π^- exists, it must contain at least one atom with $inhibited$ by (15) or (16). Hence there is no Π^- in this case.

Secondly, consider the case of $k = 1$. If Π^+ exists, then it must contain an inhibitor $inhibited(\chi, \psi)$ for some χ and ψ in N from a rule of the form (14) after some chain of triggers via (13). Then $suppressed(\psi, s)$ must be derived, but Π^+ cannot contain any other inhibitor by the assumption. By the case of $k = 0$, it is impossible to derive $suppressed(\psi, s)$ without using any inhibitor. Hence, there is no Π^+ in this case. On the other hand, Π^- must contain either an inhibitor $inhibited(\chi, s)$ for some node χ in N from a rule of the form (15) or an inhibitor $inhibited(\chi, \psi)$ for some χ and ψ in N from a rule of the form (16) after

[7] A preliminary version of the completeness property presented in [17, Proposition 3.1] presupposes that the observation is a unit clause of the form $suppressed(g, s)$ or $promoted(g, s)$, and does not properly distinguish inhibitors contained in the axiom set and inhibitors abduced in an explanation.

some chain of triggers via (17). In the latter case, $promoted(\psi, s)$ must be proved with no inhibitor, which is possible by the case of $k = 0$. In either case, Π^- exists in this case.

Next, in the case of $k = m \geq 1$, suppose the two propositions: (I) for any positive causal chain Π^+, the number m of occurrences of inhibitors in Π^+ is even; and (II) for any negative causal chain Π^-, the number m of occurrences of inhibitors in Π^- is odd.

Now, consider the case of $k = m + 1$. Then, Π^+ contains either a trigger from (13) or an inhibitor from (14). In the former case, the predicate of the other subgoal of (13) is $promoted$, but after some sequence of triggers, an inhibitor must be contained in Π^+ from (14). That is, at least one inhibitor is contained and the predicate of the rest of subgoals is $suppressed$, whose proof contains m inhibitors. In either case, by the induction hypothesis (II), m is odd so k is even. On the other hand, Π^- contains either an inhibitor from (16) or a trigger from (17). In the latter case, the predicate of the other subgoal of (17) is $suppressed$, but after some sequence of triggers, an inhibitor must be contained in Π^- from (16). That is, an inhibitor is encountered and the predicate of the other subgoals is $promoted$, whose proof contains m inhibitors. In either case, by the induction hypothesis (I), m is even, so k is odd. □

In axiomatization (12–17), the logical interpretations of $g \longleftarrow s$ and $g \longmapsto s$ in causal networks are defined as follows. A trigger $triggered(g, s)$ at the meta level is interpreted as $(g \leftarrow s) \wedge (\neg g \leftarrow \neg s)$ at the object level, which is now abbreviated as $(g \leftrightarrow s)$, then the truth value of s is preserved in the truth value of g with this trigger. On the other hand, an inhibitor $inhibited(g, s)$ at the meta level can be interpreted as $(\neg g \leftarrow s) \wedge (g \leftarrow \neg s)$ at the object level, which is now abbreviated as $(\neg g \leftrightarrow s)$, and the truth value of s is reversed in the truth value of g with this inhibitor. For any meta-level clausal theory N such that the predicate of each clause appearing in N is either $triggered$ or $inhibited$, let $\lambda(N)$ be the object-level theory obtained by replacing (i) every $triggered(t_1, t_2)$ (t_1 and t_2 are terms) appearing in N with the formula $(t_1 \leftrightarrow t_2)$ and (ii) every $inhibited(t_1, t_2)$ appearing in N with the formula $(\neg t_1 \leftrightarrow t_2)$.

Proposition 2 *Suppose a meta-level theory N, which consists of disjunctions of literals of the form $triggered(\_, \_)$ and disjunctions of literals of the form $inhibited(\_, \_)$. Let the background theory be $B = N \cup \{(12\text{–}17)\}$. Then,* (1) $B \models (promoted(g, s_1) \vee \cdots \vee promoted(g, s_n))$ *if and only if* $\lambda(N) \models (g \leftarrow s_1 \wedge \cdots \wedge s_n)$*; and* (2) $B \models (suppressed(g, s_1) \vee \cdots \vee suppressed(g, s_n))$ *if and only if* $\lambda(N) \models (\neg g \leftarrow s_1 \wedge \cdots \wedge s_n)$.

Proof Extending the proof of Proposition 1, we can prove that any (partial) proof Π_i of $promoted(g_i, s_i)$ or $suppressed(g_i, s_i)$ ($g_i \neq s_i$) from B contains m occurrences of inhibitors from N iff $\lambda(N) \models (\underbrace{\neg(\neg(\cdots\neg}_{m}(g_i))) \leftarrow s_i)$.[8] By the result and the proof of Lemma 1, such a proof of $promoted(g_i, s_i)$ (resp. $suppressed(g_i, s_i)$) ($g_i \neq s_i$) from B contains an even (resp. odd) number of occurrences of inhibitors from N. Then, (1) $B \models promoted(g_i, s_i)$ iff $\lambda(N) \models (\underbrace{\neg(\neg(\cdots\neg}_{m:\text{even}}(g_i))) \leftarrow s_i)$ iff $\lambda(N) \models (g_i \leftarrow s_i)$. On the other hand, (2) $B \models suppressed(g_i, s_i)$ iff $\lambda(N) \models (\underbrace{\neg(\neg(\cdots\neg}_{m:\text{odd}}(g_i))) \leftarrow s_i)$ iff $\lambda(N) \models (\neg g_i \leftarrow s_i)$. The proposition follows by applying these results to disjuncts of goal clauses. □

[8] To see this, if there is no inhibitor in Π_i, $(g_i \leftrightarrow s_i)$ holds. If there is one inhibitor in Π_i, only one change of negation is introduced in a chain from s_i, which is then propagated to g_i, i.e., $(\neg g_i \leftrightarrow s_i)$. Similarly, an additional introduction of an inhibitor i Π_i adds one more negation at g_i.

Proposition 2 states that the axiomatization by (12–17) is sound and complete for enumerating positive and negative causal chains. Extending deductive inference to abductive inference, meta-level abduction on causal networks with positive and negative links can now be defined by letting the abducibles Γ be those atoms with the predicates $triggered$ and/or $inhibited$: $\Gamma = \{triggered(\_, \_),\ inhibited(\_, \_)\}$, and observations or goals are given as (disjunctions of) literals either of the form $promoted(g, s)$ or of the form $suppressed(g, s)$. Hence, given positive and negative observations, we can abduce both positive and negative causes, and new nodes are produced whenever necessary.

Theorem 2 *Let N and B be the same theories as in Proposition 2. Let the abducibles be $\Gamma = \{triggered(\_, \_), inhibited(\_, \_)\}$. Then, H^+ is an abductive explanation of $O^+ = (promoted(g, s_1) \vee \cdots \vee promoted(g, s_n))$ with respect to B and Γ if and only if $\lambda(H^+)$ is a hypothesis satisfying that*

$$\lambda(N) \cup \lambda(H^+) \models (g \leftarrow s_1 \wedge \cdots \wedge s_n) \tag{18}$$

and $\lambda(N) \cup \lambda(H^+)$ is consistent. On the other hand, H^- is an abductive explanation of $O^- = (suppressed(g, s_1) \vee \cdots \vee suppressed(g, s_n))$ with respect to B and Γ if and only if $\lambda(H^-)$ is a hypothesis satisfying that

$$\lambda(N) \cup \lambda(H^-) \models (\neg g \leftarrow s_1 \wedge \cdots \wedge s_n) \tag{19}$$

and $\lambda(N) \cup \lambda(H^-)$ is consistent.

Proof The equivalence between the meta-level abductive entailment $B \cup H^+ \models O^+$ and the object-level abductive entailment (18) holds by Proposition 2 (1). Similarly, the equivalence between the relation $B \cup H^- \models O^-$ and relation (19) holds by Proposition 2 (2). The equivalence between the consistency of $B \cup H^+$ (resp. $B \cup H^-$) and the consistency of $\lambda(N) \cup \lambda(H^+)$ (resp. $\lambda(N) \cup \lambda(H^-)$) also holds: Since $B \cup H^+$ (resp. $B \cup H^-$) is consistent, so is $\lambda(N) \cup \lambda(H^+)$ (resp. $\lambda(N) \cup \lambda(H^-)$). □

Theorem 2 shows the soundness and completeness of axiomatization (12–17) for meta-level abduction of positive and negative causal effects. Moreover, the consistency of the background theory $B = N \cup \{(12\text{–}17)\}$ in Theorem 2 also holds. However, both $promoted(g, s)$ and $suppressed(g, s)$ can be explained from this B for some N at the same time. This "semantic inconsistency" is, however, inevitable in this monotonic representation, since we can answer to any virtual query supposing a source s and a goal g. Then, to prevent derivations of promotion and suppression simultaneously for the same s and g, the following integrity constraint can be placed at the meta level.

$$\leftarrow promoted(X, Y) \wedge suppressed(X, Y). \tag{20}$$

The role of (20) is to derive (minimal) nogoods, i.e., a (minimal) set of incompatible instances of abducibles. Any abductive explanation must not include any nogood. For example, given the network

$$N_1 = \{triggered(g, t)\},$$

the minimal nogoods containing at most 2 abducibles are the following 13 clauses.[9]

[9] This enumeration is verified by SOLAR. Recall (Section 2.4) that a minimal nogood can be constructed as the negation of a characteristic clause of the background theory B, which is now the union of the causal network N at the meta-level and axioms (12–17) with constraint (20). Note that the number of (minimal) nogoods becomes infinite if we do not restrict the number of abducibles appearing in each nogood.

$\{inhibited(g,t)\}$, $\{inhibited(t,g)\}$, $\{inhibited(\chi,\chi)\}^*$,
$\{triggered(g,\chi), inhibited(\chi,t)\}$, $\{inhibited(g,\chi), triggered(\chi,t),\}$,
$\{triggered(t,\chi), inhibited(g,\chi)\}$, $\{triggered(\chi,\psi), inhibited(\psi,\chi)\}^*$,
$\{triggered(g,\chi), inhibited(t,\chi)\}$, $\{triggered(t,\chi), inhibited(\chi,g)\}$,
$\{triggered(\chi,g), inhibited(\chi,t)\}$, $\{triggered(\chi,\psi), inhibited(\chi,\psi)\}^*$,
$\{inhibited(\chi,g), triggered(\chi,t),\}$, $\{triggered(\chi,g), inhibited(\chi,t)\}$,

where χ and ψ are existentially quantified variables. A set with the symbol "$*$" does not refer constants g, t, and is also a minimal nogood of the empty network $N_0 = \emptyset$. This implies that any causal network including an instance of such a set with "$*$" becomes inconsistent. Hence, introduction of (20) into axiom set (12–17) can make a causal network inconsistent, that is, *the empty set $\emptyset$ can become a nogood*. For example, given the network

$$N_2 = \{triggered(g,t), inhibited(g,s), triggered(s,t)\}, \tag{21}$$

$N_2 \cup \{(12\text{–}17)\}$ is consistent. However, $N_2 \cup \{(12\text{–}17), (20)\}$ is inconsistent, since both $promoted(g,t)$ (by (12)) and $suppressed(g,t)$ (by (16) and (12)) hold. Also, the p53 network (47), which will be given later in Section 4.1, becomes inconsistent if it is combined with this axiomatization $\{(12\text{–}17), (20)\}$. The (in)consistency of a theory with axioms (12–17) together with constraint (20) is characterized as follows.

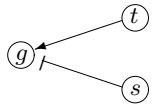
Proposition 3 *Let N be the same as in Proposition 2. Let the background theory be $B = N \cup \{(12\text{–}17), (20)\}$. Then, B is consistent if and only if there exist no nodes g and s in N such that there are causal chains Π^+ and Π^- from s to g in N such that there are an even (resp. odd) number of occurrences of inhibitors in Π^+ (resp. Π^-).*

Proof If the condition is satisfied, Π^+ is a proof of $promoted(g,s)$ and Π^- is a proof of $suppressed(g,s)$. Then B becomes inconsistent by (20). Conversely, if B is inconsistent, both $promoted(g,s)$ and $suppressed(g,s)$ can be proved from B for some nodes g and s in N by (20). Let Π^+ be a proof of $promoted(g,s)$, and Π^- a proof of $suppressed(g,s)$. Then, Π^+ (resp. Π^-) contains an even (resp. odd) number of inhibitors by Lemma 1. □

We next show a modification of this monotonic axiomatization to reduce such semantic inconsistencies using default assumptions.

