

Examen écrit 3 heures
Tout document autorisé
Tout matériel électronique interdit
Written examination 3 hours
All printed documents allowed
Any electronic device prohibited

3 Mars 2025
March 3rd 2025

Abstract

La modélisation n'est pas un procédé rectiligne. Le modéleur fait souvent face à des choix. Il est alors important qu'il puisse retarder chaque décision jusqu'au moment qui lui semblera le plus opportun. Pour l'assister dans cette tâche, nous proposons dans ce problème de concevoir une analyse statique pour des modèles paramétrés. Dans ce cadre, l'ensemble des règles du modèle pourra être ajusté selon la valeur de paramètres booléens. L'analyse statique établira alors une relation entre les motifs et les conditions sur la valeur des paramètres booléens pour que ces motifs soient accessibles.

The modelling process does not usually follow a straight line. The modeller often faces choices. It is important for him to delay his decision until the right time. To assist him in this task, we propose, in this assignment, the design of a static analysis for parametric models. In this setting, the rule set of a model can be tuned thanks to the value of some Boolean parameters. The goal of the static analysis is then to establish some relationships among the patterns and the conditions over the Boolean parameters so that these patterns are reachable.

1 Cas d'étude / Case study

1.1 Signature du modèle / Model signature

Nous considérons un modèle constitué de trois types de protéine, E , R et T . Les instances de la protéine R , activées par des instances de la protéine E acquiert la capacité de se lier à des instances de la protéine T . Il sera supposé que chaque instance de la protéine R peut être liée à quatre instances de la protéine E , par des sites d'interactions a , b , c et d , et à une instance de la protéine T , par le site d'interaction e .

We consider a model with three kinds of protein, E , R , and T . The instances of the protein R , when activated by some instances of the protein E get the capability to bind to the instances of the protein T . More precisely, we assume that each instance of the protein R may be bound to four instances of the protein E on four interaction sites a , b , c , and d , and to one instance of the protein T on the interaction site e .

Un exemple d'espèce biochimique, pour ce modèle, est donnée en Fig. 1. Celle-ci est constituée d'une instance de la protéine R , liée à exactement trois instances de la protéine E sur ces sites d'interaction a , b et d et à une instance de la protéine T .

An example of a biochemical species, for this model, is given in Fig. 1. This one is made of one instance of the protein R . This instance is bound to exactly three instances of the protein E on its interaction sites a , b , and d , and to one instance of the protein T .

1.2 Première version du modèle / First variant of the model

Dans un premier temps, nous supposons que les liaisons entre les instances de la protéine R et les instances de la protéine E se font sans conditions. En revanche, nous supposons qu'une instance de la protéine R ne peut se lier à une instance de la protéine T que si ses trois sites d'interaction a , b et d sont liés et son site

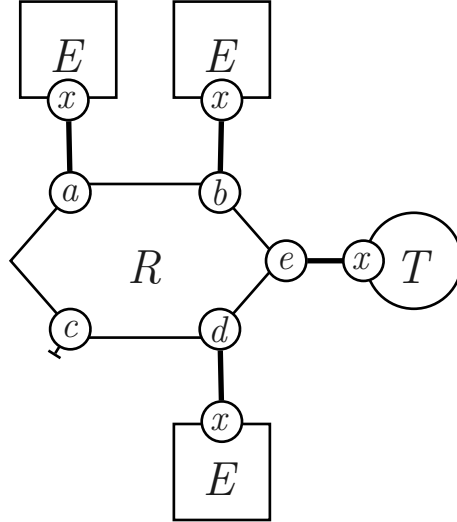


Figure 1: Un exemple d'espèce biochimique / an example of biochemical species.

d'interaction c est libre.

Firstly, we assume that the instances of the protein R and the ones of the protein E may bind unconditionally. Nevertheless, we assume that an instance of the protein R may bind to an instance of the protein T only if its three interaction sites a , b , and d are bound and its interaction site c is free.

La règle qui permet de lier les instances de la protéine R à celles de la protéine T est donnée en Fig. 2. En particulier, cette liaison n'est possible que si les sites d'interaction a , b et d de l'instance de la protéine R sont liées à des instances de la protéine E et le site d'interaction c est libre.

