Abduction Based Drug Target Discovery Using Boolean Control Network

Célia Biane, Franck Delaplace

Abduction Based Drug Target Discovery Using Boolean Control Network

Célia Biane and Franck Delaplace

IBISC, Univ Evry, Université Paris-Saclay, 91025, Evry, France
celia.biane@ibisc.univ-evry.fr
franck.delaplace@ibisc.univ-evry.fr

Abstract. A major challenge in cancer research is to determine the genetic mutations causing the cancerous phenotype of cells and conversely, the actions of drugs initiating programmed cell death in cancer cells. However, such a challenge is compounded by the complexity of the genotype-phenotype relationship and therefore, requires to relate the molecular effects of mutations and drugs to their consequences on cellular phenotypes. Discovering these complex relationships is at the root of new molecular drug targets discovery and cancer etiology investigation. In their elucidation, computational methods play a major role for the inference of the molecular causal actions from molecular and biological networks data analysis. In this article, we propose a theoretical framework where mutations and drug actions are seen as topological perturbations/actions on molecular networks inducing cell phenotype reprogramming. The framework is based on Boolean control networks where the topological network actions are modelled by control parameters. We present a new algorithm using abductive reasoning principles inferring the minimal causal topological actions leading to an expected behavior at stable state. The framework is validated on a model of network regulating the proliferation/apoptosis switch in breast cancer by automatically discovering driver genes and finding drug targets.

Keywords: dynamical system reprogramming, Boolean control network, abductive reasoning, drug target prediction, etiology of cancer.

1 Introduction

In precision medicine, the discovery of causal genes and efficient drug targets is challenged by the complexity of the genotype-phenotype relationship. A key milestone in this challenge is the ability to understand how cell behaviour arises from the synergistic effect of local molecular interactions [49]. Accordingly, cells are envisioned as a web of macromolecular interactions constituting the "interactome" from which phenotype changes are explained by perturbations of molecular interactions [50]. At the molecular level, the phenotypic changes are assessed by the measure of the state of some molecules, called biomarkers, that are defined as observable and objective characteristics of biological processes. They

are used to assess the shift between normal and pathological conditions [45] and to predict the appropriate treatment [12]. Inferring, from the interactome, the molecular causes of phenotypic switches assessed by biomarkers will thus constitute the root for the development of efficient therapies, by predicting the actions at the molecular level directing cells from a diseased toward a healthy state.

In cancer, cells acquire phenotypes with characteristic cancerous hallmarks such as uncontrolled proliferative activity, apoptosis resistance and invasiveness [18]. These phenotypes are caused by multigenic mutations altering molecular interactions. Therefore, a preliminary issue concerns the definition of the effects of mutations on the interactome. In [57], the authors relate mutations to their network effect and introduce the notion of edgetic perturbations of molecular networks: nonsense mutation, out-of-frame insertion or deletion and defective splicing are interpreted as node or arc deletions whereas missense mutation and in-frame insertion or deletion can be modelled as node or arc addition. Moreover, in [10], the authors classify mutations according to the way they affect signalling networks and distinguish mutations that constitutively activate or inhibit enzymes and mutations that rewire the network interactions. The effect of mutations on molecular networks can thus be described as elementary topological actions of deletion or insertion of nodes and arcs. Symmetrically, targeted therapies switch cancer cells phenotype toward growth arrest and apoptosis. Their actions can also be interpreted as network rewiring [12]. A phenotypic switch following mutations or targeted therapies is therefore considered as the observable trait of a *dynamical system reprogramming* caused by *topological network actions* (TN-action).

The inference of TN-actions would provide major insights for etiological investigation of disease, molecular pathogenesis and drug targets prediction by assimilating them to the effects of causal gene mutations (*a.k.a*, drivers) or actions of drugs. In this endeavour, it is worth noticing that generate-and-test method checking the TN-actions exhaustively is often pointless. Indeed, assuming that an expected phenotypic switch results from the application of a specific gene action up to m amongst n genes, then the number of trials[1] equals $\sum_{k=1}^{m} \binom{n}{k}$. For example, the number of trials for targeting up to 10% on 100 genes exceeds 19 billions[2]. Hence, automatically inferring the TN-actions from observable effects is essential to meet this challenge. By considering biomarkers as the entry point of the inference, the issue thus refers to an inverse problem (*ie.*, causes discovery from effects) deducing the sufficient TN-actions from biomarker-based properties variation at stable states.

In this article, we introduce a theoretical framework for TN-action based system reprogramming formalized by Boolean control network. Based on this framework, we develop an algorithm inferring the causal TN-actions that reprogram a Boolean network, redirecting its dynamics to fulfil an expected property. The article is organised as follows: first, we define the Boolean control network framework (Section 2), then we present the inference of causal actions represented by

---

[1] corresponding to the number of parts of size 1 to m in a set with n elements.
[2] Exactly 19 415 908 147 835 trials.

control parameters based on abduction principle (Section 3) and finally, we show its application in breast cancer (Section 4).

2 Boolean control network

In this section we first review the main theoretical elements used in this article, namely: propositional logic, Boolean network and then we introduce Boolean control network.

2.1 Propositional logic

A propositional formula is inductively constructed from atoms composed of constants (False/0, True/1) and variables V, unary negation operator $\neg$, and binary logical operators (*e.g.*, $\wedge$/conjunction/AND, $\vee$/disjunction/OR). A *literal* is either an atom or its negation. Given a formula f, $V(f)$ denotes the set of variables occurring in f. For example, let f_α be the propositional formula representing the exclusive OR between atom x_1 and the negation of atom x_2, $f_\alpha = (x_1 \veebar \neg x_2)$, the variables are $V(f_\alpha) = \{x_1, x_2\}$ and the literals are x_1 and $\neg x_2$. Let $X' \subseteq X$ be a subset of variables $f_{\downarrow X'}$ is the restriction of a formula f to the literals involving the variables of X'.

A *cube* syntactically denotes a conjunction of literals and a *clause* a disjunction. In this article, cubes and clauses will be assimilated to literal sets when needed. A *disjunctive normal form* (DNF) of a formula is a disjunction of cubes (*ie.*, $\bigvee_i \bigwedge_{j_i} l_{j_i}$) whereas a *conjunctive normal form* (CNF) is a conjunction of clauses (*ie.*, $\bigwedge \bigvee_{j_i} l_{j_i}$). Any formula can be transformed in DNF or in CNF. For example, a DNF of f_α is $(x_1 \wedge x_2) \vee (\neg x_1 \wedge \neg x_2)$ and a CNF is $(\neg x_1 \vee x_2) \wedge (x_1 \vee \neg x_2)$.

Let an *interpretation* $I : V \to \{0, 1\}$ be a mapping assigning a truth value to each variable[3], a *model* of a formula f, $I \models f$, is an interpretation such that the formula is evaluated to True and a *satisfiable* formula admits a model at least. For example, f_α is satisfiable because the interpretations $I_1 = \{x_1 = 1, x_2 = 1\}$ and $I_2 = \{x_1 = 0, x_2 = 0\}$ are both models of f_α.

Formula f_1 *entails* formula f_2, denoted by $f_1 \models f_2$, if and only if any model of f_1 is also a model of f_2 (*ie.*, $f_1 \models f_2 \overset{\text{def}}{=\!=} \forall I : I \models f_1 \implies I \models f_2$). Hence, the entailment defines a partial order on formulas.

A *minterm* C_I of an interpretation I is the unique cube such that $V(I) = V(C_I)$ fulfilling $I \models C_I$. For the example $C_1 = x_1 \wedge x_2$ and $C_2 = \neg x_1 \wedge \neg x_2$ are the minterms of I_1 and I_2 respectively. A cube C entailing a formula f is said an *implicant* of f and it is *prime* if it ceases to be one when deprived of any literal. Considering the example, C_1, C_2 are both prime implicants of f_α with I_1 and I_2 as model respectively, thus entailing f_α: $C_1 \models f_\alpha, C_2 \models f_\alpha$. Notice that by contrast to a minterm, an implicant does not necessary involve all the variables of the formula (*e.g.*, x_1 is an implicant of $(x_1 \vee x_2) \wedge (x_1 \vee x_3)$).

---

[3] A mapping will be described $x = v$ instead of $x \mapsto v$ for the sake of simplicity.