3.2 Axiomatization with Default Assumptions

Suppose two antagonistic direct causal relations appear simultaneously for the same node g as follows.



$triggered(g,t)$
$inhibited(g,s)$ (22)

Our intuition on diagram (22) is as follows. (I) If the trigger t is present and the inhibitor s is not present, then g is triggered by t (and is not inhibited); (II) Else if s is present and t is not present, then g is inhibited by s (and is not triggered). These two cases are rather clear, but what happens for g if both t and s are somehow caused? Is it triggered or inhibited? The last axiomatization in Section 3.1 concludes that g is both promoted and suppressed by s, which leads to inconsistency under the existence of the integrity constraint of the form (20). However, often it is indicated that: (III) If both t and s are present, then g is inhibited by s

(and is not triggered). Namely, *an inhibitor is preferred to a trigger*.[10] This preference rule is often used in applications, particularly in the biological literature.[11]

The last inference is *nonmonotonic*: a trigger of g works *if there is no inhibitor for* g, but if an inhibitor is added then the trigger stops working. We now show another axiomatization which reflects this principle of inhibitor preference. Depart from the monotonic axiomatization of causal chains (12–17), reasoning about networks is now made nonmonotonic and involves *default reasoning*. In the following new definitions of $promoted$ and $suppressed$, we will associate an extra condition for each trigger to work.

$$promoted(X,Y) \leftarrow triggered(X,Y) \wedge no\_inhibitor(X). \quad (23)$$
$$promoted(X,Y) \leftarrow triggered(X,Z) \wedge no\_inhibitor(X) \wedge promoted(Z,Y). \quad (24)$$
$$promoted(X,Y) \leftarrow inhibited(X,Z) \wedge suppressed(Z,Y). \quad (25)$$
$$suppressed(X,Y) \leftarrow inhibited(X,Y). \quad (26)$$
$$suppressed(X,Y) \leftarrow inhibited(X,Z) \wedge promoted(Z,Y). \quad (27)$$
$$suppressed(X,Y) \leftarrow triggered(X,Z) \wedge no\_inhibitor(X) \wedge suppressed(Z,Y). \quad (28)$$
$$\leftarrow promoted(X,Y) \wedge suppressed(X,Y). \quad (29)$$

The new axiom set for positive and negative causal chains (23–28) are the same as the monotonic version (12–17), except that each trigger to X ($triggered(X,\_)$) must not be inhibited ($no\_inhibitor(X)$) to give the positive effect to X in (23), (24) and (28).[12] Here, inclusion of $no\_inhibitor(\chi)$ in association with $triggered(\chi,\psi)$ makes those three axioms *default rules* [57,49]. The literal of the form $no\_inhibitor(\_)$ is thus treated as a *default*, which can be assumed during inference unless contradiction occurs, and constraints can also be added to reject inconsistent cases with these assumptions. Then, a trigger $triggered(g,s)$ at the meta level is now interpreted as $(no\_inhibitor(g) \rightarrow (g \leftrightarrow s))$ at the object level, that is, $(g \leftrightarrow s)$ is true if the default $no\_inhibitor(g)$ is consistent with the union of the background theory B and a constructing hypothesis H. On the other hand, an inhibitor $inhibited(g,s)$ at the meta level can be interpreted as $(\neg g \leftrightarrow s)$ at the object level, hence the effect of an inhibitor does not need addition of a default. Finally, the integrity constraint of the form (29) is the same as (20) and prohibits the presence of both positive and negative causes between any pair of nodes X and Y.

Meta-level abduction under axiomatization (23–28) and the integrity constraint of the form (29) is defined in the same way as in Section 3.1. Abduction of *joint triggers* and *joint inhibitors* can also be realized in the same way as $jointly\_connected$ in (7) by adding the meta-level axioms

$$triggered(X,Y) \vee triggered(X,Z) \leftarrow jointly\_triggered(X,Y,Z). \quad (30)$$
$$inhibited(X,Y) \vee inhibited(X,Z) \leftarrow jointly\_inhibited(X,Y,Z). \quad (31)$$

[10] In some formalization of inhibitors [32,9], an inhibitor is modeled to work against a trigger/inhibitor edge between nodes by blocking its activation/inhibition function. This results in the same effect as preferring inhibitors to triggers when an inhibitor works against activation of a trigger. Blocking inhibition by an inhibitor has the same effect as double inhibition in our formalization (see Remark 2).

[11] In some applications, it is possible to consider the preference for a trigger over an inhibitor. All theories explored in this section can be easily adapted in accordance with this reverse preference by just exchanging the roles of positive and negative causal relations.

[12] Instead of (26), $(suppressed(X,Y) \leftarrow inhibited(X,Y) \wedge no\_inhibitor(Y))$ was used in an earlier version of this paper (`http://ilp2010.dsi.unifi.it/pdf/ilp2010_submission_43.pdf`). This axiom was only applied at an inhibitor to X to which $no\_inhibitor(Y)$ is assumed. Although this is a possible formalization, it introduced an inelegant asymmetry between positive and negative influences.

to the background theory B and the literals of the form $jointly\_triggered(\_,\_,\_)$ and of the form $jointly\_inhibited(\_,\_,\_)$) to the abducibles Γ. Note that one of definitions (30,31) of these meta-predicates can be omitted by simulating one by the other. For example, abduction of $jointly\_inhibited(g,s,t)$ can be simulated by abduction of $inhibited(g,s\text{-}t) \wedge jointly\_triggered(s\text{-}t,s,t)$, where s-t is an intermediate complex.

Remark 1 Unlike axiom (7) of $jointly\_connected$ in Section 2.2, we can propagate negative effects from sources to goals through joint triggers and joint inhibitors by the axioms:

$$suppressed(G,X) \leftarrow jointly\_triggered(G,A,B) \wedge suppressed(A,X). \quad (32)$$
$$suppressed(G,X) \leftarrow jointly\_triggered(G,A,B) \wedge suppressed(B,X). \quad (33)$$
$$promoted(G,X) \leftarrow jointly\_inhibited(G,A,B) \wedge suppressed(A,X). \quad (34)$$
$$promoted(G,X) \leftarrow jointly\_inhibited(G,A,B) \wedge suppressed(B,X). \quad (35)$$

Axioms (32) and (33) show that suppression of one of the joint triggers causes suppression of the goal, while axioms (34) and (35) show that suppression of one of the joint inhibitors causes promotion of the goal. All these inverse effects can be caused by the suppression of one of the joint links regardless of the status of the other link.

As for default assumptions of the form $no\_inhibitor(\_)$, *default reasoning* can be implemented by assuming those literals whenever necessary during inference, and the consistency of such assumptions is checked each time they are added to the current set of abduced literals. This is a simple yet powerful method for default reasoning in the case of so-called *normal defaults (without prerequisites)* [57,49]. Hence, the abducibles Γ now also contain the literals of the form $no\_inhibitor(\_)$ along with $triggered(\_,\_)$ and $inhibited(\_,\_)$. For example, the abducibles allowing joint triggers can be defined as

$$\Gamma = \{\, triggered(\_,\_),\ inhibited(\_,\_),\ jointly\_triggered(\_,\_,\_),\ no\_inhibitor(\_)\,\}.$$

Note that addition of default literals through the abducibles $\{no\_inhibitor(\_)\}$ is necessary not only for abduction with defaults but for deduction involving defaults.

The soundness and completeness of meta-level abduction with the new axiomatization is guaranteed as in Theorem 2. Let M be a meta-level clausal theory such that the predicate of any literal appearing in M is either $triggered$, $inhibited$, or $no\_inhibitor$. Now $\lambda^\dagger(M)$ is defined as the object-level theory obtained by replacing (i) every $triggered(t_1,t_2)$ (t_1 and t_2 are terms) appearing in M with the formula $((t_1 \leftrightarrow t_2) \wedge t_1^*)$ (t_1^* is a new term uniquely associated with t_1), (ii) every $inhibited(t_1,t_2)$ appearing in M with the formula $(\neg t_1 \leftrightarrow t_2)$, and (iii) every $no\_inhibitor(t)$ appearing in M with t^*. Note that t^* represents that t is not inhibited by default.

Theorem 3 *Suppose a meta-level theory N, which consists of disjunctions of literals of the form $triggered(\_,\_)$ and disjunctions of literals of the form $inhibited(\_,\_)$. Let the background theory be $B = N \cup \{(23\text{–}28)\}$. Let the abducibles be*

$$\Gamma = \{triggered(\_,\_), inhibited(\_,\_), no\_inhibitor(\_)\}.$$

Then, H^+ is an abductive explanation of $O^+ = (promoted(g,s_1) \vee \cdots \vee promoted(g,s_n))$ with respect to B and Γ if and only if $\lambda^\dagger(H^+)$ is a hypothesis satisfying that

$$\lambda^\dagger(N) \cup \lambda^\dagger(H^+) \models (g \leftarrow s_1 \wedge \cdots \wedge s_n) \quad (36)$$

and $\lambda^\dagger(N) \cup \lambda^\dagger(H^+)$ *is consistent. On the other hand,* H^- *is an abductive explanation of* $O^- = (suppressed(g, s_1) \vee \cdots \vee suppressed(g, s_n))$ *with respect to* B *and* Γ *if and only if* $\lambda^\dagger(H^-)$ *is a hypothesis satisfying that*

$$\lambda^\dagger(N) \cup \lambda^\dagger(H^-) \models (\neg g \leftarrow s_1 \wedge \cdots \wedge s_n) \tag{37}$$

and $\lambda^\dagger(N) \cup \lambda^\dagger(H^-)$ *is consistent.*

Proof Each explanation H obtained with the monotonic axioms (12–17) in Theorem 2 can also be obtained by incorporating defaults of the form $no\_inhibitor(t)$ in the corresponding explanation H^* with axioms (23–28), which are translated to atoms of the form t^* in $\lambda^\dagger(H^*)$. Then, H satisfies (18) (resp. (19)) iff H^* satisfies (36) (resp. (37)). □

A background theory B can also include several knowledge about defaults, and here are some examples. First, when we are sure that there is no inhibitor for a node s, we can include the fact $no\_inhibitor(s)$ in the background theory B. For instance, $no\_inhibitor(s)$ can be declared to be true in B if s is a *terminal source node*:

$$no\_inhibitor(X) \leftarrow source(X), \tag{38}$$

where each source s is declared as $source(s)$ in B too. Such a terminal source node has the same effect as an abducible fact $abd(s)$ in Section 2.3. We will see an example to use $source(\_)$ and constraint (38) in Section 4.1. Second, a meta-level constraint:

$$\leftarrow no\_inhibitor(X) \wedge inhibited(X, Y). \tag{39}$$

blocks to assume a default $no\_inhibitor(g)$ for any node g to which an inhibitor is connected. This constraint is natural in many cases, but is only optional because $inhibited(g, k)$ and $no\_inhibitor(g)$ may coexist when, for example, there are multiple supports through derivations of $promoted(g, h_1), \ldots, promoted(g, h_n)$ $(n \geq 2)$ that share $no\_inhibitor(g)$, yet only one $inhibited(g, k)$ exists.[13]

Remark 2 Introduction of (39) prunes a hypothesis containing $triggered(g, \_)$ whenever an inhibitor $inhibited(g, \_)$ exists. For example, for the causal network

$$N_3 = \{triggered(g, t), inhibited(g, s)\},$$

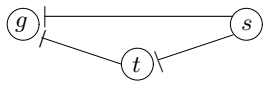
and the observation $O_2 = promoted(g, t)$, the set $E = \{no\_inhibitor(g)\}$ can be obtained by (23) to explain O_2 but is not consistent with (39), so is not an explanation of O_2. Instead, $F = \{inhibited(s, t)\}$ is an explanation of O_2 obtained by (25) and (26), i.e., $g \longmapsto s \longmapsto t$. Here, we can again see that an inhibitor is preferred to a trigger; for activation of g in the presence of an inhibitor of g, introduction of a new trigger of g is not enough, but *suppression of the existing inhibitor is really effective*. This is an example of how *double inhibitions* works as a promoter, which is actually seen in many biological systems, e.g., the p53 network (47) given in Section 4.1.