The rule given in Fig. 2 binds an instance of the protein R to an instance of the protein T , provided that the three interaction sites a , b , and d of the instance of the protein R are bound to some instances of the protein E , and its interaction site c is free.

Question 1 (*) *Écrire en Kappa, les quatre autres règles de liaisons de ce modèle.*

Write in Kappa the four other rules of the model.

Pour simplifier, et dans la totalité de ce problème, seules les règles de liaison seront considérées (on ne représentera donc pas de règles de dissociation).

To ease the things, and throughout this assignment, we will consider only binding rules (we ignore potential unbinding rules).

Question 2 (*) *Partant d'un état initial avec quatre instances de la protéine E , une instance de la protéine R et une instance de la protéine T , existe-t-il une trace d'exécution dans laquelle l'instance de la protéine R et l'instance de la protéine T se lient ensemble ? Donner en une preuve.*

Starting from an initial state made of four instances of the protein E , one instance of the protein R , and one instance of the protein T , does there exist an execution trace in which the instance of the protein R and the instance of the protein T bind to each other? Give a proof of your answer.

1.3 Deuxième version du modèle / Second variant of the model

Le modéleur trouve des hypothèses supplémentaires dans deux articles.

Reading through the literature, the modeller discovers two additional assumptions.

- Selon le premier, les sites d'interaction a et b dans les instances de la protéine R sont en conflit (ainsi, dans une instance de la protéine R , lorsque le site d'interaction a (respectivement le site d'interaction

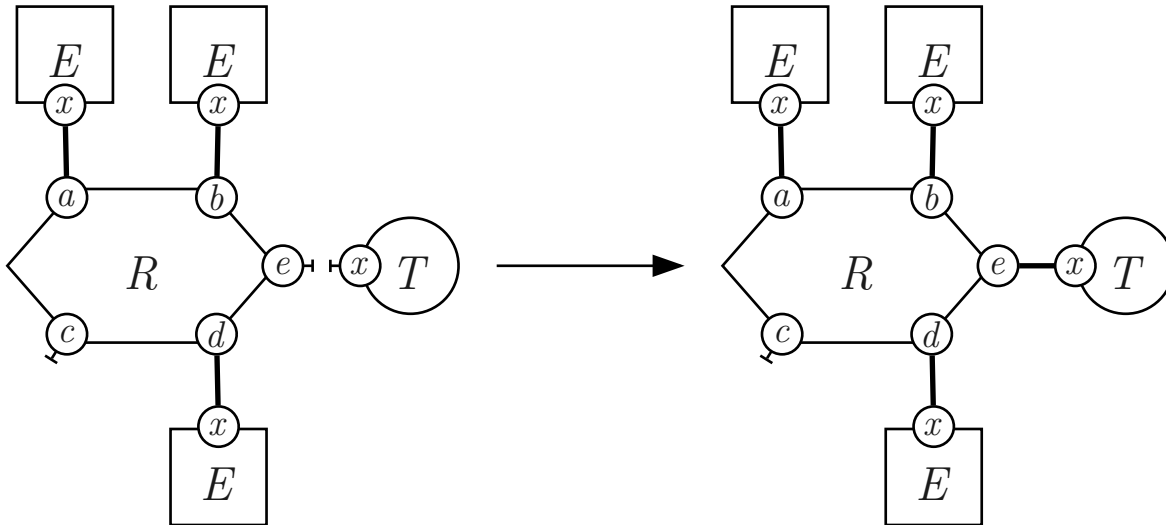


Figure 2: Un exemple de règle / an example of a rule.

b) est lié, il est impossible de lier le site d'interaction b (respectivement le site d'interaction a)).

According to the first one, the interaction sites a and b in the instances of the protein R are in conflict (this way, in an instance of the protein R , whenever the interaction site a (respectively the interaction site b) is bound, the interaction site b (respectively the interaction site a) cannot get bound).

- Selon le second, la liaison des sites d'interaction c et d dans les instances de la protéine R se font séquentiellement (ainsi, le site d'interaction d ne peut se lier que si le site d'interaction c l'est déjà).
According to the second one, the bindings of the interaction sites c and d in the instances of the protein R are established sequentially (this way, the interaction site d can get bound only if the interaction site c is bound already).