2.2 Boolean network

A *Boolean network* is a discrete dynamical system operating on Boolean variables X that determines the *state* evolution of variables $x_i \in X$. It is defined as a system of Boolean equations of the form: $x_i = f_i(x_1, \ldots, x_n), 1 \leq i \leq n$ where each f_i is a propositional formula. A Boolean state of s is an interpretation of the variables (*ie.*, $s : X \to \mathbb{B}$) and S_X will denote the set of all states for variables of X.

$$F = \{x_1 = x_2 \vee x_3, x_2 = \neg x_3, x_3 = \neg x_2 \wedge x_1\}$$
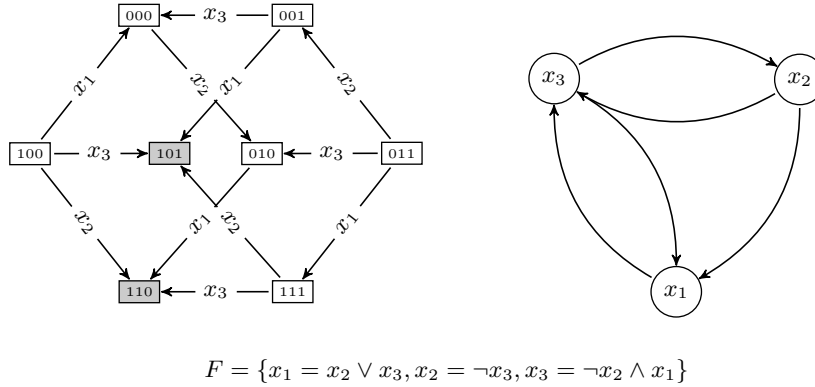
Fig. 1. Model of asynchronous dynamics and interaction graph.

The *model of dynamics* of a Boolean network describes all the trajectories of the system by a labelled transition system. For each transition the states of agents are updated with respect to a predefined updating policy. For the *asynchronous* updating used in the article, one agent only is updated per transition. Hence, the labelled transition system for the asynchronous updating is $\langle \longrightarrow, X, \mathbb{B}^n \rangle$ where the transition relation $\longrightarrow \subseteq S_X \times X \times S_X$ is labelled by the updated agent, $\xrightarrow{x_i}$ such that:

$$s_1 \xrightarrow{x_i} s_2 \stackrel{\text{def}}{=\!=} s_1 \neq s_2 \wedge s_2(x_i) = f_i(s_1) \wedge \forall x_j \in X \setminus \{x_i\} : s_2(x_j) = s_1(x_j).$$

We denote $\longrightarrow = \bigcup_{x_i \in X} \xrightarrow{x_i}$. A state s_2 is said *reachable* from state s_1 if and only if there exists a trajectory defined by the reflexive and transitive closure of the transition relation connecting s_1 to s_2, $s_1 \longrightarrow^* s_2$.

A state s is an *equilibrium* for $\longrightarrow$, if it can be infinitely reached once met, *ie.*, $\forall s' \in S_X : s \longrightarrow^* s' \implies s' \longrightarrow^* s$. An *attractor* is a set of equilibria that are mutually reachable and a *stable state* is an attractor of cardinality 1. In Figure 1, the states 101 and 110 in grey are stable. Stable states remain identical whatever the updating policy as they comply to Definition 1:

$$\text{STBL}_F(s) \stackrel{\text{def}}{=\!=} \forall 1 \leq i \leq n : f_i(s) = s(x_i). \tag{1}$$

An *interaction graph* $\langle X, \longrightarrow \rangle$ portrays the causal interactions between variables of a Boolean network (*cf.*, Figure 1). An interaction $x_i \dashrightarrow x_j$ exists if and only if x_i occurs as literal in a minimal DNF form of f_j, *ie.*,

$$x_i \dashrightarrow x_j \stackrel{\text{def}}{=\!=} x_i \in V(\text{DNF}(f_j)).$$

2.3 Boolean Control Network

Boolean Control Network (BCN) extends Boolean network by adding *control parameters* that are Boolean variables, $u_i \in U$ without equation definition. Hence, a BCN is defined as a function generating Boolean network parametrized by an interpretation of control parameters $\mu \in S_U$, called a *control input*: $F_u : S_U \to (S_X \to S_X)$. For example, an extension of the Boolean network in Figure 1 to a BCN by adding four control parameters u_1, u_2, u_3, u_4 is:

$$F_{u_1, u_2, u_3, u_4} = \begin{cases} x_1 = (x_2 \wedge u_1) \vee x_3, \\ x_2 = \neg(x_3 \vee \neg u_2), \\ x_3 = ((\neg x_2 \wedge x_1) \vee \neg u_3) \wedge u_4 \end{cases} \tag{2}$$

The application of a control input μ to a Boolean control network F_μ therefore reprograms the dynamics. Figure 2 describes the dynamics resulting from the application[4] of two control inputs $\mu_1 = \{u_1 = 0, u_2 = 1, u_3 = 1, u_4 = 1\}$ and $\mu_2 = \{u_1 = 1, u_2 = 1, u_3 = 1, u_4 = 0\}$.

Boolean control network provides a general framework for dynamical system reprogramming. Indeed, let F be an initial Boolean network reprogrammed into an other Boolean network G where the equations are modified, then the Boolean control network $F_u = (u \wedge F) \vee (\neg u \wedge G)$ behaves as F if $u = 1$ and as G if $u = 0$. The switch between F and its reprogramming G now depends on the value of u only. This encoding can be trivially extended to address a family of dynamical systems viewed as the different outcomes of the reprogramming by triggering each particular system from a particular valuation of several control parameters, *e.g.*, $F_{u_1, u_2} = (u_1 \wedge u_2 \wedge F) \vee (\neg u_1 \wedge u_2 \wedge G_1) \vee (u_1 \wedge \neg u_2 \wedge G_2) \vee (\neg u_1 \wedge \neg u_2 \wedge G_3)$ with G_1, G_2, G_3 as reprogramming outcomes. However, the control will be practically specified in another way in order to represent the effective control operated in the real system (Section 2.4).

Finally, a Boolean control network can be associated to a *control constraint* $\Phi : U^m \to \mathbb{B}$ fixing the allowed control inputs.

2.4 Control-freezing category

Amongst the different possibilities to control a Boolean network, we focus on a particular category called *control-freezing* where the control action fixes (freezes)

---

[4] The formulas resulting from the instantiation of the BCN by a control input are simplified.

$$F_{\mu_1} = \begin{cases} x_1 = x_3, \\ x_2 = \neg x_3, \\ x_3 = \neg x_2 \wedge x_1 \end{cases} \qquad\qquad F_{\mu_2} = \begin{cases} x_1 = x_2 \vee x_3, \\ x_2 = \neg x_3, \\ x_3 = 1 \end{cases}$$

$$\mu_1 = \{u_1 = 0, u_2 = 1, u_3 = 1, u_4 = 1\} \qquad \mu_2 = \{u_1 = 1, u_2 = 1, u_3 = 1, u_4 = 0\}$$
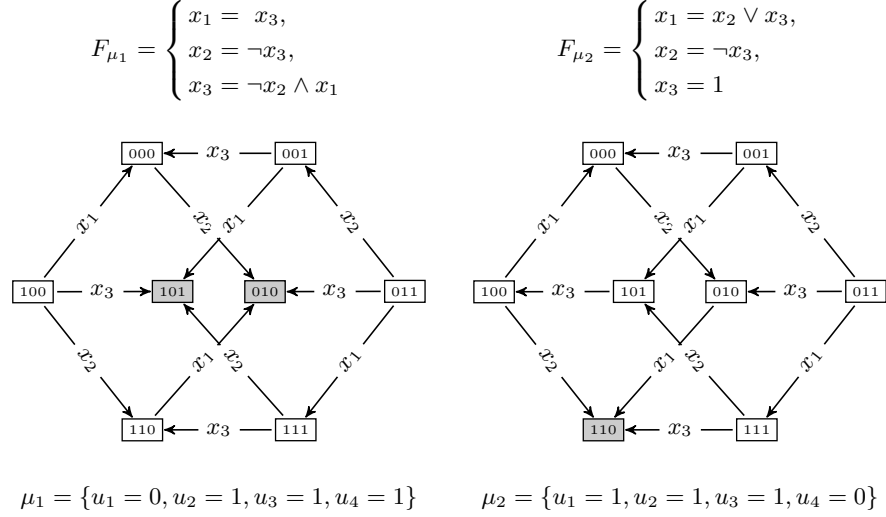
Fig. 2. Modification of the dynamics by control inputs for the example of Figure 1.

the variable states to a specific value. This category models the dynamical after-maths on Boolean network of the TN-actions on the interaction graph. We define two categories of control actions: *Definition-freezing* (D-freezing) that controls the definition of a variable and *Use-freezing* (U-freezing) controlling the use of a variable in an equation defining another variable. Therefore, D-freezing directly assigns an invariant value to variables whereas U-freezing sets locally an invariant value for their use in an equation. The immediate consequence on the interaction graph of a freezing is to totally disconnect a node from its inputs for D-freezing and to remove an arc for U-freezing. Therefore, D-freezing control models node action whereas U-freezing control represents arc action (*cf.*, Section 4 for their interpretation in biological network). The D-freezing parameter governing the freeze of variable x_i will be denoted d_i and the U-freezing parameter is denoted $u_{i,j}$ standing for the control by freeze of the variable x_i in its use in f_j. Moreover, each control parameter has two distinct regimes: either it freezes the variable to a specific value or remains idle. The convention, inspired by the freezing temper-ature of water $0°C$, is as follows: the freezing action is triggered when the control parameter is set to 0 whereas the idle situation corresponds to 1. As the value of a parameter indicates the freezing activity (active or idle), the two possible freezing outcomes 0 or 1 are supported by two distinct parameters respectively denoted $d_i^0, u_{i,j}^0$ and $d_i^1, u_{i,j}^1$. For example, by considering the following controlled equation $x_1 = (\neg x_2) \wedge d_1^0$, d_i^0 will freeze x_1 to 0 if $d_1^0 = 0$ otherwise x_1 behaves as the negation of x_2 (See also (7)).