[13] This is a matter of majority when there are more triggers than inhibitors for the same node. In such a case, we could still prefer an inhibitor by introducing constraint (39), but alternatively might want to allow promotion by omitting this constraint. See also the discussion in Section 6.5.

3.3 Consistency of Nonmonotonic Axiomatization

When constraint (29) is incorporated into the background theory B containing a causal network N and new axioms (23–28), the consistency of B is "more" guaranteed than in the case of Section 3.1. This is because, unlike axioms (12–17), the new axioms contain additional defaults of the form $no\_inhibitor(\_)$. For example, the causal network N_2 (21) is inconsistent with the previous axiom set $\{(12\text{–}17), (20)\}$, but $B_2 = N_2 \cup \{(23\text{–}28), (29)\}$ is now consistent. Still, both $promoted(g, t)$ and $suppressed(g, t)$ can be explained from B_2 and Γ, but their explanations are not the same: $\{no\_inhibitor(g)\}$ explains the former, while $\{no\_inhibitor(s)\}$ explains the latter. Again, the role of (29) is to identify each nogood to prune any incompatible combination of defaults and abducibles. That is, abducing literals with the predicates $triggered$ and $inhibited$ involves default assumptions of the form $no\_inhibitor(\_)$, and any inconsistent set of abducibles can be detected by subsumption checking with nogoods. In N_2, $\{no\_inhibitor(g), no\_inhibitor(s)\}$ becomes a nogood.

However, there are some networks that can still be proved inconsistent under the new axiom set. Suppose the network



$$N_4 = \{inhibited(g, s), inhibited(g, t), inhibited(t, s)\}, \tag{40}$$

where the first inhibitor $inhibited(g, s)$ is a suppressor for g (by (26)) and the latter two consecutive inhibitors work as a promoter for g by (25) and (26). Since there is no trigger in N_4, no default is involved for proving $promoted(g, s)$ and $suppressed(g, s)$, and neither goal can be preferred to the other. Hence, N_4 is proved inconsistent by (29).

The necessary and sufficient condition for the consistency of a meta-level background theory can be characterized by networks with multiple "inhibitor-only paths" as follows.

Proposition 4 *Suppose a meta-level theory N, which is the same as in Theorem 3. Let the background theory be $B = N \cup \{(23\text{–}28), (29)\}$. Then, B is consistent if and only if there are no nodes g and s in N such that there exist a proof Π^+ of $promoted(g, s)$ and a proof Π^- of $suppressed(g, s)$ satisfying that* (i) *both Π^+ and Π^- have no triggers;* (ii) *there are an even number of occurrences of inhibitors in Π^+; and* (iii) *there are an odd number of occurrences of inhibitors in Π^-.*

Proof Suppose a proof Π^+ of $promoted(g, s)$ and a proof Π^- of $suppressed(g, s)$ satisfying the conditions (i), (ii) and (iii). Then, by (29), B becomes inconsistent. Conversely, if B is inconsistent, both $promoted(g, s)$ and $suppressed(g, s)$ can be proved from B by (29). However, those proofs Π^+, Π^- do not use axioms (23), (24) and (28), since the literal of the form $no\_inhibitor(\_)$ cannot be proved. Hence, no trigger appears in either Π^+ or Π^-, and (i) holds. By Lemma 1, (ii) and (iii) hold. □

It is easy to see that the causal network N_4 (40) is inconsistent with $\{(23\text{–}28), (29)\}$ by Proposition 4. Actually, N_4 is a minimal inconsistent network under axioms (23–28) and (29), since $\{inhibited(\gamma, \psi), inhibited(\gamma, \tau), inhibited(\tau, \psi)\}$ is a minimal nogood of $N_0 = \emptyset$ with those axioms. Note that the smallest inconsistent network is

$$N_5 = \{inhibited(g, g)\}. \tag{41}$$

In N_5, a self-loop on g suppresses g if it passes the inhibitor an odd number of times, while it promotes g if it counts the inhibitor an even number of times. Remark 3 below suggests

a minimal modification of the axiom set to guarantee the consistency, and Section 6.5 will discuss how to resolve such inconsistencies in general.

Since inhibitors are preferred to triggers, abduction of promotion is blocked by abduction of suppression if the explanation of the former is subsumed by that of the latter.

Proposition 5 *Let N, B, and g, s be the same as in Proposition 4. Suppose the abducibles $\Gamma = \{triggered(\_,\_), inhibited(\_,\_), no\_inhibitor(\_)\}$ and the observations $O^+ = promoted(g,s)$ and $O^- = suppressed(g,s)$. If E^+ is an explanation of O^+ with respect to B and Γ, then there is no explanation E^- of O^- with respect to B and Γ such that* (i) *all defaults of the form $no\_inhibitor(\_)$ in E^- are included in E^+ and* (ii) *all inhibitors in E^- are included in E^+.*

Proof Suppose that E^+ is an explanation of O^+. Assume further that an explanation E^- of O^- satisfying the conditions of the theorem exists. Since O^+ and O^- violate constraint (29), $E^+ \cup E^-$ is a nogood. The condition (i) implies that the triggers in E^- are also included in E^+ because abduction of a trigger always involves a default by axioms (23), (24) and (28). By this and the condition (ii), $E^- \subseteq E^+$ holds. Then, $E^+ \cup E^- = E^+$ holds, and thus E^+ is a nogood. This contradicts with the supposition that E^+ is an explanation of O^+. □

Proposition 5 implies that, whenever both $promoted(g,s)$ and $suppressed(g,s)$ are explained, their explanations must contain different triggers and inhibitors, and the union of the two explanations is a nogood. For example, for the network

$$N_6 = \{triggered(g,s), inhibited(g,s)\}. \tag{42}$$

$\{no\_inhibitor(g)\}$ cannot be an explanation of $promoted(g,s)$ because it is a superset of $\emptyset$, an explanation of $suppressed(g,s)$, so that $\{no\_inhibitor(g)\}$ becomes a nogood. This is a typical effect of preferring inhibitors.

Remark 3 Proposition 4 captures the nature of inconsistent networks under axioms (23–28) and constraint (29), which minimizes the inconsistent cases only in the existence of two nodes connected by multiple inhibitor chains with odd and even numbers of occurrences of inhibitors. If we would further like to resolve inconsistencies in such cases, we could attach another default to axiom (25) as

$$promoted(X,Y) \leftarrow inhibited(X,Z) \wedge suppressed(Z,Y) \wedge \textit{unsuppressed}(X), \tag{43}$$

and replace (25) by (43). Here, a literal of the form *unsuppressed*$(\_)$ can be assumed whenever (43) is applied by putting it into the abducibles Γ. Then, any even number of occurrences of inhibitors must contain defaults of the form *unsuppressed*$(\_)$, but an odd number of inhibitor occurrences may not need such a default. Hence, no network can be inconsistent under axioms (23,24), (43), (26–28) and constraint (29). For example, in networks (40) and (41), the minimal nogood is obtained as $\{$*unsuppressed*$(g)\}$, which implies that only suppression is achieved on g in both cases, resulting in preferring inhibitors again.

However, this inconsistency resolution is not always welcome in real applications. A *negative feedback loop* like (41) is known to cause periodic oscillation between positive and negative states, and then it is harmful to completely ignore the possibility of promotion in such a case. This problem can only be solved by introducing time in the axioms, and will be discussed in Section 6.5. So far, we have a choice between (25) and (43), depending on our strategy to allow or disallow minimal inconsistencies in inhibitor-only paths.

Table 1 Correspondence between object-level inference and meta-level consequence finding

object-level inference	top clause in SOLAR	production field
rule verification	$\leftarrow caused(g, s).$	$\langle\emptyset\rangle$
fact abduction	$ans(X) \leftarrow caused(g, X).$	$\langle\{ans(\_)\}\rangle$
fact prediction	$ans(X) \leftarrow caused(X, s).$	$\langle\{ans(\_)\}\rangle$
rule prediction	unspecified	$\langle\{promoted(\_,\_), suppressed(\_,\_)\}\rangle$
rule abduction	$\leftarrow caused(g, s).$	$\langle\{\neg triggered(\_,\_), \neg inhibited(\_,\_)\}\rangle$
abducing rules and facts	$ans(X) \leftarrow caused(g, X).$	$\langle\{\neg triggered(\_,\_), \neg inhibited(\_,\_), ans(\_)\}\rangle$
fact prediction + rule abduction	$ans(X) \leftarrow caused(X, s).$	$\langle\{\neg triggered(\_,\_), \neg inhibited(\_,\_), ans(\_)\}\rangle$
rule prediction + rule abduction	unspecified	$\langle\{\neg triggered(\_,\_), \neg inhibited(\_,\_), promoted(\_,\_), suppressed(\_,\_)\}\rangle$

3.4 Various Inferences on Causal Networks

As in Section 2.3, axiomatization in Sections 3.1 and 3.2 can be further combined with fact abduction to allow mixed forms of inferences. For example, to abduce a source node which promotes (resp. suppresses) some goal node g, clause (10) can be rewritten to clause (44) (resp. (45)):

$$ans(X) \leftarrow promoted(g, X) \wedge source(X). \tag{44}$$

$$ans(X) \leftarrow suppressed(g, X) \wedge source(X). \tag{45}$$

Here, a source node represented by X corresponds to an abducible with the predicate $abd(\_)$, and both the predicates $promoted(\_,\_)$ and $suppressed(\_,\_)$ correspond to the predicate $caused(\_,\_)$ in (10). Then a source node at the object level is obtained as a ground term of X by answer extraction with these formulas. Combination of rule abduction and fact abduction can also be realized using consequence-finding techniques as in Section 2.3.

On the other hand, those nodes that can be promoted or suppressed by some source node s can be obtained by answer extraction for variable Y in the query of the form $promoted(Y, s)$ or $suppressed(Y, s)$. This last inference is called *prediction* rather than abduction, and can be realized by consequence finding too. Table 1 summarizes the correspondence between object-level and meta-level inferences. All types of meta-level inferences, possibly involving generation of existentially quantified hypotheses, can be realized by SOLAR. Recall that, in the context of inverse entailment (11), the negation of a goal is set as a *top clause*, and the negation of each abducible is given in a production field in SOLAR. In Table 1, " $\leftarrow caused(\_,\_)$" in a "top clause" column is instantiated by either " $\leftarrow promoted(\_,\_)$" or " $\leftarrow suppressed(\_,\_)$", and "$ans(\_)$" is an answer predicate to collect answer substitutions. In abducing object-level facts, a top clause can be further conditioned with the abducible literal "$abd(X)$" after $caused(g, X)$ if the list of abducibles is given in the background theory. Like (44) and (45), abducibles are often represented as literals of the form $source(\_)$.

In Table 1, "Rule verification" verifies if a given causal chain can be derived or not. "Fact prediction" computes *ramification* of a source s, i.e., to derive facts that can be caused by s. "Rule prediction" enumerates possible causal chains derivable from the given causal network, so a top clause is not provided in this case and characteristic clauses with respect to $\langle\{promoted(\_,\_), suppressed(\_,\_)\}\rangle$ are just computed. The last three are mixed forms of inference with the conditional format; rules and facts are abduced or predicted under some completed rules by rule abduction. A complex example of abducing rules and facts will

be shown in Section 4.1. The soundness and completeness of those meta-level inferences by SOLAR with respect to object-level inferences listed in Table 1 are guaranteed due to completeness of consequence finding in SOLAR [13,45] and completeness of (conditional) answer extraction [23,20].