Question 3 (*) Écrire en Kappa, les règles de cette nouvelle version du modèle.

Write in Kappa the rules of this new variant of the model.

Question 4 ()** Partant du même état initial que pour la question 2, existe-t-il une trace d'exécution dans laquelle l'instance de la protéine R et l'instance de la protéine T se lient ensemble ? Donner une preuve de votre réponse.

Starting from the same initial state as in question 2, does it exist an execution trace in which an instance of the protein R and an instance of the protein T get bound to each other? Give a proof of your answer.

2 Quelques directives / Some directives

Le modéleur n'est pas sûr de pouvoir croire en ces hypothèses. Il préfère considérer non pas un modèle, mais quatre modèles selon que chacune de ses deux hypothèses soit vérifiée ou non. Cependant, plutôt que d'avoir à tenir à jour ces quatre modèles, il préfère un seul modèle paramétré par deux valeurs booléennes, une par hypothèse. Dans ce cadre, chaque règle sera préfixée par une expression booléenne sur la valeur des paramètres, qui indiquera pour quelles conditions sur la valeur des paramètres booléens, cette règle fait partie du modèle. Par défaut, les règles sont préfixées par la tautologie 'T'.

The modeller is not sure whether he should believe in those assumptions. He would rather keep four models (according to which assumptions are kept), rather than one. Yet, instead of having to maintain these four models, he would prefer a parametric model, with two Boolean parameters (one per assumption). In that

context, each rule will be prefixed with a Boolean expression over the value of the Boolean parameters. This expression indicates under which conditions over the value of the Boolean parameters this rule belongs to the model. By default, the rules are prefixed by the tautology '.T.'

Le but de cette section est de proposer des directives pour définir que certains sites d'interaction sont en conflit ou doivent se lier séquentiellement, puis une étape de compilation qui va poser un paramètre booléen frais par directive avant de dupliquer et raffiner les règles concernées selon la valeur effective de ce paramètre.

The goal of this section is to provide some syntactic directives to specify that a pair of interaction sites may be in conflict, or shall be sequentially bound, and then a compilation pass that will set a fresh Boolean parameter per directive, before duplicating and refining the rules that are concerned with the corresponding assumption according to the potential values of this parameter.

2.1 Conflits / Conflicts

Question 5 (*) Proposer une syntaxe pour une directive permettant d'indiquer que le modeleur hésite à supposer que deux sites d'interaction donnés dans les instances d'une protéine donnée sont en conflit, tout en posant un paramètre booléen (supposé frais) pour indiquer si cette hypothèse est considérée ou non dans la version correspondante du modèle.

Propose a syntax for a directive that would specify that the modeler wonder whether two given interaction sites in the instances of a given protein are in conflict, while declaring a Boolean parameter (assumed to be fresh) to indicate whether this assumption is considered in the corresponding variant of the model.

Question 6 ()** Proposer une procédure de compilation qui prends en argument, un paramètre Booléen P , un type de protéine A , deux sites d'interaction différents x et y de ce type de protéine, et une règle r préfixée de la précondition $pred$, et qui raffine la règle r et sa précondition $pred$ pour tenir compte du fait que lorsque la condition booléenne P prends la valeur '.T.', l'application de la règle r ne doit pas amener à ce qu'une instance de la protéine A ait ses deux sites d'interaction x et y tous les deux liés.

Cette procédure pourra éventuellement dupliquer et raffiner la règle pour les deux valeurs potentielles du paramètre P .

Suggest a compilation procedure which would take as arguments: a Boolean parameter P , a protein kind A , two different interaction sites x and y of this protein kind, and a rule with a precondition $pred$, and that refines the rule r and its precondition $pred$ so as to take into account that when the Boolean condition is satisfied, the application of the rule r shall not lead to a state where both interaction sites x and y in an instance of the protein A are both bound.

This procedure may duplicate and refine the rule r depending on whether the Boolean condition P is satisfied, or not.

2.2 Liaisons séquentielle / Sequential bindings

Question 7 (*) Proposer une syntaxe pour une directive permettant d'indiquer que le modeleur hésite à supposer que deux sites d'interaction donnés se lient toujours dans un ordre donné dans chaque instance d'une protéine donnée, tout en posant un paramètre booléen (supposé frais) pour indiquer si cette hypothèse est considérée ou non dans la version correspondante du modèle.