Control-freezing implementation to Boolean network. The implementation of the freezing control on a Boolean network extends the formulas to obtain the expected control behaviour depending on the type of control parameters: D^0, D^1 or U^0, U^1.

D-freezing control implementation. The D-freezing control of variable x_i consists in adding a D-freezing parameter to formula f_i such that setting $\mu(d_i^k) = 0, k \in \{0, 1\}$ leads to freeze variable x_i to k and remains idle otherwise $(\mu(d_i^k) = 1)$. Formula f_i is completed according to this control behaviour:

$$x_i = f_i(x_1, \ldots, x_n) \wedge d_i^0 \qquad \text{for freezing to 0} \qquad (3)$$

$$x_i = f_i(x_1, \ldots, x_n) \vee \neg d_i^1 \qquad \text{for freezing to 1} \qquad (4)$$

D^0 and D^1 freezing parameters can be combined to trigger the freeze to different values. To avoid a contradictory freeze to 0 and 1 simultaneously, the constraint $\Phi = d_i^0 \vee d_i^1$ is added ensuring the mutual exclusion of the parameter activities.

U-freezing control implementation. The U-freezing control application follows the same principles as the D-freezing control but applied on the occurrence of variables in the equations of other variables.

$$x_j = f_j(x_1, \ldots, x_i \wedge u_{i,j}^0, \ldots, x_n) \qquad \text{for freezing to 0} \qquad (5)$$

$$x_j = f_j(x_1, \ldots, x_i \vee \neg u_{i,j}^1, \ldots, x_n) \qquad \text{for freezing to 1} \qquad (6)$$

Both controls can be also combined with a constraint avoiding to trigger contradictory freezing controls simultaneously (*ie.*, $\Phi = u_{i,j}^0 \vee u_{i,j}^1$).

In Example (2), u_1 is assimilated to the U-freezing parameter of x_2 to 0 ($u_1 = u_{2,1}^0$) used in x_1 definition, u_2 can be interpreted as the U-freezing parameter of x_3 ($u_2 = u_{3,2}^1$), and u_3, u_4 are the D-freezing parameters of x_3 freezing the variable to 1 and 0 respectively ($u_3 = d_3^1, u_4 = d_3^0$). Consequently, the BCN (2) can be rewritten using the appropriate naming convention as:

$$F_{u_{2,1}^0, d_2^0, d_3^2, d_3^1} = \begin{cases} x_1 = \left(x_2 \wedge u_{2,1}^0\right) \vee x_3, \\ x_2 = \neg(x_3 \vee \neg u_{3,2}^1), \\ x_3 = \left((\neg x_2 \wedge x_1) \vee \neg d_3^1\right) \wedge d_3^0 \end{cases} \qquad (7)$$

The control activity is thus fully determined by the parameters assigned to 0 in a control input μ. The *set of active control parameters* collect these parameters to trace the control activity (*ie.*, $\{u_i \in U \mid \mu(u_i) = 0\}$). In the sequel U will represent the set of the freezing control parameters indifferently and $u_i \in U$ a generic freezing control parameter.

3 Control parameters inference

The issue is to formally characterize the basic patterns specifying the changes of the observable molecular traits resulting from biological system reprogramming. Such variations will be questioned at equilibrium conditions in a twofold way: either finding a particular property in some stable states, or finding a particular property in all of them. We thus define two modalities: the *possibility of meeting a property* in at least one stable state (PoP) and the *necessity of meeting a property* in all stable states (NoP). Let p be a Boolean function on states ($p : S_X \to \mathbb{B}$) standing for a property, the PoP and NoP inference problems are defined as follows:

Find a control input μ fulfilling the constraints of Φ such that:

$$\exists s \in S_X : \text{STBL}_{F_\mu}(s) \wedge p(s). \qquad \text{(PoP)} \qquad (8)$$

$$\forall s \in S_X : \text{STBL}_{F_\mu}(s) \implies p(s). \qquad \text{(NoP)} \qquad (9)$$

Different control inputs may be suitable as solutions. For instance, gaining stable state 010 for Boolean network of Figure 1 with parameters defined in (7) can be obtained with the following control inputs:

$$\left\{ u_{2,1}^0 = 0, u_{3,2}^1 = 1, d_3^1 = 1, d_3^0 = 1 \right\}$$
$$\left\{ u_{2,1}^0 = 0, u_{3,2}^1 = 1, d_3^1 = 1, d_3^0 = 0 \right\}$$
$$\left\{ u_{2,1}^0 = 0, u_{3,2}^1 = 1, d_3^1 = 0, d_3^0 = 0 \right\}$$

The plurality of solutions raises the question of their interpretation for identifying the root factors causing the expected effects. The causal factors are defined as the essential actions shifting the dynamics to the objective whereas the casual factors behave neutrally and do not interfere with the objective whatever their valuation. Focusing on the active parameters, only $u_{2,1}^0 = 0$ matters for shifting the dynamics to gain 010 (first solution) since it is shared by all solutions, and without this assignment the system reprogramming fail to reach the expected objective. The other parameters becoming active are casual because they can be set to 0 or 1 without deviating the dynamics to the result.

The set of causal control parameters forms a *core* K^* defined as a minimal active parameter set under the inclusion which is equivalent to the entailment order for cubes. Considering the example, the core $K^* = \{u_{2,1}^0\}$ is included in all other active parameter sets.

Several cores may be found for a given problem. For example, three different cores $\{d_3^1\}, \{u_{2,1}^0\}, \{u_{3,2}^1\}$ enable the loss of equilibrium 110. Hence, the inference algorithm aims at finding all the cores in regards to a reprogramming query formulated by the possibility or the necessity of meeting a property at steady-state.

3.1 Abduction based core inference

Inferring a core corresponds to the determination of control parameters producing an expected effect. In logic finding causes from effects is an abduction

problem. Abduction is a method of reasoning proposing hypotheses that provide the best explanation for observable facts in regards to knowledge of the problem constituting the theory [38, 42, 31]. In propositional logic, a cube C is an abductive explanation of a formula f formalizing the facts with respect to another formula Φ representing the theory if and only if: $C \wedge \Phi \models f$ and C is consistent with Φ (ie., $\Phi \wedge C$ is satisfied). Finding a parsimonious hypothesis introduces the notion of minimal solution which is usually assimilated to a prime implicant. Within this framework, the possibility and the necessity of property (8, 9) are formulated as abduction problems in propositional logic (10,11) by considering that p is a propositional formula. Lemma 1 demonstrates this equivalence.

Find a cube C_μ such that:

$$(C_s \wedge C_\mu) \wedge \phi \models (\text{STBL}_{F_u} \wedge p); \qquad \text{(PoP)} \qquad (10)$$

$$C_\mu \wedge \phi \models (\text{STBL}_{F_u} \implies p); \qquad \text{(NoP)} \qquad (11)$$

where C_s and C_μ are consistent with Φ, $V(C_\mu) = U, V(C_s) = X$ and the stability condition is defined as:

$$\text{STBL}_{F_u} \overset{\text{def}}{=} \bigwedge_{i=1}^{n} (x_i \iff f_i(x_1, \ldots, x_n, u_1, \ldots, u_m)).$$

In Example (7), the components of the problem for gaining state 010 (Figure 2, μ_1) are:

$$
\begin{aligned}
\text{STBL}_{F_u} = \quad & x_1 \iff \left(x_2 \wedge u_{2,1}^0\right) \vee x_3 && \text{Stability condition} \\
\wedge \ & x_2 \iff \neg(x_3 \vee \neg u_{3,2}^1) \\
\wedge \ & x_3 \iff \left((\neg x_2 \wedge x_1) \vee \neg d_3^1\right) \wedge d_3^0 \\
\Phi = \ & d_3^0 \vee d_3^1 && \text{Exclusive activity of } d_3 \\
p = \ & \neg x_1 \wedge x_2 \wedge \neg x_3 && \text{Minterm of } s = 010
\end{aligned}
$$

For the loss of stable state 101 (Figure 2, μ_2), only the property differs, now defined as: $p = \neg(x_1 \wedge \neg x_2 \wedge x_3)$ corresponding to the negation of the minterm of 101.