Abduction with default assumptions is also implemented in SOLAR [20]. Membership of a clause F in an *extension* of a default theory [57,49][14] is guaranteed for each obtained consequence

$$F \leftarrow no\_inhibitor(t_1) \wedge \cdots \wedge no\_inhibitor(t_m) \tag{46}$$

if $\{no\_inhibitor(t_1), \ldots, no\_inhibitor(t_m)\}$ is not a nogood [20, Theorem 4.5]. In our meta-level abduction, F is a clause consisting of literals of such forms as $\neg triggered(\_, \_)$, $\neg inhibited(\_, \_)$ and $ans(\_)$. Given a top clause C and a production field $\mathcal{P}$, SOLAR outputs the new characteristic clauses $Newcarc(B, C, \mathcal{P})$, and hence each new produced clause is always checked if it is not subsumed by any nogood in $Carc(B, \mathcal{P})$ (see Section 2.4). Hence, each clause F in a consequence of the form (46) belongs to some extension.

4 Case Study: p53 Signal Networks

In this section, we will see that meta-level abduction can be well applied to completion of *signaling networks*. The importance of network completion in signaling networks has been recognized, since it is hard to observe activity levels and quantities of proteins in living organisms [1].[15] Moreover, reporter proteins/genes are usually employed in signaling pathways, but designing and introducing reporter proteins are hard tasks. This is contrasted to the case of genetic networks, in which expression levels of most genes can be observed using DNA microarray/chip technologies.

As case studies of meta-level abduction, we consider two signaling networks containing the *p53 protein* [50], but use it for different purposes. Although these networks are rather simple, they illustrate one of the most fundamental inference problems in Systems Biology: Given an incomplete biological network N, infer possible connections to promote or suppress the function associated with a node in N. We consider two target functions: suppression of tumors in cancer [65] (Section 4.1) and switching DNA synthesis on and off [62] (Section 4.2). Both problems are examples of *goal-oriented abduction*, in which additional links and nodes are abduced to realize given goals rather than to explain observed data.

4.1 Enumerating Tumor Suppressors

This subsection examines the p53 signal network presented in [65] by meta-level abduction. The p53 protein plays a central role as a tumor suppressor and is subjected to tight control through a complex mechanism involving several proteins [36].

The p53 protein has the transactivator domain, which bounds to the promoters of target genes, then leads to protect the cell from cancer. The level and activity of p53 in the cell is

[14] Simply, a *(normal) default theory* is a pair (B, D), where B and D are sets of first-order formulas representing a background theory and a set of defaults, respectively. An *extension* of a default theory (B, D) is the set of all consequences of $B \cup S$, where S is a maximal subset of D such that $B \cup S$ is consistent. Then, a formula F is considered as a rational consequence of (B, D) if F belongs to some extension of (B, D).

[15] This feature of non-observability also exists in completing causal networks with respect to physical skills for cello playing [18], in which internal causation in human bodies cannot be observed directly.

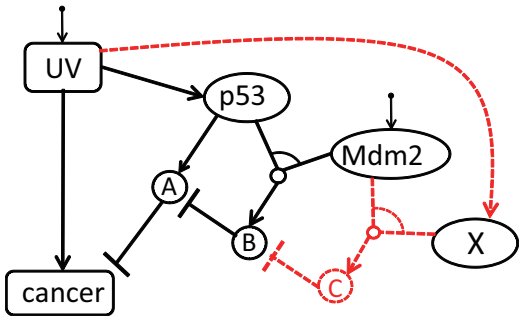


Fig. 1 Causal network of the p53 pathway

influenced by its interactions with other proteins. Tumor suppression is enabled if the interacting partners of p53 do not inhibit the functionality of the transactivator domain. Mdm2 binds to the transactivator domain of p53, thus inhibiting the p53 from tumor suppression. UV (ultraviolet light) causes stress, which may induce the upregulation of p53. However, stress can also influence the growth of tumors.

These relations can be represented in solid lines of the causal network in Fig. 1. The corresponding formulas at the meta level can be simply represented by the clauses:

$$\begin{aligned} &triggered(\texttt{cancer},\texttt{uv}), \quad triggered(\texttt{p53},\texttt{uv}),\\ &inhibited(\texttt{cancer},a), \quad triggered(a,\texttt{p53}),\\ &inhibited(a,b), \quad jointly\_triggered(b,\texttt{p53},\texttt{mdm2}), \end{aligned} \tag{47}$$

where a (“A” in Fig. 1) is the inhibitory domain of p53, and b (“B” in Fig. 1) is the complex p53-mdm2. The unit clause $jointly\_triggered(b,\texttt{p53},\texttt{mdm2})$ can be replaced by the clause $(triggered(b,\texttt{p53}) \vee triggered(b,\texttt{mdm2}))$.

As in the setting of [65], we consider a tumor suppressor gene X (e.g., MdmX) such that mutants of X are highly susceptible to cancer. Suppose in some experiments that exposure of the cell to high level UV does not lead to cancer, given that the initial concentration of Mdm2 is high. Those initial conditions are represented as two facts,

$$source(\texttt{uv}), \quad source(\texttt{mdm2}), \tag{48}$$

that is, both UV and Mdm2 can be abduced whenever necessary.[16] The meta-predicate $source$ thus behaves like the abducible predicate abd. Some other meta-level axioms can be introduced, e.g., $(no\_inhibitor(S) \leftarrow source(S))$ (38). Supposing further that a high level of gene expression of the X protein is also observed, our objective in this experiment is to hypothesize about the various possible influences of X on the p53 pathway, thereby explaining how the cell can avoid cancer.

Our goal is now expressed as $\exists S(suppressed(\texttt{cancer},S) \wedge source(S))$. Like (45) in Section 3.4, the top clause is then given in SOLAR as:

$$ans(S) \leftarrow suppressed(\texttt{cancer},S) \wedge source(S). \tag{49}$$

The background theory B is now defined as the set consisting of the above causal network (47,48), the top clause (49), the meta-axioms (23–28), the integrity constraint (29),

[16] Each source is depicted with the symbol “↓” in the figure of a causal network.

the axioms for joint triggers (30,32,33), constraints for defaults (38,39), two facts for defaults: $no\_inhibitor(\mathtt{p53})$ and $no\_inhibitor(\mathtt{x})$, and domain constraints for pruning such as $\neg inhibited(\mathtt{uv}, Z)$ and $\neg inhibited(\mathtt{mdm2}, Z)$. Let the abducibles be

$$\Gamma_{\mathtt{x}} = \{triggered(\_, \_), inhibited(\_, \_), jointly\_triggered(\_, \_, \mathtt{x})\},$$

expecting that a mutant of X bound to some Z is produced in suppressing the cancer from some source. Then the production field $\mathcal{P}$ is set as:

$$\mathcal{P} = \langle \overline{\Gamma_{\mathtt{x}}} \cup \{ans(\_), \neg no\_inhibitor(\_)\}, Cond \rangle,$$

where $Cond$ is the length conditions such that each produced clause C must satisfy $|C \cap \{\neg triggered(\_, \_)\}| \leq 1$, $|C \cap \{\neg inhibited(\_, \_)\}| \leq 1$ and $|C \cap \{\neg jointly\_triggered(\_, \_, \mathtt{x})\}| \leq 1$. Then, SOLAR produces the 26 new characteristic clauses in 10 seconds using a PC with Core 2 Duo 3.06GHz and 4GB RAM.

In these 26 consequences of SOLAR, the following two clauses are included:

$$ans(\mathtt{uv}) \leftarrow triggered(\mathtt{x}, \mathtt{uv}) \wedge jointly\_triggered(Y, \mathtt{p53}, \mathtt{x}) \wedge inhibited(b, Y) \wedge no\_inhibitor(Y) \quad (50)$$

$$ans(\mathtt{uv}) \vee ans(\mathtt{mdm2}) \leftarrow triggered(\mathtt{x}, \mathtt{uv}) \wedge jointly\_triggered(Y, \mathtt{mdm2}, \mathtt{x}) \wedge inhibited(b, Y) \wedge no\_inhibitor(Y) \quad (51)$$

Both (50) and (51) give conditional answers. Consequence (50) represents a ***definite answer*** indicating that the p53-X complex has the unique source UV since both p53 and X are caused by the same source UV. On the other hand, (51) represents a ***disjunctive answer***: X is activated by UV but Mdm2 itself is assumed to be a source, hence the Mdm2-X complex has two sources. In fact, it takes more time to find consequence (51) than to find (50) in SOLAR. Those two formulas respectively correspond to the following hypotheses:

(I) $triggered(\mathtt{x}, \mathtt{uv}) \wedge \exists Y (jointly\_triggered(Y, \mathtt{p53}, \mathtt{x}) \wedge inhibited(b, Y))$,
(II) $triggered(\mathtt{x}, \mathtt{uv}) \wedge \exists Y (jointly\_triggered(Y, \mathtt{mdm2}, \mathtt{x}) \wedge inhibited(b, Y))$.

The variable Y in (I) or (II) represents a new complex synthesized from X and either p53 or Mdm2, respectively. That is, meta-level abduction generates a new mutant Y ("C" in Fig. 1) by combining X and either p53 or Mdm2, which then inhibits the existing complex b. Those two hypotheses are actually suggested in [65]: (I) X directly influences p53 protein stability: UV causes stress then induces high expression of X, which then binds to p53, so p53 is stabilized and formation of Mdm2-p53 complex is prevented; (II) X is a negative regulator of Mdm2: UV causes stress then induces high expression of X, which then binds to Mdm2, which competes against inhibiting the Mdm2-p53 interaction (depicted in dashed lines in Fig. 1). In both cases, p53 (or "A") can be functional as a tumor suppressor. In the biological viewpoint, however, the hypothesis (I) seems preferred because p53 has more chances to be bound to other proteins.

We again stress that all abduced links in hypotheses (I) and (II) are newly generated by meta-level abduction and that the new node Y is automatically invented during inference. In contrast, Tran and Baral [65] consider an action theory for this p53 network, and prepare all possible ground candidate nodes and links as abducibles in advance. Hence, no new node is invented in the framework of [65].

Concerning other 24 explanations obtained by SOLAR, 14 solutions promote a by inhibiting b, and 10 inhibit cancer directly. Although direct inhibition of cancer in the latter

case is feasible, we consider that revival of p53 as a tumor suppressor by promoting a is more important, since it relies on *double inhibition* of a through inhibition of b, which is biologically more admissible. Then, in such 14 explanations, 6 solutions have the definite answer $ans(\mathtt{uv})$, 6 have the definite answer $ans(\mathtt{mdm2})$, and 2 have the disjunctive answer $ans(\mathtt{uv}) \vee ans(\mathtt{mdm2})$. In these solutions, there are simple ones that never produce complexes with X, e.g.,

$$ans(\mathtt{uv}) \leftarrow inhibited(b, \mathtt{x}) \wedge triggered(\mathtt{x}, \mathtt{uv}),$$

but they are not very useful, since *inhibition is usually achieved by a mutant of some protein combined with another protein* in biological systems. One of the disjunctive answers has the same abducibles as in (50) except that X is triggered by Mdm2 instead of UV:

$$ans(\mathtt{uv}) \vee ans(\mathtt{mdm2}) \leftarrow triggered(\mathtt{x}, \mathtt{mdm2}) \wedge jointly\_triggered(Y, \mathtt{p53}, \mathtt{x}) \\ \wedge inhibited(b, Y) \wedge no\_inhibitor(Y).$$

One of other interesting solutions is

$$ans(\mathtt{uv}) \leftarrow jointly\_triggered(b, Z, \mathtt{x}) \wedge inhibited(Z, \mathtt{p53}),$$

which represents that suppression of b is realized by a complex of X and a new protein Z that is inhibited by p53. In this explanation, the additional axiom (32) works well.

The p53 regulatory network includes a complex array of upstream regulators and downstream effectors. There are versatile functions of p53 in the suppression of tumors; deriving the cell to apoptosis and ensuring the normal cell growth as a guardian of life. The stabilization of p53 seems to be the result of inhibition by Mdm2 and Mdm4, while activators such as HIPK2 and DYRK2 increase the p53 response. This complex system of activation/inhibition shows that p53 is a bottleneck of many regulatory mechanisms. The results obtained in this section are important in the sense that the activation/inhibition mechanism of p53 is linked to some proteins that might not have been found out yet. Meta-level abduction is thus crucial for this discovery task, and inferred hypotheses can suggest to scientists necessary experiments with gene knockout mice as minimally as possible.