Propose a syntax for a directive that would specify that the modeler wonder whether two given interaction sites always bind in a given order in the instances of a given protein, while declaring a Boolean parameter (assumed to be fresh) to indicate whether this assumption is considered in the corresponding variant of the model.

Question 8 ()** Proposer une procédure de compilation qui prends en argument, un paramètre Booléen P , un type de protéine A , deux sites d'interaction différents x et y de ce type de protéine, et une règle r préfixée de la précondition $pred$, et qui raffine la règle r et sa précondition $pred$ pour tenir compte du fait que lorsque la condition booléenne P prends la valeur '.T.', l'application de la règle r ne doit pas amener à ce qu'une instance de la protéine A ait son site d'interaction y lié, alors que son site d'interaction x est libre.

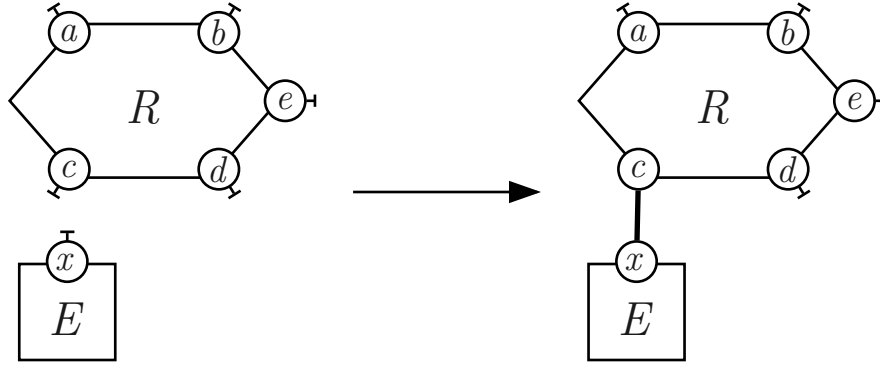


Figure 3: Un second exemple de réaction / A second example of reaction.

Cette procédure pourra éventuellement dupliquer et raffiner la règle pour les deux valeurs potentielles du paramètre P .

Suggest a compilation procedure which would take as arguments: a Boolean parameter P , a protein kind A , two different interaction sites x and y of this protein kind, and a rule with a precondition $pred$, and that refines the rule r and its precondition $pred$ so as to take into account that when the Boolean condition is satisfied, the application of the rule r shall not lead to a state where an instance of the protein A has its site y bound with its interaction site x free.

This procedure may duplicate and refine the rule r depending on whether the Boolean condition P is satisfied, or not.

3 Analyse statique de modèles paramétrés / Static analysis of parametric models

3.1 Analyse des espèces biochimiques accessibles / Reachable biochemical species analysis

Le but de cette sous-section est de concevoir une analyse statique qui associe à chaque espèce biochimique potentielle une formule sur les paramètres booléens qui indiquent pour quelles hypothèses il est possible de former cette espèce.

The objective of this subsection is the design of a static analysis that maps each potential biochemical species to a Boolean formula over the Boolean parameters indicating under which assumptions this biochemical species may potentially be formed.

Nous rappelons qu'une réaction biochimique est une règle d'interaction dans laquelle les composants biochimiques sont entièrement spécifiés. Son membre gauche (les réactifs) est donc un multi-ensemble d'espèces biochimiques, ainsi que son membre droit (les produits).

As a reminder, a biochemical reaction is an interaction rule where all the biochemical components are fully specified. This way, its left hand (the reactants) is a multi-set of biochemical species, and the right-hand side (the products) as well.

La règle d'interaction donnée en Fig. 2 est une réaction biochimique. Un autre exemple de réaction biochimique est donnée en Fig. 3. Cette dernière est un raffinement de la règle d'interaction qui lie le site d'interaction c d'une instance de la protéine R , au cas où tous les autres sites d'interactions de cette instance de protéine sont également libres.