Lemma 1. *(10) and (11) define the PoP (8) and NoP (9) problems as abductive problems in propositional logic. (See Appendix for the proof.)*

3.2 Core inference algorithm

For a formula f, the core inference consists in finding a satisfiable implicant C^* fulfilling $C^* \models f$ that minimizes the number of negative control parameters $(\neg u_i)$ with respect to the inclusion. The resulting core K^* is trivially deduced by collecting the negative control parameters of C^*. Computing a core is an NP-Hard problem[5]. In this section, we present an algorithm adapted from the

---

[5] by reduction of the minimum hitting set problem.

method developed for prime implicants computation in [41] and based on $0\text{-}1$ Integer Linear Programming ($0\text{-}1$ ILP). A $0\text{-}1$ ILP problem is formulated as:

$$\text{Minimize } \sum_{j=1}^{h} m_j.y_j, \text{subject to } \sum_{j=1}^{h} W_{i,j}.y_j \leq v_i, \text{for } 1 \leq i \leq r, y \in \{0,1\}^h.$$

where y is the unknown vector, and m, v vectors, W matrix are the parameters of the problem.

The method, called ILP-CORE, operates on a formula f in CNF and computes the set of all the cores $\mathcal{K}^*$. The method is based on the translation of the constraints related to core definition into $0\text{-}1$ ILP constraints such that a solution y is a binary representation of an implicant C^*. The algorithm is outlined in Algorithm 1 and the main steps are fully described in the proof of Theorem 2.

Function ILP-CORE(*f: CNF formula*)

    $(\min m.y^T, Wy \leq v) = $ Describe constraints on core as $0\text{-}1$ ILP problem ;
    // $C^* \models f$ `minimizing the number of negative control parameters.`
    $\mathcal{K}^* = \emptyset$;
    repeat
        $y = $ Solve $(\min m.y^T, Wy \leq v)$ with a $0\text{-}1$ ILP solver ;
        if *a solution y is found* **then**
            $K^* = $ Collect the negative control parameters from y;
            $\mathcal{K}^* = \mathcal{K}^* \cup \{K^*\}$;
            Exclude all solutions $K, K^* \subseteq K$ by adding constraints to $Wy \leq v$;
        end
    until *No solution y is found*;
    return $\mathcal{K}^*$ `// the set of all cores`
end

Algorithm 1: Outline of the ILP-CORE algorithm.

Theorem 2. *The* ILP-CORE *algorithm finds all and only the cores. (See Appendix for the proof.)*

To properly specify the PoP and NoP resolutions, the method is called with different formulas specifying the query. Applied to PoP (10), the complete formula is passed as parameter since literals of C^* contain control parameters as well as variables identifying the state. For NoP (11), as C^* must contain control parameters only, each clause is then restricted to control parameters by removing the literals involving state variables (ie., $x_i \in X$). The constraints on control parameters Φ are already in CNF form by definition (Section 2.4).

$$\text{ILP-CORE}(\text{CNF}(\text{STBL}_{F_u} \wedge p) \wedge \Phi) \qquad \text{(PoP)}$$
$$\text{ILP-CORE}(\text{CNF}(\text{STBL}_{F_u} \implies p)_{\downarrow U} \wedge \Phi) \text{ (NoP)}$$

3.3 Related works

BCN was recently introduced in systems biology to provide the theoretical foundations and computational methods for investigating cell fate reprogramming and therapeutic target discovery. In [25] the authors apply a stuck-at fault model to simulate drug intervention in an acyclic growth factors pathway by a generate-and-test method. stuck-at fault model mimics the defects on combinatorial logic circuit which were assimilated here to malignant mutations. Based on this model, authors identify drug actions for single mutations by correcting all possible single faults. This framework was improved by [27] using a Max-SAT based method dedicated to acyclic networks in order to directly compute the control parameter values and final states. Inferring the drug targets on a network is also developed by [35] using algebraic techniques (Gröbner basis) in order to modify the system dynamics for creating or avoiding particular stable states. In [54], the authors propose a heuristic method with the same goal but focused on the control of key-nodes stabilizing "motifs" identifying sub-networks. Finally, we have introduced the principle of the abductive inference of cores for drug target discovery in [3] which is significantly extended here, in particular with the formalization and the generalization of the TN-actions as control freezing, and with a more efficient method for the core inference.

Our approach follows a similar orientation of these works by using BCN for modelling disease and drug actions. By comparison, the target discovery is modelled in an original way as an abductive problem. The resulting framework supports any kind of networks including cycles with actions applied on both nodes and arcs and find multiple targets qualifying the parsimonious TN-actions (cores) reprogramming the system. The proposed algorithm infers the causes of expected properties met at stable states and we formalize their query in a general setting using propositional formulas with the Necessity and Possibility modalities.

4 Application to breast cancer

This section shows the application of TN-actions inference for the study of breast cancer. Mainly, cancer cells differ from normal cells by their uncontrolled proliferation and apoptotic evasion. Accordingly, targeted drugs aim at inducing apoptosis or stop the proliferation of cancer cells [18]. We therefore developed a model (Section 4.1) focusing on the regulation of division and apoptosis. We infer the causal TN-actions leading to a loss or gain of apoptosis (Section 4.2) and then analyse the results (Section 4.3).

4.1 Aptoptosis/Cell Division Boolean network

The model focuses on the regulation of cell division and apoptosis by the EGFR signalling pathway and a BRCA1/TP53 DNA damage response module. These genes have been identified as central in the process of tumor formation in breast

cancer [23, 36]. The model incorporates the positive and negative interactions between nuclear TP53 and MDM2 described by [8], the main messengers of the PI3K/AKT and MAPK signalling following EGFR activation described by [52] and adds BRCA1 and PARP1 regulation of DNA damage. These pathways are gathered into a unique Boolean network through the lens of their role in the regulation of the G1/S transition and the triggering of apoptosis in case of DNA damage. The corresponding Boolean network[6], constructed from published litterature and signalling pathways databases (KEGG [21] and Signor [39]), is shown in Figure 3 and the molecular mechanism for each interaction is detailed and referenced in Appendix (see Table 2). The Boolean dynamics is bistable characterizing two cellular functions in normal cells: either (1) the cell enters division by activation of the G1/S transition and inhibition of apoptosis, or (2) it enters in apoptosis and arrest the cell cycle.

4.2 Inference query

As network reprogramming effects biomarker profile changes, it is required to 1) identify the biomarkers discriminating phenotypes and 2) define the reprogramming queries based on these biomarkers for causal genes and drug actions inference.

Since the proliferative activity of cells depends on the balance between division and apoptosis, we selected CYCLIN D1 and BAX as biomarkers as they are the key effector of the G1/S transition of cell division and initiation of apoptosis [2, 17]. The pair $(CYCLIN\,D1, BAX)$ distinguishes four phenotypes: apoptosis, division, quiescence (apoptosis balanced by division) and dormancy (neither apoptosis nor division) [43] through to the following signatures: $(0, 1)$ for apoptosis, $(1, 0)$ for division, $(0, 0)$ for quiescence and $(1, 1)$ for dormancy.

Since cancer cells are characterized by their inability to trigger apoptosis, the reprogramming query for the inference of causal genes corresponds to the loss of apoptosis. Conversely, as drugs induce apoptosis in cancer cells, the reprogramming query for the inference of drug actions corresponds to the gain of apoptosis. Apoptosis is formalized as a property by the minterm of $(0, 1)$ signature: $p = \neg CYCD1 \wedge BAX$. The loss of apoptosis thus corresponds to the necessity of $\neg p$ since the apoptosis must not occur in any stable state. To recover this marking, the query can be either the necessity or the possibility of p. We have tested both and the solutions providing stable states are the same.