4.2 Recovering Links in CDK Networks

The next case study is completion in the switch network of *cyclin-dependent kinases* (CDKs) [62]. CDKs are kinase enzymes that modify other proteins by chemically adding phosphate groups to them (phosphorylation), and are involved in the regulation of the cell cycle. A CDK is activated by association with a cyclin, forming a cyclin-CDK complex (Fig. 2).

The cdk2/cyclin E complex inactivates the retinoblastoma (Rb) protein, which then releases cells into the synthesis phase. Cdk2/cyclin E is regulated by both the positive switch called CAK (cdk activating kinase) and the negative switch p21/WAF1. p21/WAF1 is activated by p53, but p53 can also inhibit cyclin H, which is a source of the positive regulator of cdk2/cyclin E. The negative regulation from p53 works as a defensive system in the cells: when DNA damage occurs, it triggers p53, which then turns on the negative regulation to stop DNA synthesis, so that the damage should be repaired before DNA replication to prevent passing damaged genetic materials onto the next generation.

For the CDK network represented in Fig. 2 with $source(\mathtt{dna\_damage})$, several experimental problems are designed by removing some links from Fig. 2. Then, meta-level abduction is applied to verify how those removed links can be recovered and whether other

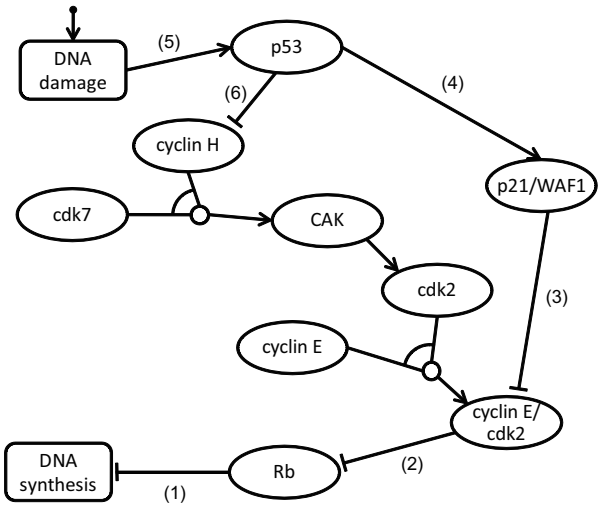


Fig. 2 Causal network of the CDK pathway

interesting links and unknown missing links can be inferred in explaining the observation

$$suppressed(\texttt{dna\_synthesis}, \texttt{dna\_damage}).$$

The objective of this experiment is to show how meta-level abduction can be well applied to complete incomplete networks. Recovery of removed links is a good testbed for this purpose because the existing natural system can be considered as an ideal solution. Yet, looking at other hypotheses, we understand that the same functions can be realized in different ways.

Table 2 shows a part of experimental results. All experiments are done in the environment of Mac mini with Core 2 Duo 1.83GHz and 2GB RAM. The maximum search depth of SOLAR is set to 5 for computing the characteristic clauses and 10 for computing the new characteristic clauses. The table shows 6 problems, each of which is given as a network obtained by removing the links shown in the table. In the table, the production field for each problem is defined as follows:

$$\mathcal{P} = \langle\, \{\neg triggered(\_,\_), \neg inhibited(\_,\_), \neg no\_inhibitor(\_)\},\ LCond \wedge Occ\text{-}Cond \,\rangle,$$

where $LCond$ is a condition on the maximum length len of each consequence (described in the table) and $Occ\text{-}Cond = |C \cap \{\neg no\_inhibitor(\_)\}| \leq 1$, that is, the number of occurrences of (the negation of) defaults in each consequence C is at most 1. The "#H" column shows the number of new characteristic clauses (minimal explanations) obtained by SOLAR. "Time" is computation time to obtain all minimal hypotheses and to check their consistencies.

The results of recovering removed links are shown in the table. The symbol "+" means the default $no\_inhibitor(\texttt{cdk2})$, and link (7) is $inhibited(\texttt{cyclin\_e}, \texttt{p53})$. Two nodes are declared to have no inhibitors in the axioms so that they do not need to be included in each explanation: $no\_inhibitor(\texttt{p53})$ and $no\_inhibitor(\texttt{p21\_WAF1})$. Here, among the number #H of explanations, we pick up three important groups of solutions depending on suppression paths from `dna_damage` to `dna_synthesis`. The first solution group goes through the right path via p21/WAF1, i.e., (5)-(4)-(3)-(2)-(1). The second one uses two cyclin-cdk joint triggers via inhibition (6) of cyclin H by p53. The third one generates the new link (7)

Table 2 Results of recovering links in the CDK pathway

Removed links	$LCond$	#H	Time [sec]	Minimal recovered links		
				p21-mediated	cdk2 activate	cyclin E inhibit
(1)	$len \leq 1$	4	10.7	(1)	—	—
(1), (2)	$len \leq 2$	9	18.8	(1g), (2g)	—	—
(1), (3)	$len \leq 2$	17	18.5	(1), (3)	(1), +	(1), (7)
(2), (4)	$len \leq 2$	24	20.3	(2), (4)	(2), +	(2), (7)
(4), (6)	$len \leq 2$	22	22.3	(4)	(6), +	(7)
(1), (2), (3) (4), (5)	$len \leq 5$	392	138.4	(1g), (2g), (3g) (4), (5)	(1g), (2g), (5), +	(1g), (2g), (5), (7)

which inhibits cyclin E by p53. For instance, when links $\{(1), (3)\}$ are removed, there exists the explanation in the first group that contains exactly the same links as the removed ones among the 17 hypotheses. For a link (N) for $\mathrm{N} = 1, 2, 3$, the recovered link (Ng) means that a more general hypothesis than (N) is obtained. For example, when $\{(1), (2)\}$, i.e., $inhibited(\texttt{dna\_synthesis}, \texttt{rb}) \wedge inhibited(\texttt{rb}, \texttt{cyclin\_e/cdk2})$ is removed, $\{(1\mathrm{g}), (2\mathrm{g})\}$, which is $\exists X(inhibited(\texttt{dna\_synthesis}, X) \wedge inhibited(X, \texttt{cyclin\_e/cdk2}))$, is recovered. Under axiomatization (23–28), explanations in the second group always contain the default $no\_inhibitor(\texttt{cdk2})$ (+). By this reason, when links $\{(1)\}$ or $\{(1), (2)\}$ are removed, no minimal explanation via the second path is obtained because any explanation is always subsumed by an explanation in the first group without the default. We can observe that two consecutive links are replaced by general ones with existentially quantified variables at recovery, e.g., $\{(1), (2)\}$, but links that are not connected are recovered as they are in the first group, e.g., $\{(1), (3)\}$. Notice that only either (4) or (6) is recovered at the first or second group, respectively, when links $\{(4), (6)\}$ are removed. Actually, the negative regulation by p53 has the two paths to suppress cdk2/cyclin E via (4) and (6), which correspond to the first and second paths, respectively. This shows that the biological network in Fig. 2 is *robust*, and the system can still work well even if one of the two paths is cut. In other words, if these two are cut simultaneously, only one recovery is logically sufficient to realize this function, and SOLAR outputs such minimally sufficient hypotheses in this case. Interestingly, we can also find other recoveries by connecting nodes that have not been originally connected. The third solution group realizes such a case by newly imposing inhibitor (7) (Fig. 3).

A well-known effect of p53 expression is a block in the cell division cycle, and induction of p53 leads to cell growth arrest and apoptosis. Although the function of p53 to induce a G1 arrest by binding with p21 has been well defined, the control of cyclin E by p53 has not been established yet. In our experiment, we have applied abductive inference on an incomplete p53 network related to its activation by DNA damage, and one of the results obtained in Fig. 3 clearly shows that p53 degrades the cyclin E protein by finding (7) and consequently imposes cycle arrest. This is a biologically important function verified by our method. A possible explanation here is based on *analogical reasoning* such that the inhibition of cyclin H from p53 in the same network can suggest a possible inhibition of cyclin E from p53. This analogy could be regarded as a form of inductive reasoning because it strives to provide understanding of what is likely to be true. We note that link (7) is automatically generated in our experiment without giving such an analogical bias.[17] In the literature, examples of this function can be found in genetic analysis of *Drosophila* melanogaster [34], which has shown that disruption of the mitochondrial electron transport chain activates a G1–S checkpoint as a result of control of cyclin E by p53, and in a recent evidence [35] that p53 modulates the

[17] We can also supply such a bias as $inhibited(X, \texttt{p53}) \wedge cyclin(X)$ in our framework by giving facts $cyclin(\texttt{cyclin\_h})$ and $cyclin(\texttt{cyclin\_e})$.

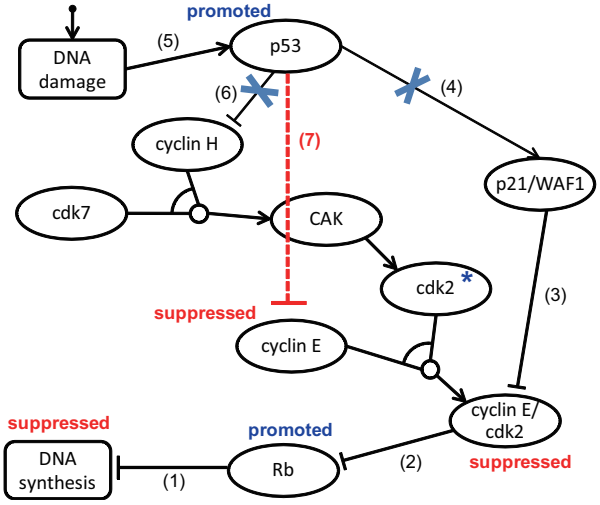


Fig. 3 p53-cyclin E-mediated pathway

activity of the ubiquitin-proteasome system to degrade cyclin E and thereby imposes cell cycle arrest. Our results hence explain those results from a logical point of view.

In this subsection, we have verified that previously known solutions, either well-known or just reported in the literature, can be generated by our abductive method. In other words, meta-level abduction has the ability to recover or reproduce such solutions *from an incomplete network*. The experiments show that such known solutions can be found among many other solutions, so recovery rate would be low if our task were to recover exactly the same removed links. From the biological viewpoint, however, this phenomenon is not surprising and can be explained as follows. The biological mechanism is so complex that any configuration of a biological network cannot be optimally organized *per se*. Once a biological function is lost in a network for some reason, it is not naturally recovered as it was. Actually, there are many possibilities to realize the same function, since the inverse problem does not have a unique solution in general. Moreover, biological systems are *robust* and have much *functional redundancy*, and survival of individuals and species preservation are of overriding importance for them [66]. In other words, a biological system is *resilient*, that is, it has the ability to endure and successfully recover from perturbations such as genetic and environmental changes. Having multiple stable states (solutions) in a biological system, transition to a new stable state occurs in the face of changes so that they can survive and evolve. Our results having many alternative solutions can thus be understood as a logical account for an example of biological resilience. Now, *the purpose of abduction here is not recovery of particular links but recovery of the function itself*, i.e., suppression of DNA synthesis. Hence, it is essential for the abductive system to guarantee that important solutions are never lost. Under this completeness condition, it is better to restrict the number of possible solutions as many as possible, and we will discuss this important issue in Section 7. Finally, biologists often decompose networks into smaller sub-networks to make biological analysis easier, and thus even those simplified p53 networks in this section are useful for considering all possible connectivities that are hard to find for human experts.

5 Experiments of Network Completion

This section addresses the issue of scalability of meta-level abduction, which is important when applications involve large networks. To this end, we show experiments of network completion by meta-level abduction on networks randomly generated with some parameters varied, and see scalability of the method.

For experiments, networks are generated randomly, each of which consists of the specified numbers of nodes and edges including both positive and negative links. We consider 70 combinations of the numbers of nodes and edges, whose ratios are among 3:1, 2:1, 1:1, 1:2, 1:3, 1:4 and 1:5 (7 ratios), each varies from 10 nodes to 100 nodes (10 cases). Note that each ratio determines the average *degree* (the number of adjacent nodes connected to a node), but it does not mean that the degree is always fixed to every node. The consistency of each network instance together with axioms (23–28) and constraints (29, 39) is checked when the network is generated; if a randomly generated network is proved inconsistent by SOLAR, that is, the empty clause is produced as the unique characteristic clause, the next network is randomly generated until a consistent one is found. By this way, 30 consistent random network instances are generated for each test case. We set the timeout for each run of SOLAR on one network instance to 1,800 seconds. Then, problems with ratio 1:5 and 100 nodes (500 edges) cannot be solved within the time limit. We also do not allow bidirectional links between any pair of nodes like $triggered(a, b) \wedge inhibited(b, a)$, so that there is at most one link, which is either a trigger or an inhibitor, is allowed between two nodes. Then, there are ${}_{10}C_2 = 45$ edges for any complete graph with 10 nodes, and we cannot generate any network for problems with ratio 1:5 and 10 nodes (50 edges). Hence, we test $7 \times 10 - 2 = 68$ cases, and $68 \times 30 = 2,040$ network instances are generated.

The observation for each network instance is given as $suppressed(n - 1, 0)$ where n is the number of nodes in the network, supposing that 0 is the source node and $n - 1$ is the goal node. We limit the maximum length of hypotheses to 3, which is considered reasonable in many real applications. The maximum search depth of SOLAR in running each network instance is set to 5 for computing the characteristic clauses and 8 for computing the new characteristic clauses. Running environment is Core 2 Duo 1.83GHz with 2GB RAM. Fig. 4 is the graph of results with all 7 nodes-edges ratios on the average computation time to compute all explanations for each test case with 30 network instances. Each time includes computation of generating all possible hypotheses and their consistency checking. Fig. 5 is the graph of results on the average number of explanations generated by network completion for each test case. The ratio of m nodes to n edges is denoted as "$m : n$" in these graphs.

From Figs. 4 and 5, it is observed that, as the number of nodes (and edges) increases, both the CPU time and the number of hypotheses grow in all cases of ratios. As the degree of a network increases, completion of intermediate nodes in the network occurs more frequently and then time and hypotheses also increase. In particular, when new nodes are created, possible new connections between the source and the goal become larger. However, as the ratio of nodes to edges becomes smaller, e.g., 1:4 and 1:5, that is, *as the average degree of a network increases, the growth rates of hypotheses decrease and tend to converge*. This is a remarkable result in this experiment, and this dynamics is explained as follows. Under the limit of inference depth and maximum length of consequences, there is no room for further addition of new edges into such a "dense" network by keeping the consistency, so the number of hypotheses does not increase too much.

It has been recognized that the complexity of *enumeration problems* may not be defined in terms of functions of the size of the input because the number of solutions can be exponential in the input size [64]. For example, [24] defines that an enumeration problem can be

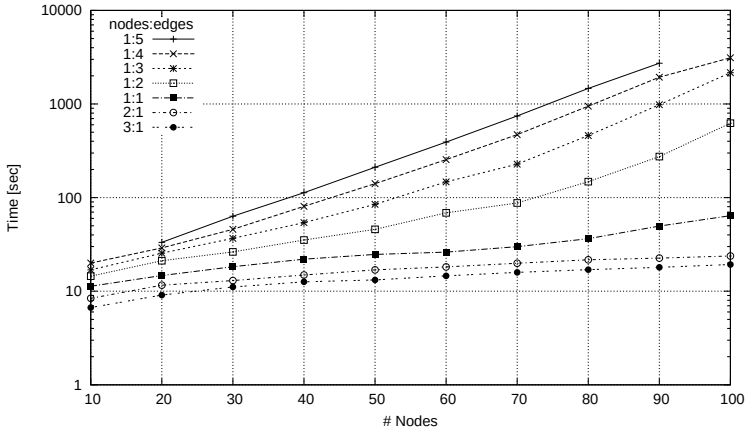


Fig. 4 Total computation time of network completion

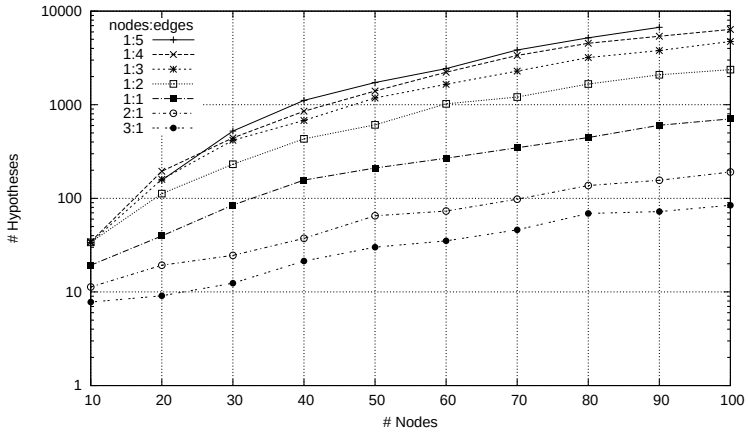


Fig. 5 The number of hypotheses generated by network completion

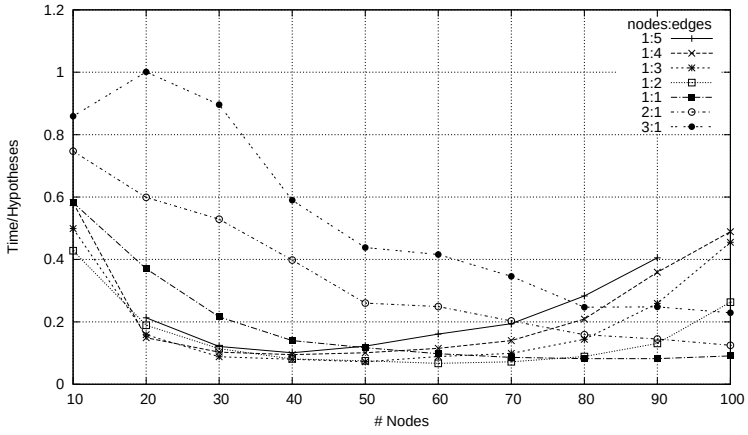


Fig. 6 Average time required to find one hypothesis in network completion

solved in polynomial total time if an algorithm exists with running time bounded by a polynomial function of the combined size of the input and the output. With this regard, Fig. 6 shows the graph of average time to generate one hypothesis, which is total running time divided by the number of generated hypotheses. The figure indicates that time to compute each hypothesis in enumeration does not grow exponentially with increase of input network sizes for all cases of networks. In Fig. 6, when there are fewer nodes, e.g., 10–30, the running time is short and then there are some overheads of file read and start-up of Java Virtual Machine. On the other hand, when many nodes exist (90–100), time for subsumption checking (consistency checking) increases due to a large number of hypotheses and then average time to compute one hypothesis also increases, but this happens only when the ratio of nodes to edges becomes smaller (1:3–1:5). Hence, it is possible to argue that enumeration scalability can be observed in this experiment of network completion.

Finally, it should be noted that no domain-dependent knowledge other than an initial network is incorporated in each run of this experiment. In real-world applications, however, we expect that the number of hypotheses can be decreased in general by incorporating more background knowledge and constraints and by specifying more restricted production fields.

6 Discussion and Related Work

6.1 Rule Abduction

The method of *rule abduction* by means of meta-level abduction was firstly introduced in [18], and this paper has extended it to deal with positive and negative causal links. Although few works on rule abduction exist previously, they focus on positive effects only unlike our work. Moreover, the patterns of *abducible rules* must be determined in advance. The framework of abductive systems with such abducible rules were firstly considered in [49], which associates a unique name with each ground instance of a rule schema that is predetermined as possible hypotheses, and those names are treated as abducible atoms in abductive reasoning. This is a convenient method when we know exact patterns of rules as strong biases. However, it is impractical or impossible to prepare all patterns of rules in advance in order to abduce missing rules, and predicate invention is basically impossible by this method.

Meta-reasoning has been discussed intensively in logic programming [29, 12, 4], but rule abduction has never been considered in the literature of meta-reasoning. While the importance of both abduction and meta-reasoning is discussed in [29], their combination has not been discussed explicitly. Actually it is possible to embed abductive procedures into meta-programming to perform abduction, but this is distinguished from abduction at the meta level. Christiansen [2] computes both abduction and induction in a unified system of meta-programming, and uses separate forms of reasoning in their actual computation followed by the system. In [2], the $demo$ predicate is used to generate parts of programs that are necessary to derive the goal based on techniques of constraint logic programming.

As a predecessor of this work, applications of SOLAR to complete networks are discussed in [53], but negative causes and effects are not handled and joint (positive) causes are not considered there. CF-induction [14] can induce explanatory rules, and its applications to completion of causal rules and abduction of network status in metabolic pathways are shown in [68]. CF-induction can directly induce first-order full clausal theories at the object level, but predicate invention and hypothesis enumeration at the object level are not easier than those by meta-level abduction. This is because predicate invention is realized by *inverse resolution* [41] and the search space for hypothesis enumeration is huge in general.

6.2 Combining Abduction and Induction

Abduction is also used to induce rules in several ILP systems other than CF-induction [14]. Earlier works are summarized in several papers in [10], see, e.g., [37,58]. Most such previous systems use abduction to compute inductive hypotheses efficiently, and such integration is useful in theory refinement. More recently, those works [67,52,27,3] use procedures for *abductive logic programming* (ALP), which are different from SOLAR that is used for full clausal abductive theories as in our meta-level abduction. Progol [39] is the first ILP system based on inverse entailment, and [40] improves it to realize theory completion. ILP systems in [67,52,27] extend Progol's applicability to larger classes of induced logic programs, which are yet limited to a set of Horn clauses. Such extensions do not necessarily depend on ALP, and [54] uses SOLAR instead of ALP to deal with non-Horn clauses. These systems abduce part of hypotheses from the background theory together with the negation of an observation. In [3], an ILP problem is fully translated into an equivalent ALP problem, and then an ALP procedure is used to abduce a hypothesis for an observation from the background theory and an inductive bias called a *top theory*. Their system TAL is an extension of TopLog [43] to allow for multiple clause hypotheses with negation-as-failure (or default negation). To abduce rules in ALP, TAL uses the naming method as in [49], thus depends on a strong bias to induce rules, while our meta-level abduction defines the abducibles as meta-predicates only.

There are other types of works on combining ALP and ILP. Dimopoulos and Kakas [6] propose a general framework in which a new abductive logic program is learned, given a previous abductive program as background knowledge and training examples. Their framework is followed by [25,31]. On the other hand, Inoue and Haneda learn an abductive program from a non-abductive logic program with positive and negative examples, and new abducibles are acquired there [19]. The covering relation in [6,25,31] is defined based on *abductive entailment*, while [19] did not change the entailment relation from the ordinary one in extended logic programs.

6.3 Abduction and Induction in Causal Theories

Some attempts in ALP, ILP and AI have contributed to abduction and induction in causal theories. In these works, a causal theory is formally represented in a more specialized way as a set of individual rules between causes and their effects. The *event calculus* [30] is a meta-theory for reasoning about time and action in the framework of logic programming. *Abductive event calculus* is an abductive extension of event calculus [8], and can thus be regarded as a kind of meta-level abduction. Abductive event calculus has been extended for applications to planning, e.g., [61], but has never been used for abducing causal theories.

Moyle [38] uses a theory completion technique of [40,67] to learn a causal theory in the form of logic programs based on event calculus, given positive examples of input-output relations. In this work, a complete initial state is required as an input and a complete set of *narrative* facts is computed in advance, and thus observations handled in our work cannot be explained. Otero [47] considers causal theories represented in logic programs in the case of incomplete narratives. These previous works need either *frame axioms* or *inertia rules* in logic programs. The former causes the frame problem and the latter requires induction in nonmonotonic logic programs. Inoue *et al.* [16] induce causal theories represented in an *action language* given an incomplete action description and observations, but requires an algorithm to learn finite automata to compute hypotheses, which may search the space of

possible permutations of actions. Unlike our meta-level abduction, all these previous works on induction of action theories do not consider invention of new events or objects.