Finally, the genetic events are modelled by control parameters as follows: the loss of expression of a gene following loss-of-function mutations or other genetic events such as gene deletion corresponds to D^0-freezing; gene over-expression following gain-of-function mutations or other genetic events such as gene amplification are represented by D^1-freezing; and the loss of interactions between two molecules is interpreted as U^0-freezing. The Boolean network (Figure 3) is automatically completed with control parameters by following the rules set out

---

[6] For the sake of simplicity, the names of genes (by convention written in upper case letters) can also denominate the proteins they encode.

$$EGFR = \neg BRCA1$$
$$ERK1/2 = EGFR$$
$$PI3K = \neg PTEN \wedge EGFR$$
$$AKT = PI3K$$
$$GSK3\beta = \neg AKT$$
$$MDM2 = AKT \wedge TP53$$
$$TP53 = \neg MDM2 \wedge (BRCA1 \vee \neg PARP1)$$
$$PTEN = TP53$$
$$PARP1 = ERK1/2$$
$$BRCA1 = \neg CYCD1$$
$$BCL\text{-}2 = AKT$$
$$BAX = \neg BCL\text{-}2 \wedge TP53$$
$$CYCD1 = (\neg GSK3\beta \wedge ERK1/2) \vee (\neg BRCA1 \wedge PARP1)$$

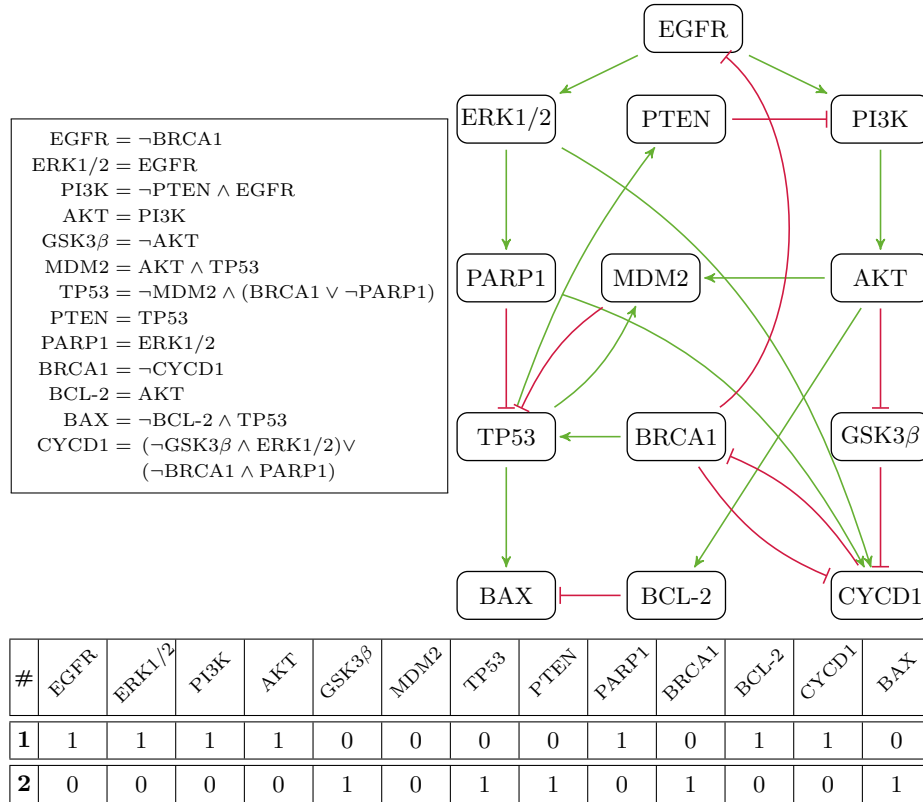| #	EGFR	ERK1/2	PI3K	AKT	GSK3β	MDM2	TP53	PTEN	PARP1	BRCA1	BCL-2	CYCD1	BAX
1	1	1	1	1	0	0	0	0	1	0	1	1	0
2	0	0	0	0	1	0	1	1	0	1	0	0	1

Fig. 3. Boolean network (left) with its regulatory graph (right) representing the acti-vatory (green) and inhibitory (red) interactions, and stable states (below).

in Section 2.4. Notice that U^1-freezing does not seem interpretable in terms of biological events and not used here.

4.3 Analysis of the results.

We inferred the actions from combination of D^0/D^1-freezing on all variables (molecules) except markers and the U^0-freezing on all interactions separately to compare them. The computed TN-actions are shown in Table 1. The TN-actions for the gain of apoptosis have been inferred from the model with BRCA1-deficiency (BRCA1 = 0).

Applied to the loss of apoptosis with $D-$freezing, the method retrieves the main driver genes identified in breast cancer namely BRCA1, TP53, PI3K and EGFR [20, 5]. Moreover, it segregates tumor suppressor genes (*ie.*, frequently affected by gain-of-function mutations in cancers) from oncogenes (*ie.*, frequently affected by loss-of-function mutations in cancers) [11, 29]: D^0-frozen genes all

- Health → Cancer: necessary loss of apoptosis -

<table>
<tr><td rowspan="9" style="writing-mode:vertical">NODE ACTION</td><td colspan="2">*Single D-freezing*</td><td rowspan="12" style="writing-mode:vertical">ARC ACTION</td></tr>
</table>

	Single D-freezing
	BRCA1 = 0
	TP53 = 0
	PI3K = 1
	AKT = 1
	BCL-2 = 1
	MDM2 = 1

	Double D-freezing
	GSK3β = 0, ERK1/2 = 1
	PTEN = 0, EGFR = 1
	GSK3β = 0, EGFR = 1

Single U^0-freezing
TP53 $\longrightarrow$ BAX

Double U^0-freezing
BRCA1 $\longrightarrow$ EGFR, TP53 $\longrightarrow$ PTEN
BRCA1 $\longrightarrow$ EGFR, BRCA1 $\longrightarrow$ CYCD1
BRCA1 $\longrightarrow$ EGFR, BRCA1 $\longrightarrow$ TP53
BRCA1 $\longrightarrow$ EGFR, PTEN $\longrightarrow$ PI3K
BRCA1 $\longrightarrow$ EGFR, GSK3β $\longrightarrow$ CYCD1

- BRCA1 mutation (Cancer) → Cell death: possible gain of apoptosis -

	Single D-freezing
	BRCA1 = 1
	PARP1 = 0
	ERK1/2 = 0
	EGFR = 0

Single U^0-freezing
ERK1/2 $\longrightarrow$ PARP1
EGFR $\longrightarrow$ ERK1/2

Double U^0-freezing
PARP1 $\longrightarrow$ CYCD1, PARP1 $\longrightarrow$ TP53

Table 1. Freezing actions causing the gain or loss of apoptosis.

correspond to tumour suppressors and D^1-frozen genes to oncogenes. For the gain of apoptosis after application of BRCA1 deficiency, the single D-freezing inferred actions recover the necessity of blocking PARP1, the synthetic lethal partner of BRCA1. The pair BRCA1/PARP1 are called synthetic lethal partners because the use of PARP inhibitors in patients with BRCA1-deficiency prevents any possibility of DNA-repair resulting in permanent DNA damage inducing apoptosis of the cancer cell [16, 28]. Finding such partnerships is critical for anticancer treatment [19] but since the cancer target differs from the drug target, they are hard to recover experimentally and computationally.

The algorithm also predicts double D-freezing actions for the necessary loss of apoptosis which suggest that overexpression of EGFR alone would not be sufficient to provoke a cancerous phenotype and must be combined with either loss of PTEN or GSK3β. The validation of such result is less obvious than the former and is based on the concomittent overexpression of EGFR and loss of PTEN/GSK3β. Work in [26] confirms the existence of a co-occurence of EGFR over-expression and loss of PTEN in 20% of the tumors of the studied population. Moreover, authors also show that PTEN loss is associated to resistance to EGFR inhibitors. Similarly in erlotinib resistant model cell lines [4] it has been observed

that GSK3β was upregulated. Thus, these works suggest the existence of the predicted cooperation between these genes.

It is also predicted that EGFR inhibition would be synthetic lethal with BRCA1 mutations. This is supported by the observation that the proliferation properties of BRCA-deficient cells are sensitive to EGFR inhibition by erlotinib [5]. We found no published work suggesting that ERK1/2 inhibition in such cells would be synthetic lethal.

In summary, in the studied model the method accurately predicts cancerous genes and drug targets and segregate oncogenes from tumor suppressors. The inference also recovers cooperative gene mutations and synthetic lethal partnerships. The double freezing results provide some insights on the necessary cooperative combination of perturbations that are difficult to assess experimentally [53, 40]. Moreover by inferring cores, the method separate causal genes to casual ones (passengers) and determine frequent drivers as well as rare ones which is more difficult to obtain by statistical analysis that prioritize genes from the frequency of their occurrence [51]. Usually, drivers are classified in subtypes where a specific drug target is associated for each subtype. In the proposed approach the drug target may be directly inferred from the application of the TN-actions corresponding to drivers on the initial boolean network. Finally, arc inference (U^0−freezing) refines the results on nodes (D^0−freezing) and, to the best of our knowledge, the resulting predictions are not experimentally confirmed.

5   Conclusion

In this article, we have proposed a modelling framework discovering the reprogramming actions of a dynamical system using BCN and designed a new inference method based on abduction that identifies the minimal causes reprogramming the network. A library called PROTAXION was developed in Mathematica to support the application on concrete cases. It has been validated on a breast cancer model and has shown that the method can retrieve driver genes and drug targets.

A perspective of this work is to include the notion of resistance in the inference. Two sorts of resistances were established: the primary arising prior to a classical treatment and the secondary which is an adaptive negative response to a treatment. As the method infers all the causes responsible for a biomarker profile shift, the primary resistance is interpreted in our framework as the variation of the input Boolean network of a patient in comparison to a generic one in which the drug targets were deduced. In this context, we need to specialize the network to a patient. The issue for the secondary resistance is more complex and necessitates to predict the further alterations of the network once a TN-action is applied. The prediction of secondary resistance requires to extend the BCN model by including the notion of temporal sequence of control inputs instead of a single control input.

Appendix

Proofs

Proof (Lemma 1). To prove the equivalence of PoP (10) and NoP (11) problems with Definitions (10,11) based on propositional logic, we need to demonstrate that 1) the stability condition can be formulated as a propositional formula 2) we can find an equivalent formulation for these Definitions using abduction on propositional formulas.

Stability condition defined by a propositional formula. By definition (1), the stability condition for a BCN involving control parameters $u = (u_1, \ldots, u_m)$ is defined as: $\forall\ 1 \leq i \leq n : x_i = f_i(x, u)$. As the set of equations is finite, the condition can be rephrased as:

$$\bigwedge_{i=1}^{n} (x_i = f_i(x, u)).$$

As the equivalence $p_1 \iff p_2$ is satisfied if and only if $I(p_1) = I(p_2)$, the equality can be formulated by an equivalence, finally defining the stability condition as:

$$\bigwedge_{i=1}^{n} \left(x_i \iff f_i(x, u) \right).$$

Now we examine the equivalent formulation of PoP and NoP definitions given in (8,9) using abduction on propositional formulas and leading to Definitions (10, 11).

PoP defined as adbuctive problem in propositional logic. Definition (8) can be reformulated by introducing the entailment as:

$$\exists s \in S_X : (C_s \wedge C_\mu) \wedge \phi \models \text{STBL}_{F_\mu} \wedge p,$$

where C_s and C_μ are the minterms of a state s and a control input μ respectively. As $\text{STBL}_{F_\mu} \wedge p$ contains state variables and control parameters an implicant C fulfilling $C \wedge \phi \models \text{STBL}_{F_\mu} \wedge p$ can be divided into $C = C_s \wedge C_\mu$. C can be possibly completed by missing variables if needed. C_s is the minterm of an interpretation $s \in S_X$. Therefore determining the existence of an implicant C involving the variables of X and U is equivalent of proving the existence of its interpretation (*ie.*, $(C \models f) \iff (\exists s, s \cup \mu \models f)$). Thus, the existential quantifier of s can be removed.

NoP defined as adbuctive problem in propositional logic.

($\Longrightarrow$) Definition (9) of NoP can be expressed using formula entailment as:

$$\forall s \in S_X : (C_s \wedge C_\mu) \wedge \phi \models (\mathrm{STBL}_{F_\mu} \implies p),$$

where C_s and C_μ are the minterms of s and μ respectively.

We define $P = (\mathrm{STBL}_{F_\mu} \implies p)$ and s_{-i} a state s deprived of its ith component.

Let C_μ be a solution, for all states s there exists a state $(s_{-i}, \neg s_i)$ such that both fulfil the following $C_s \wedge C_\mu \wedge \phi \models P$ and $(C_{(s_{-i}, \neg s_i)} \wedge C_\mu \wedge \phi \models P$ respectively as the property holds for all states. By applying the consensus theorem [42], we deduce that $C_{s_{-i}} \wedge C_\mu$ is also an implicant by removing s_i since it appears positively and negatively in two implicants. As this simplification can be applied for all states $s \in S_x$, we finally have:

$$\forall s_{-i} : (C_{s_{-i}} \wedge C_\mu) \wedge \phi \models P.$$

By following the same reasoning for all s_{-i} states with the deprivation of $s_j, j \neq i$, we deduce the same conclusion for $s_{-\{i,j\}}, \forall j \neq i$. This can be repeated until no variables exist $(C_\emptyset)$. As $C_\emptyset = 1$ by definition, we thus conclude that:

$$C_\mu \wedge \phi \models P.$$

($\Longleftarrow$) Conversely, assume that $C_\mu \wedge \phi \models P$ with $V(C_\mu) = U$ we deduce that for all C such that C is consistent with Φ we have $C \wedge C_\mu \wedge \phi \models P$. As for all $s \in S_X, C_s$ only involves variables of X and Φ is a constraint on parameters with variables of U then $V(C_s) \cap V(\Phi) = \emptyset$. Two satisfiable cubes with distinct variables are necessary consistent. Thus, we conclude that: $\forall s \in S_X : (C_s \wedge C_\mu) \wedge \phi \models P$.

In conclusion, the solution can always be determined by focusing on control parameters only for NoP. $\qquad\square$

Proof (Theorem 2). The proof of the Theorem lies in the translation of the different constraints related to the determination of the cores in integer linear constraints.

Translation of constraints on C^* to linear constraints: The main issue of the method is to translate the required constraints applied on C^* as a set of linear constraints.

Let $f = C_1 \wedge C_2 \wedge \ldots \wedge C_r$ be the input formula in CNF where each C_i is a clause, we define $L_f = \bigcup_{i=1}^{r} C_i$ the set of literals appearing in f. We associate to each literal $l_j \in L_f$ a $\{0,1\}$-variable denoted y_{l_j} meaning that the vector y is indexed by the literals of L_f.

Objective function. Let $\bar{N}$ be the set of control parameters occurring negatively in L_f, ie., $\bar{N} = \{u_j \in U \mid \neg u_j \in L_f\}$, as each core should minimize their occurrences, the objective function is the sum of these negative control parameters:

$$\sum_{u_j \in \bar{N}} y_{\neg u_j}$$

.

Clauses defined by inequalities. By definition of abduction, if there exists an implicant C^* of a formula f which is consistent with a theory, then C^* is satisfiable and thus also f. A formula in CNF is satisfiable if and only if all its clauses are satisfiable and a clause is satisfiable if and only if at least one of its literal is satisfiable. Therefore, an implicant of this formula is a cube formed by taking at least one literal from each clause. This condition is formulated by a constraint for each clause, such that:

$$\forall C_i : \sum_{l_j \in C_i} y_{l_j} \geq 1$$

Satisfiability of C^* defined by inequalities. As the implicant C^* is satisfiable, it cannot contain both a literal and its negation leading to an antilogy. Let $P = \{v_j \in X \cup U \mid v_j \in L_f \wedge \neg v_j \in L_f\}$ be the set of variables occurring both positively and negatively in L_f, then we have the following constraints excluding at least the positive or the negative literal for the variables of P:

$$\forall v_j \in P : y_{v_j} + y_{\neg v_j} \leq 1$$

Conversion of y into a core K^*. y is a binary representation of a cube C^* where $y_{l_j} = 0$ means that l_j does not belong to C^* and $y_{l_j} = 1$ means that l_j is contained in C^*, namely $C^* = \{l_i \mid y_{l_i} = 1\}$. A core K^*, being the set of negative control parameters in C^*, is deduced from y as follows:

$$K^* = \{u_j \in U \mid y_{\neg u_j} = 1\}.$$

Exclusion of all sets including the core as further solutions. The application of the algorithm computes one solution y from which a core K^* is deduced. Thus, we need to exclude it and any set including it to find other solutions. For this, we add the constraints such that a solution cannot contain exactly the same negative control parameters as the found core. Therefore the sum of the y values of the negative control parameters belonging to the core must be less than its cardinality. This linear constraint is expressed from K^* as follows:

$$\sum_{u_j \in K^*} y_{\neg u_j} \leq |K^*| - 1.$$

The method is then iterated until no more core is found to finally provide the set of all cores. □

Molecular mechanisms of the Boolean network model

EGFR is a member of the epidermal growth factors (EGF) receptors family, it responds to extracellular stimulation at the cellular membrane. In turn, it activates the MAPK pathway (represented by ERK1/2) and the PI3K pathway. These two pathways stimulate the entry in the cell division cycle and inhibit apoptosis. To do so, ERK1/2 activates CYCLIN D1 (an effector of the G1/S transition of the cell cycle) and inhibits apoptosis through the activation of PARP1 while PI3K signalling is mediated by AKT that 1) releases CYCLIN D1 of its inhibition by GSK3β and 2) activates BCL-2, an inhibitor of BAX (an effector of apoptosis). In the case of DNA damage, a BRCA1/TP53 module is responsible for the control of cell cycle arrest and triggering of apoptosis. To do so, TP53 activates BAX, inhibits the PI3K pathways through the activation of PTEN (an inhibitor of PI3K). Conversely, TP53 is inhibited by the PI3K pathway at the level of AKT, that activates MDM2. BRCA1 is involved in an inhibitory loop with CYCLIN D1 : active BRCA1 provokes cell cycle arrest at the G1/S transition checkpoint, and avoid subsequent activation by inhibiting EGFR. Morevoer, BRCA1 has been shown to activate TP53 and its subsequent activation of BAX [34, 55]. Finally, PARP1, which is involved in efficient DNA repair, is activated by the MAPK pathway and subsequently inhibits TP53 and activates CYCLIN D1.