Tran and Baral [65] use an action language which formalizes notions such as *causal rules* of the form "X causes Y", *trigger rules* of the form "X triggers Y", and *inhibition rules* of the form "X inhibits Y" to model cell biochemistry, and apply it to hypothesize about signaling networks presented in Section 4.1. As discussed in Section 4.1, all ground candidate nodes and links to be added are prepared as abducible causal/trigger/inhibition rules in advance in [65]. In contrast, all abduced causal relations as well as new nodes are automatically generated in our meta-level abduction.

6.4 Completion of Biological Networks

There are several works on completing biological networks. Notably, work on Robot Scientist [28] adopts abduction to complete biochemical pathways. The abduction mechanism of Robot Scientist detailed in [56] suggests the use of SOLDR resolution [67], which is a version of SOL resolution [13] restricted for Horn theories. In [56], a *reaction* is defined as a pair of sets of compounds representing the substrates and the products of the reaction, and a *metabolic graph* is defined in such a way that each node is given as a set of compounds available by sequences of reactions. This representation is used to deal with biochemical networks represented as hyper-graphs and to jointly propagate causes of the parent nodes to those of their child nodes. However, since each node does not correspond to one compound but represents a set of compounds, the number of nodes in a metabolic graph becomes large. In contrast, we adopt causal networks, which are simpler than metabolic graphs, and the meta-predicate $jointly\_connected$ or a disjunction of $connected$ literals is used to represent a direct multi-causal relationship by (7) or (4). Moreover, our causal network deals with inhibition, while [56] does not take negative effects into account.

Metabolic pathways are updated by several reasons. As discussed in Section 1, known metabolic pathways are often incomplete so that they should be completed [28]. New experimental data often reveal that previously known pathways are incorrect so that they should be revised as opposed to merely extended [55]. Moreover, biological systems are robust to perturbations and can survive in critical situations like loss of some important genes or links by finding bypasses in pathways [22] and by creating novel reactions that are not normally used [46]. In computational models, metabolic pathways are completed in [59] and are revised in [55] using *answer set programming* (ASP). These works do not invent new nodes, and the work [59] does not consider inhibition. Furthermore, in revising pathways in [55], deletion is realized through addition of atom δ_R representing retractability of rule R, where negation-as-failure $not\ \delta_R$ appears in the body of each retractable rule R represented in ASP. This kind of deletion is only conceivable under the assumption that we can declare retractable links that are subject to change in advance. However, real deletion of connections between metabolites is generally impossible *in vivo* once they have been established. Instead of deleting links from a network, our method *adds* new links and nodes by abduction, yet the switch between presence (activation) and absence (inhibition) of a node can be controlled by our method due to the nonmonotonic axiomatization (23–28). Our abductive method is then more realistic, since we can simulate the biological effect of double inhibition by inhibiting inhibitors in the p53 network in Section 4.1. Abduction in metabolic pathways with inhibition is also considered in [63], although the problem setting in [63] is to infer the state of each reaction, which is different from network completion and no new links/nodes are abduced there. Moreover, this kind of qualitative abstraction is a very difficult task, and in-

deed the axiomatization in [63] can be inconsistent with some networks and observed data, as is discussed in [51].

For *gene regulatory networks*, Gat-Viks and Shamir [11] determine a class of regulation functions, by which regulators determine transcription, and analyze their complexity. Zupan *et al.* [69] construct networks from mutant data using abduction, but use experts' heuristic rules for construction. These works are different from ours in that they use more specific methods at the object level to reconstruct networks, depending on the problem domains. On the other hand, completion of *signaling networks* is analyzed in a general way by Akutsu *et al.* [1], in which unknown Boolean functions are guessed in a *Boolean network* [26] whose network topology is known. This contrasts with our setting that a network is incomplete and its topology is not fixed.

6.5 Extension of Network Representation

Besides completion of biological networks, logical inference on networks has been considered with several different formalizations in the literature. Leitgeb [32] defines *inhibition nets* as a special case of artificial neural networks to represent inhibitory connections between nodes and excitatory connections. Inhibition nets can represent some nonmonotonic logics and can simulate logic programs. In other words, inhibitors can be used to realize nonmonotonic reasoning, and can replace negation-as-failure in logic programs to some extent. On the other hand, Fayruzov *et al.* [9] formalize inference on gene regulatory networks using ASP by utilizing nonmonotonic behavior with negation-as-failure. In [9], each particular inference problem is translated into an equivalent computational problem in ASP at the meta level. Inoue [15] shows that logic programs with negation-as-failure can simulate Boolean networks [26] and vice versa. These works support a claim that inhibition and negation-as-failure can simulate each other, and that inference on networks with inhibitors is nonmonotonic.

Although we have shown two axiomatizations in Section 3, other formalizations can be still considered for controlling inference in different ways. In fact, we have noticed that there are at least two arguable cases using the axiomatization in this paper. The first one is a competing case that there are multiple triggers and one or fewer inhibitors for a node g. Then, instead of preferring the unique inhibitor to all triggers as in Section 3.2, we could use the *majority function* or assign *probabilities* to them to determine the value of g. Setting a *threshold* on a node is also useful, so that a node is activated if the sum of input signals for the node exceeds the threshold. Such a mechanism is implemented using ASP in [9]. Alternatively, guess of a Boolean function for updating the truth value of g in the framework of Boolean networks is another interesting direction in this case, as shown in [1]. Network N_4 (40), which is inconsistent with (23–28) and (29), can also be handled in such a way. The use of *paraconsistent logic* [5] would also help to infer some meaningful information from part of knowledge that are not related to inconsistency.

The second important case is a *negative feedback loop*, in which a node g depends negatively on itself. The smallest example of negative feedback loops is $N_5 = \{inhibited(g, g)\}$ (41), which is inconsistent by Proposition 4. Interestingly, if we interpret N_5 as a Boolean network, then any stable state (called a *point attractor*) of N_5 is characterized by a model of the Boolean equation $(\neg g \equiv g)$ [15]. Since this is unsatisfiable, N_5 has no stable state. Actually, a negative loop is the source of periodic oscillation (called a *cycle attractor*) in Boolean networks. Hence, it is admissible to return no explanation for an observation in a network containing negative feedback loops, as long as we are concerned with solutions in a static or

equilibrium state. On the other hand, to represent *dynamic behaviors* in our framework, we need to introduce *time* for causal axioms. For example, axiom (25) can be reexpressed as

$$promoted(X, Y, T + 1) \leftarrow inhibited(X, Z) \wedge suppressed(Z, Y, T).$$

by incorporating the time argument. Lejay *et al.* [33] have examined such an extended axiomatization in an application of meta-level abduction to treatment of hypertension, but a formal work is needed to explore this extension.

7 Conclusion

The method of meta-level abduction [18] has been extensively explored in this paper. In particular, we have allowed representation of positive and negative effects in causal networks, and deductive and abductive methods have been investigated for reasoning about networks. Two axiomatizations for causal networks have been presented: one treats positive and negative effects equally and the other prefers inhibitors to triggers. The latter involves default reasoning on triggers. With this extension, nonmonotonic reasoning in causal networks and completion of positive and negative links are now possible by meta-level abduction.

Besides reasoning with positive and negative causal effects discussed in this paper, meta-level abduction has several advantages in abduction. For example, multiple observations are explained at once, full clausal theories can be allowed for background knowledge, both rules and facts can be abduced at the object level, and predicate invention is partially realized as existentially quantified hypotheses.

Problem solving with meta-level abduction consists of three steps: (I) design of meta-level axioms; (II) representation of domain knowledge at the meta level; and (III) restriction of the search space to treat large knowledge. This work supposes an incomplete network for Step (II), and hence representation of a problem is rather tractable. On the other hand, we have made great effort on Step (I), by gradually extending the positive causal networks (Section 2) to alternating axioms (Section 3.1) and then to nonmonotonic formalization (Section 3.2). This formalization process was based on trial and error by running SOLAR many times, and case studies in Section 4 were used as testbeds for this purpose. Step (III) is important not only for more efficient computation of hypotheses but for making the user of meta-level abduction to select appropriate hypotheses more easily. This goal is achieved by introducing more constraints, and it is important to explore more useful methods for inducing such constraints. In [33], a general method to generate constraints is proposed to prune many useless solutions by assuming that given link information does not change so that a trigger cannot become an inhibitor and vice versa. For example, if $triggered(g, t)$ is true in the given network, then $\neg inhibited(g, t)$ is assumed automatically. If we further assume non-existence of negative feedbacks, the constraint $\neg inhibited(t, g)$ is further added there. These two additional constraints are, however, still weak because they can be also obtained as characteristic clauses for network N_1 in Section 3.1. Then, a stronger constraint $\neg triggered(t, g)$ is introduced by assuming that the direction of a link is not changed, but is often too restrictive and cannot be applied in a complicated case. Domain-dependent constraints are more powerful, and we have seen in Section 4.1 which hypotheses are more important than others, and have shown that analogical biases are useful in Section 4.2. Incorporation of these guides would be a promising way to focus on particular solutions. Thus, discovering a more general and useful method for constraint generation is left as a future work.

One very important topic that we could not describe much in this paper is evaluation of logically possible hypotheses obtained in meta-level abduction. To evaluate hypotheses,

some statistical methods have been proposed. For example, hypotheses can be ranked according to frequencies of literals appearing in them and corresponding paths [21], and can be given their scores according to their fitness with observed data [11]. In general, obtained hypotheses can be used to explain more than given observations. For example, cross-validation can be used to evaluate hypotheses as long as observed data can be used for both training and testing, although case studies in Section 4 are not of this kind. Often, within a certain maximum number of alternative solutions, providing a set of *diverse solutions* rather than a few most promising solutions is rather a good idea. To this end, we need more domain-dependent or general constraints and heuristics to restrict the search space and the number of possible solutions, so the future work mentioned above is also important here.

This paper has shown applications of meta-level abduction to signaling pathways, but we expect that the proposed method can also be applied to abduction in metabolic pathways with inhibition [63, 68, 55] as well as transcription networks [11]. Another interesting application is to understanding and modeling of robustness of biological systems during severe and unexpected environments [66, 22, 46]. We hope that meta-level abduction could contribute to a future breakthrough in sciences including biology and medicine by helping scientists to discover important missing and unknown networks.

Acknowledgements The authors would like to thank many people who had discussions on a wide range of topics related to this work: Tomoya Baba and Kenji Nakahigashi for their comments from viewpoints of biologists, Domenico Corapi, Koichi Furukawa, Bob Kowalski and Stephen Muggleton for discussions on integration of meta-reasoning and abduction, Jean-Charles Faye and Olivier Sordet from L'Institut Claudius Regaud for their verification of our results on completed p53 pathways, Koji Iwanuma and Oliver Ray for their contribution on conditional query answering and disjunctive answers, Hiroshi Maruyama and colleagues of the Systems Resilience project for discussions on robustness, Philip Reiser for his early attempt to complete biological networks by SOLAR, and the reviewers of this paper for their many comments and suggestions.