Boolean function	Molecular mechanisms	References
EGFR = ¬BRCA1	BRCA1 inhibits EGFR through transcriptional, post-transcriptional and post-translational mechanisms.	[5, 24]
ERK1/2 = EGFR	EGFR activation induce an increase of enzymatic activity of ERK1 and ERK2.	[56, 15]
PI3K = ¬PTEN ∧ EGFR	EGFR activation leads to the activation of PI3K by phopshorylation, PTEN inhibits PI3K activity by catalyzing its inverse reaction.	[13, 15]
AKT = PI3K	AKT is activated via PI3K.	[13, 15]
GSK3β = ¬AKT	AKT inhibits GSK3β by phopshorylation	[6, 13]
MDM2 = AKT ∧ TP53	MDM2 gene transcription is activated by nuclear TP53, AKT is necessary for MDM2 to be translocated into the nucleus where it binds TP53	[32, 33]
TP53 = ¬MDM2∧ (BRCA1 ∨ ¬PARP1)	The binding of MDM2 to TP53 blocks TP53 transcriptional activity and MDM2 exports TP53 to the cytoplasm and targets it for proteasomal degradation. BRCA1 stimulates TP53 transcription, while PARP1 inhibits it.	[32–34, 55, 30, 37]
PTEN = TP53	TP53 activates PTEN transcription	[32, 44]
PARP1 = ERK1/2	ERK1 activates PARP1 by phosphorylation.	[9, 46]
BRCA1 = ¬CYCD1	CYCLIN D1/Cdk4 complex inhibits BRCA1 by phopshorylation	[22]
BCL-2 = AKT	AKT activates BCL-2 by phosphorylation of Bad (not shown) consequently releasing BCL-2 inhibition by Bad	[7]
BAX = ¬BCL-2 ∧ TP53	BAX gene transcription is directly activated by TP53 and its translocation to its active site (mitochondria) is blocked by BCL-2.	[47, 48]
CYCD1 = (¬GSK3β ∧ ERK1/2) ∨(¬BRCA1 ∧ PARP1)	CYCLIN D1 degradation is regulated both dependently and independently of GSK3 levels (OR), BRCA1 provokes cell cycle arrest at the G1/S transition checkpoint, PARP inhibition induces cell cycle arrest suggesting that PARP activates cell cycle transitions	[37, 34, 14, 1]

Table 2. Detailed molecular mechanisms driving the Boolean network construction

References

1. John P Alao. The regulation of cyclin D1 degradation: roles in cancer development and the potential for therapeutic invention. *Molecular cancer*, 6(1):1, 2007.

2. V Baldin, J Lukas, MJ Marcote, M Pagano, and G Draetta. Cyclin D1 is a nuclear protein required for cell cycle progression in g1. *Genes & development*, 7(5):812–821, 1993.

3. Célia Biane, Franck Delaplace, and Tarek Melliti. Abductive network action inference for targeted therapy. In *Static Analysis and Systems Biology*, 2016.

4. Gregory M Botting, Ichwaku Rastogi, Gagan Chhabra, Marie Nlend, and Neelu Puri. Mechanism of resistance and novel targets mediating resistance to EGFR and c-Met tyrosine kinase inhibitors in non-small cell lung cancer. *PloS one*, 10(8):e0136155, 2015.

5. Laura N. Burga, Hai Hu, Ashish Juvekar, Nadine M. Tung, Susan L. Troyan, Erin W. Hofstatter, and Gerburg M. Wulf. Loss of BRCA1 leads to an increase in epidermal growth factor receptor expression in mammary epithelial cells, and epidermal growth factor receptor inhibition prevents estrogen receptor-negative cancers in BRCA1-mutant mice. *Breast Cancer Research*, 13(2):R30, 2011.

6. L C Cantley and B G Neel. New insights into tumor suppression: PTEN suppresses tumor formation by restraining the phosphoinositide 3-kinase/AKT pathway. *Proceedings of the National Academy of Sciences of the United States of America*, 96(8):4240–4245, 1999.

7. F Chang, JT Lee, PM Navolanic, LS Steelman, JG Shelton, WL Blalock, RA Franklin, and JA McCubrey. Involvement of PI3K/Akt pathway in cell cycle progression, apoptosis, and neoplastic transformation: a target for cancer chemotherapy. *Leukemia*, 17(3):590–603, 2003.

8. Andrea Ciliberto, Béla Novák, and John J Tyson. Steady states and oscillations in the p53/Mdm2 network. *Cell cycle*, 4(3):488–493, 2005.

9. Malka Cohen-Armon, Leonid Visochek, Dana Rozensal, Adi Kalal, Ilona Geistrikh, Rodika Klein, Sarit Bendetz-Nezer, Zhong Yao, and Rony Seger. DNA-Independent PARP-1 Activation by Phosphorylated ERK2 Increases Elk1 Activity: A Link to Histone Acetylation. *Molecular Cell*, 25(2):297–308, 2007.

10. Pau Creixell, Erwin M. Schoof, Craig D. Simpson, James Longden, Chad J. Miller, Hua Jane Lou, Lara Perryman, Thomas R. Cox, Nevena Zivanovic, Antonio Palmeri, Agata Wesolowska-Andersen, Manuela Helmer-Citterich, Jesper Ferkinghoff-Borg, Hiroaki Itamochi, Bernd Bodenmiller, Janine T. Erler, Benjamin E. Turk, and Rune Linding. Kinome-wide Decoding of Network-Attacking Mutations Rewiring Cancer Signaling. *Cell*, 163(1):202–217, 2015.

11. Carlo M. Croce. Oncogenes and cancer. *New England Journal of Medicine*, 358(5):502–511, 2008. PMID: 18234754.

12. Peter Csermely, Tamàs Korcsmàros, Huba J M Kiss, Gàbor London, and Ruth Nussinov. Structure and dynamics of molecular networks: A novel paradigm of drug discovery: A comprehensive review. *Pharmacology and Therapeutics*, 138(3):333–408, 2013.

13. Megan Cully, Han You, Arnold J Levine, and Tak W Mak. Beyond PTEN mutations: the PI3K pathway as an integrator of multiple inputs during tumorigenesis. *Nature reviews. Cancer*, 6(3):184–192, 2006.

14. Chu-Xia Deng. Brca1: cell cycle checkpoint, genetic instability, DNA damage response and cancer evolution. *Nucleic acids research*, 34(5):1416–1426, 2006.

15. Pilar Eroles, Ana Bosch, J. Alejandro Pérez-Fidalgo, and Ana Lluch. Molecular biology in breast cancer: Intrinsic subtypes and signaling pathways. *Cancer Treatment Reviews*, 38(6):698–707, 2012.

16. Hannah Farmer, Nuala McCabe, Christopher J Lord, Andrew NJ Tutt, Damian A Johnson, Tobias B Richardson, Manuela Santarosa, Krystyna J Dillon, Ian Hickson,

Charlotte Knights, et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature*, 434(7035):917–921, 2005.

17. Sudhir Gupta. Molecular signaling in death receptor and mitochondrial pathways of apoptosis (review). *International journal of oncology*, 22(1):15–20, 2003.

18. Douglas Hanahan and Robert A. Weinberg. Hallmarks of cancer: The next generation. *Cell*, 144(5):646–674, 2011.

19. William G. Kaelin. The concept of synthetic lethality in the context of anticancer therapy. *Nature Reviews Cancer*, 5(9):689–98, 2005.

20. Cyriac Kandoth, Michael D McLellan, Fabio Vandin, Kai Ye, Beifang Niu, Charles Lu, Mingchao Xie, Qunyuan Zhang, Joshua F McMichael, Matthew A Wyczalkowski, et al. Mutational landscape and significance across 12 major cancer types. *Nature*, 502(7471):333–339, 2013.

21. Minoru Kanehisa, Miho Furumichi, Mao Tanabe, Yoko Sato, and Kanae Morishima. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research*, 45(D1):D353–D361, 2017.