References

1. Akutsu, T., Tamura, T. and Horimoto, K., Completing networks using observed data, *Proceedings of ALT '09*, LNAI 5809, pp.126–140. Springer (2009)
2. Christiansen, H., Abduction and induction combined in a metalogic framework, in: [10], pp.195–211 (2000)
3. Corapi, D., Russo, A. and Lupu, E.C., Inductive logic programming as abductive search, *Technical Communications of ICLP '10*, pp.54–63 (2010)
4. Costantini, S., Meta-reasoning: A survey, in: Kakas, A.C. and Sadri, F. (eds.), *Computational Logic: Logic Programming and Beyond, Essays in Honour of Robert A. Kowalski, Part II*, LNAI 2408, pp.253–288. Springer (2002)
5. Damásio, C.V. and Pereira, L.M., A survey of paraconsistent semantics for logic programs, in: Gabbay, D.M. and Smets, P. (eds.), *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, Vol.2, pp.241–320. Springer (1998)
6. Dimopoulos, Y. and Kakas, A., Abduction and inductive learning, in: De Raedt, L. (ed.), *Advances in Inductive Logic Programming*, pp.144–171. IOS Press (1996)
7. Dietterich, T., Domingos, P., Getoor, L., Muggleton, S. and Tadepalli, P., Structured machine learning: The next ten years, *Machine Learning*, 73, 3–23 (2008)
8. Eshghi, K., Abductive planning with event calculus, in: *Proceedings of the 5th International Conference on Logic Programming*, pp.562–579 (1988).
9. Fayruzov, T., Janssen, J., Cornelis, C., Vermeir, D. and De Cock, M., Extending Boolean regulatory network models with answer set programming, *Proceedings of 2010 IEEE International Conference on Bioinformatics and Biomedicine*, pp.207–212 (2010)
10. Flach, P.A. and Kakas A.C. (eds.), *Abduction and Induction—Essays on their Relation and Integration*, Kluwer Academic (2000)
11. Gat-Viks, I. and Shamir, R., Chain functions and scoring functions in genetic networks, *Bioinformatics*, 19(Suppl.1), i108–i117 (2003)

12. Hill, P.M. and Gallagher, J., Meta-programming in logic programming, in: Gabbay, D.M., Hogger, C.J. and Robinson, J.A. (eds.), *Handbook of Logic in Artificial Intelligence and Logic Programming*, Vol.5, pp.421–497. Oxford University Press (1998)
13. Inoue, K., Linear resolution for consequence finding, *Artificial Intelligence*, 56, 301–353 (1992)
14. Inoue, K., Induction as consequence finding, *Machine Learning*, 55, 109–135 (2004)
15. Inoue, K., Logic programming for Boolean networks, in: *Proceedings of IJCAI-11*, pp.924–930 (2011)
16. Inoue, K., Bando, H. and Nabeshima, H., Inducing causal laws by regular inference, in: *Proceedings of ILP '05*, LNAI 3625, pp.154–171. Springer (2005)
17. Inoue, K., Doncescu, A. and Nabeshima, H., Hypothesizing about networks with positive and negative effects by meta-level abduction, in: Frasconi, P. and Lisi, F.A. (eds.), *Inductive Logic Programming: Revised Papers from the 20th International Conference* (ILP '10), LNAI 6489, pp.114–129. Springer (2011)
18. Inoue, K., Furukawa, K., Kobayashi, I. and Nabeshima, H., Discovering rules by meta-level abduction, in: De Raedt, L. (ed.), *Inductive Logic Programming: Revised Papers from the 19th International Conference (ILP '09)*, LNAI 5989, pp.49–64. Springer (2010)
19. Inoue, K. and Haneda, H., Learning abductive and nonmonotonic logic programs, in: [10], pp.213–231 (2000)
20. Inoue, K., Iwanuma, K. and Nabeshima, H., Consequence finding and computing answers with defaults, *Journal of Intelligent Information Systems*, 26, 41–58 (2006)
21. Inoue, K., Sato, T., Ishihata, M., Kameya, Y. and Nabeshima, H., Evaluating abductive hypotheses using an EM algorithm on BDDs, in: *Proceedings of IJCAI-09*, pp. 810–815 (2009)
22. Ishii, N., Nakahigashi, K., Baba, T. *et al.*, Multiple high-throughput analyses monitor the response of *E. coli* to perturbations, *Science*, 316, 593–597 (2007)
23. Iwanuma, K. and Inoue, K., Minimal answer computation and SOL, in: *Proceedings of JELIA '02*, LNAI 2424, pp. 245–257. Springer (2002)
24. Johnson, D., Yannakakis, M. and Papadimitriou, C., On generating all maximal independent sets, *Information Processing Letters*, 27(3), 119–123 (1988)
25. Kakas, A.C. and Riguzzi, F., Learning with abduction, in: *Proceedings of the 7th International Workshop on Inductive Logic Programming*, LNAI 1297, pp.181–188. Springer (1997)
26. Kauffman, S.A., *The Origins of Order: Self-Organization and Selection in Evolution*, Oxford University Press (1993)
27. Kimber, T., Broda, K. and Russo, A., Induction on failure: Learning connected Horn theories, in: *Proceedings of LPNMR '09*, LNAI 5753, pp.169–181. Springer (2009)
28. King, R.D., Whelan, K.E., Jones, F.M., Reiser, P.G.K., Bryant, C.H., Muggleton, S.H., Kell, D.B. and Oliver, S.G., Functional genomic hypothesis generation and experimentation by a robot scientist, *Nature*, 427, 247–252 (2004)
29. Kowalski, R.A., Problems and promises of computational logic, in: Lloyd, J.W. (ed.), *Computational Logic*, pp.1–36 (1990)
30. Kowalski, R.A. and Sergot, M.J., A logic-based calculus of events, *New Generation Computing*, 4(1):67–95 (1986)
31. Lamma, E., Mello, P., Milano, M., Riguzzi, F., Esposito, F., Ferilli, S. and Semeraro, G., Cooperation of abduction and induction in logic programming, in: [10], pp.233–252 (2000)
32. Leitgeb, H., Nonmonotonic reasoning by inhibition nets, *Artificial Intelligence*, 128:161–201 (2001)
33. Lejay, G., Inoue, K. and Doncescu, A., Application of meta-level abduction for the treatment of hypertension using SOLAR, *Proceedings of the AINA-2011 Workshop on Bioinformatics and Life Science Modeling and Computing*, pp.495–500. IEEE Computer Society (2011)
34. Mandal, S., Guptan, P., Owusu-Ansah, E. and Banerjee, U., Mitochondrial regulation of cell cycle progression during development as revealed by the tenured mutation in *Drosophila*, *Developmental Cell*, 9, 843–54 (2005)
35. Mandal, S., Freije, W.A., Guptan, P., and Banerjee, U., Metabolic control of G1–S transition: Cyclin E degradation by p53-induced activation of the ubiquitin-proteasome system, *J. Cell Biol.*, 188(4), 473–479 (2010)
36. Meek, D.W., Tumour suppression by p53: a role for the DNA damage response? *Nature Reviews Cancer*, 9, 714–723 (2009)
37. Mooney, R., Integrating abduction and induction in machine learning, in: [10], pp.181–191 (2000)
38. Moyle, S., Using theory completion to learn a robot navigation control programs, in: *Proceedings of ILP '02*, LNAI 2583, pp.182–197. Springer (2003)
39. Muggleton, S., Inverse entailment and Progol, *New Generation Computing*, 13(3/4), 245–286 (1995)
40. Muggleton, S. and Bryant, C., Theory completion and inverse entailment, in: *Proceedings of ILP 2000*, LNAI 1866, pp.130–146. Springer (2000)
41. Muggleton, S. and Buntine, W., Machine invention of first-order predicate by inverting resolution, in: *Proceedings of the 5th International Workshop on Machine Learning*, pp.339–351. Morgan Kaufmann (1988)

42. Muggleton, S.H., De Raedt, L., Poole, D., Bratko, I., Flach, P., Inoue, K. and Srinivasan, A., ILP turns 20—Biography and future challenges. *Machine Learning*, 86(1):3–23 (2012)
43. Muggleton, S., Santos, J.C.A. and Tamaddoni-Nezhad, A., Toplog: ILP using a logic program declarative bias, in: *Proceedings of ICLP '08*, LCNS 5366, pp.687–692. Springer (2008)
44. Nabeshima, H., Iwanuma, K. and Inoue, K., SOLAR: A consequence finding system for advanced reasoning, in: *Proceedings of TABLEAUX '03*, LNAI 2796, pp.257–263. Springer (2003)
45. Nabeshima, H., Iwanuma, K., Inoue, K. and Ray, O., SOLAR: An automated deduction system for consequence finding, *AI Communications*, 23(2–3), 183–203 (2010)
46. Nakahigashi, K., Toya, Y., Ishii, N., Soga, T., Hasegawa, M., Watanabe, H., Takai, Y., Honma, M., Mori, H. and Tomita, M., Systematic phenome analysis of *Escherichia coli* multiple-knockout mutants reveals hidden reactions in central carbon metabolism, *Molecular Systems Biology*, 5(306) (2009)
47. Otero, R.P., Induction of the indirect effects of actions by monotonic methods, in: *Proceedings of ILP '05*, LNAI 3625, pp.279–294. Springer (2005)
48. Pearl, J., *Causality: Models, Reasoning, and Inference*, 2nd Ed., Cambridge (2009)
49. Poole, D., A logical framework for default reasoning, *Artificial Intelligence*, 36, 27–47 (1988)
50. Prives, C and Hall, P.A., The p53 pathway, *Journal of Pathology*, 187, 112–126 (1999)
51. Ray, O., Towards a rational approach for the logical modelling of inhibition in metabolic networks, in: *Proceedings of the 2009 IEEE International Workshop on Bioinformatics and Life Science Modeling and Computing*, pp.18–923 (2009)
52. Ray, O., Broda, K. and Russo, A., Hybrid abductive inductive learning: A generalisation of Progol, in: *Proceedings of ILP '03*, LNAI 2835, pp.311–328. Springer (2003)
53. Ray, O. and Inoue, K., A consequence finding approach for full clausal abduction, in: *Proceedings of the 10th International Conference on Discovery Science*, LNAI 4755, pp.173–184. Springer (2007)
54. Ray, O. and Inoue, K., Mode-directed inverse entailment for full clausal theories, in: *Proceedings of ILP '07*, LNAI 4894, pp.225–238. Springer (2008)
55. Ray, O., Whelan, K. and King, R., Logic-based steady-state analysis and revision of metabolic networks with inhibition, in: *Proceedings of the 3rd International Workshop on Intelligent Informatics in Biology and Medicine*, pp.661–666 (2010)
56. Reiser, P.G.K., King, R.D., Kell, D.B., Muggleton, S.H., Bryant, C.H. and Oliver, S.G., Developing a logical model of yeast metabolism, *Electronic Transactions in Artificial Intelligence*, 5-B2(024), 223–244 (2001)
57. Reiter, R., A logic for default reasoning, *Artificial Intelligence*, 13, 81–132 (1980)
58. Sakama, C., Abductive generalization and specialization, in: [10], pp.253–265 (2000)
59. Schaub, T. and Thiele, S., Metabolic network expansion with answer set programming, in: *Proceedings of ICLP '09*, LNAI 5649, pp.312–326. Springer (2009)
60. Schurz, G., Patterns of abduction, *Synthese*, 164(2), 201–234 (2008)
61. Shanahan, M., An abductive event calculus planner, *Journal of Logic Programming*, 44:207–239 (2000)
62. Shmulevich, I., Dougherty, E.R., Kim, S. and Zhang, W., Probabilistic Boolean networks: A rule-based uncertainty model for gene regulatory networks, *Bioinformatics*, 18(2), 261–274 (2002)
63. Tamaddoni-Nezhad, A., Chaleil, R., Kakas, A. and Muggleton, S. Application of abductive ILP to learning metabolic network inhibition from temporal data, *Machine Learning*, 65, 209–230 (2006)
64. T'kindt, V., Bouibede-Hocine, K. and Esswein, C., Counting and enumeration complexity with application to multicriteria scheduling, *Annals of Operations Research*, 153, 215–234 (2007)
65. Tran, N. and Baral, C., Hypothesizing about signaling networks, *Journal of Applied Logic*, 7(3), 253–274 (2009)
66. Wagner, A., *Robustness and Evolvability in Living Systems*, Princeton University Press (2005)
67. Yamamoto, A., Using abduction for induction based on bottom generalization, in: [10], pp.267–280 (2000)
68. Yamamoto, Y., Inoue, K. and Doncescu, A., Integrating abduction and induction in biological inference using CF-Induction, in: Lodhi, H. and Muggleton, S. (eds.), *Elements of Computational Systems Biology*, pp.213–234. John Wiley & Sons (2010)
69. Zupan, B., Demsar, J., Bratko, I., Juvan, P., Halter, J. A., Kuspa, A. and Shaulsky, G. GenePath: A system for automated construction of genetic networks from mutant data, *Bioinformatics*, 19(3), 383–389 (2003)