22. K Kehn, R Berro, A Alhaj, ME Bottazzi, WI Yeh, Z Klase, R Van Duyne, S Fu, and F Kashanchi. Functional consequences of cyclin D1/BRCA1 interaction in breast cancer cells. *Oncogene*, 26(35):5060–5069, 2007.

23. Walter Kolch, Melinda Halasz, Marina Granovskaya, and Boris N Kholodenko. The dynamic control of signal transduction networks in cancer cells. *Nature Publishing Group*, (August), 2015.

24. Easwari Kumaraswamy, Karen L Wendt, Laura A Augustine, Shane R Stecklein, Emily C Sibala, Dan Li, Sumedha Gunewardena, and Roy A Jensen. BRCA1 regulation of epidermal growth factor receptor (EGFR) expression in human breast cancer cells involves microRNA-146a and is critical for its tumor suppressor function. *Oncogene*, 34(33):4333–4346, 2015.

25. Ritwik Layek, Aniruddha Datta, Michael Bittner, and Edward R Dougherty. Cancer therapy design based on pathway logic. *Bioinformatics*, 27(4):548–555, 2011.

26. Ji Yun Lee, Mineui Hong, Seung Tae Kim, Se Hoon Park, Won Ki Kang, Kyoung-Mee Kim, and Jeeyun Lee. The impact of concomitant genomic alterations on treatment outcome for trastuzumab therapy in HER2-positive gastric cancer. *Scientific reports*, 5:9289, 2015.

27. Pey-Chang Kent Lin and Sunil P Khatri. Application of Max-SAT-based ATPG to optimal cancer therapy design. *BMC genomics*, 13 Suppl 6(Suppl 6):S5, 2012.

28. Luca Livraghi and Judy E Garber. PARP inhibitors in the management of breast cancer: current data and future prospects. *BMC medicine*, 13(1):1, 2015.

29. Harvey Lodish and S Lawrence Zipursky. Molecular cell biology. *Biochemistry and Molecular Biology Education*, 29:126–133, 2001.

30. Maria Malanga, Jutta M Pleschke, Hanna E Kleczkowska, and Felix R Althaus. Poly (ADP-ribose) binds to specific domains of p53 and alters its DNA binding functions. *Journal of Biological Chemistry*, 273(19):11839–11843, 1998.

31. P Marquis. Extending abduction from propositional logic to first order logic. *Lecture Notes in Artificial Intelligence*, 535:141–155, 1991.

32. Lindsey D. Mayo and David B. Donner. The PTEN, Mdm2, p53 tumor suppressor-oncoprotein network. *Trends in Biochemical Sciences*, 27(9):462–467, 2002.

33. Ute M. Moll and Oleksi Petrenko. The MDM2-p53 Interaction. *Molecular Cancer Research*, 1(14):1001–1008, 2004.

34. PB Mullan, JE Quinn, and DP Harkin. The role of BRCA1 in transcriptional regulation and cell cycle control. *Oncogene*, 25(43):5854–5863, 2006.

35. David Murrugarra, Alan Veliz-Cuba, Boris Aguilar, and Reinhard Laubenbacher. Identification of control targets in boolean molecular network models via computational algebra. *BMC Systems Biology*, 10(1):94, 2016.

36. Steven A Narod and William D Foulkes. BRCA1 and BRCA2: 1994 and beyond. *Nature Reviews — CANCER*, 4(9):665–676, 2004.

37. Diana Nguyen, Maria Zajac-Kaye, Larry Rubinstein, Donna Voeller, Joseph E Tomaszewski, Shivaani Kummar, Alice P Chen, Yves Pommier, James H Doroshow, and Sherry X Yang. Poly (ADP-ribose) polymerase inhibition enhances p53-dependent and-independent DNA damage responses induced by DNA damaging agent. *Cell Cycle*, 10(23):4074–4082, 2011.

38. Charles Sanders Peirce. On the natural classification of arguments. In *Proceedings of the American Academy of Arts and Sciences*, volume 7, pages 261–287, 1867.

39. Livia Perfetto, Leonardo Briganti, Alberto Calderone, Andrea Cerquone Perpetuini, Marta Iannuccelli, Francesca Langone, Luana Licata, Milica Marinkovic, Anna Mattioni, Theodora Pavlidou, Daniele Peluso, Lucia Lisa Petrilli, Stefano Pirro, Daniela Posca, Elena Santonico, Alessandra Silvestri, Filomena Spada, Luisa Castagnoli, and Gianni Cesareni. SIGNOR: A database of causal relationships between biological entities. *Nucleic Acids Research*, 44(D1):D548–D554, 2016.

40. Patrick C Phillips. Epistasis - the essential role of gene interactions in the structure and evolution of genetic systems. *Nature Reviews Genetics*, 9(11):855–867, 2008.

41. C. Pizzuti. Computing prime implicants by integer programming. In *Proceedings Eighth IEEE International Conference on Tools with Artificial Intelligence*, pages 332–336. IEEE Comput. Soc. Press, 1996.

42. WV Quine. On cores and prime implicants of truth functions. *The American Mathematical Monthly*, 66(9):755–760, 1959.

43. Maria Spiliotaki, Dimitris Mavroudis, Kyriaki Kapranou, Harris Markomanolaki, Galatea Kallergi, Filippos Koinis, Kostas Kalbakis, Vassilis Georgoulias, and Sofia Agelaki. Evaluation of proliferation and apoptosis markers in circulating tumor cells of women with early breast cancer who are candidates for tumor dormancy. *Breast Cancer Research*, 16(6):485, 2014.

44. V Stambolic, D MacPherson, D Sas, Y Lin, B Snow, Y Jang, S Benchimol, and TW Mak. Regulation of PTEN transcription by p53. *Molecular cell*, 8(2):317–325, 2001.

45. Kyle Strimbu and Jorge a Tavel. What are Biomarkers? *Curr Opin HIV AIDS*, 5(6):463–466, 2011.

46. M. Strnisková, M. Barančík, and T. Ravingerová. MItogen-activated protein kinases and their role in regulation of cellular processes. *General Physiology and Biophysics*, 21(3):231–255, 2002.

47. Miyashita Toshiyuki and John C Reed. Tumor suppressor p53 is a direct transcriptional activator of the human bax gene. *Cell*, 80(2):293 – 299, 1995.

48. Fuminori Tsuruta, Norihisa Masuyama, and Yukiko Gotoh. The phosphatidylinositol 3-kinase (PI3K)-Akt pathway suppresses Bax translocation to mitochondria. *Journal of Biological Chemistry*, 277(16):14040–14047, 2002.

49. Marc Vidal. A unifying view of 21st century systems biology. *FEBS Letters*, 583(24):3891–3894, 2009.

50. Marc Vidal, Michael E Cusick, and Albert-László Barabási. Interactome networks and human disease. *Cell*, 144(6):986–998, 2011.

51. Bert Vogelstein, Nickolas Papadopoulos, Victor E Velculescu, Shibin Zhou, Luis A Diaz, and Kenneth W Kinzler. Cancer genome landscapes. *science*, 339(6127):1546–1558, 2013.

52. Silvia Von der Heyde, Christian Bender, Frauke Henjes, Johanna Sonntag, Ulrike Korf, and Tim Beissbarth. Boolean ErbB network reconstructions and perturbation simulations reveal individual drug response in different breast cancer cell lines. *BMC systems biology*, 8(1):75, 2014.

53. Xiaoyue Wang, Audrey Q Fu, Megan E McNerney, and Kevin P White. Widespread genetic epistasis among cancer genes. *Nature communications*, 5, 2014.

54. Jorge GT Zanudo and Réka Albert. Cell fate reprogramming by control of intracellular network dynamics. *PLoS Comput Biol*, 11(4):e1004193, 2015.

55. H Zhang, K Somasundaram, Y Peng, H Tian, H Zhang, D Bi, B L Weber, and W S El-Deiry. BRCA1 physically associates with p53 and stimulates its transcriptional activity. *Oncogene*, 16(13):1713–1721, 1998.

56. Wei Zhang and Hui Tu Liu. MAPK signal pathways in the regulation of cell proliferation in mammalian cells. *Cell research*, 12(1):9–18, 2002.

57. Quan Zhong, Nicolas Simonis, Qian-Ru Li, Benoit Charloteaux, Fabien Heuze, Niels Klitgord, Stanley Tam, Haiyuan Yu, Kavitha Venkatesan, Danny Mou, Venus Swearingen, Muhammed a Yildirim, Han Yan, Amélie Dricot, David Szeto, Chenwei Lin, Tong Hao, Changyu Fan, Stuart Milstein, Denis Dupuy, Robert Brasseur, David E Hill, Michael E Cusick, and Marc Vidal. Edgetic perturbation models of human inherited disorders. *Molecular systems biology*, 5(321):321, 2009.