The interaction rule, given in Fig. 2 is a biochemical reaction. Another example of biochemical reaction is given in Fig. 3. This one is a refinement of the interaction rule that binds the interaction site c in an instance of the protein R to the specific case where all the other interaction sites of this instance are free as

well.

Question 9 ()** *Étant donné un ensemble initial d'espèces biochimiques et un ensemble de réactions biochimiques préfixées par des préconditions portant sur la valeur des paramètres booléens, proposer un algorithme itératif qui associe à chaque espèce biochimique E une formule booléenne ϕ_E , qui indique pour quelles valeurs des paramètres, l'espèce biochimique E peut potentiellement se former.*

Given an initial set of biochemical species and a set of biochemical reactions, each one prefixed with a precondition over the Boolean parameters, suggest an iterative algorithm which maps every biochemical species S to a Boolean formula ϕ_S indicating under which conditions over the values of the Boolean parameters, the biochemical species S may potentially be formed.

On supposera dans cette sous-section que l'ensemble des espèces biochimiques est fini.

In this subsection, the set of biochemical species will be assumed to be finite.

Question 10 ()** *Appliquer l'analyse proposée à la question précédente au cas d'étude.*
Apply the analysis of the previous question on the case study.

3.2 Analyse des motifs accessibles / Reachable pattern analysis

On souhaite dans cette sous-section raisonner sur des motifs d'intérêt plutôt que sur des espèces biochimiques entièrement spécifiées. On pose donc \mathcal{M} un ensemble de motifs d'intérêt. Pour notre cas d'étude, nous prendrons celui qui est décrit en Fig. 4.

We want, in this subsection, to reason on patterns of interest, rather than on fully specified biochemical species. This way, we consider given a set \mathcal{P} of patterns of interest. In our case study, we take the set of patterns that is depicted in Fig. 4.

Nous construisons une analyse statique basée sur des règles d'inférence préfixées par des préconditions Booléennes. Deux exemples de règles d'inférences sont données en Fig. 5. Chaque règle comprend une prémisse (au dessus de la barre horizontale), sous la forme d'un sous-ensemble de l'ensemble des motifs d'intérêt (une partie de l'ensemble \mathcal{M} donc) et une conclusion (sous la barre horizontale), sous la forme d'un motif d'intérêt (un élément de l'ensemble \mathcal{M} donc). La précondition (sur le côté gauche) précise pour quelles valeurs des paramètres booléens la règle d'inférence peut être utilisée.

We design a static analysis based on some inference rules, prefixed by some Boolean preconditions. Two examples of inference rules are given in Fig. 5. Each rule is made of a premise (above the horizontal line), as a subset of the set of the patterns of interest (a part of the set \mathcal{P} thus) and a conclusion (bellow the horizontal line) as a pattern of interest (an element of the set \mathcal{P} thus). The precondition (on the left side) specifies for which values of the Boolean parameters the rule is applicable.

1. La règle d'inférence donnée en Fig. 5(a) stipule que s'il est possible de construire le motif formé d'une instance de la protéine R avec le site d'interaction c libre et le site d'interaction d lié, il est alors possible de produire le motif formé d'une instance de la protéine R avec les sites d'interaction c et d tous deux liés. Cette règle d'inférence découle de la règle de liaison du site d'interaction c d'une instance de la protéine R dans le cas spécifique où le site d'interaction d de cette instance de protéine est déjà lié. Comme cette règle d'interaction est présente dans tous les variants du modèle (i.e. quelque soit la valeur des paramètres booléens), cette règle d'inférence est préfixée de la tautologie $.T..$

The inference rule that is given in Fig. 5(a) states that, if it is possible to build the pattern made of one instance of the protein R with the interaction site c free and the interaction site d bound, then it is also possible to build the pattern made of one instance of the protein R with both interaction sites c and d bound. This rule comes from the rule to bind the interaction site c in the instances of the protein R , in the specific case where the interaction site d of this protein instance is bound already. Since this rule is valid in every variant of the model (i.e. whatever the value of the Boolean paramters is), this inference rule is prefixed by the tautology $.T..$

2. La règle d'inférence donnée en Fig. 5(b) stipule que s'il est possible de construire les trois motifs formés de (a) d'une instance de la protéine R avec les deux sites d'interaction a et b liés (b) d'une instance

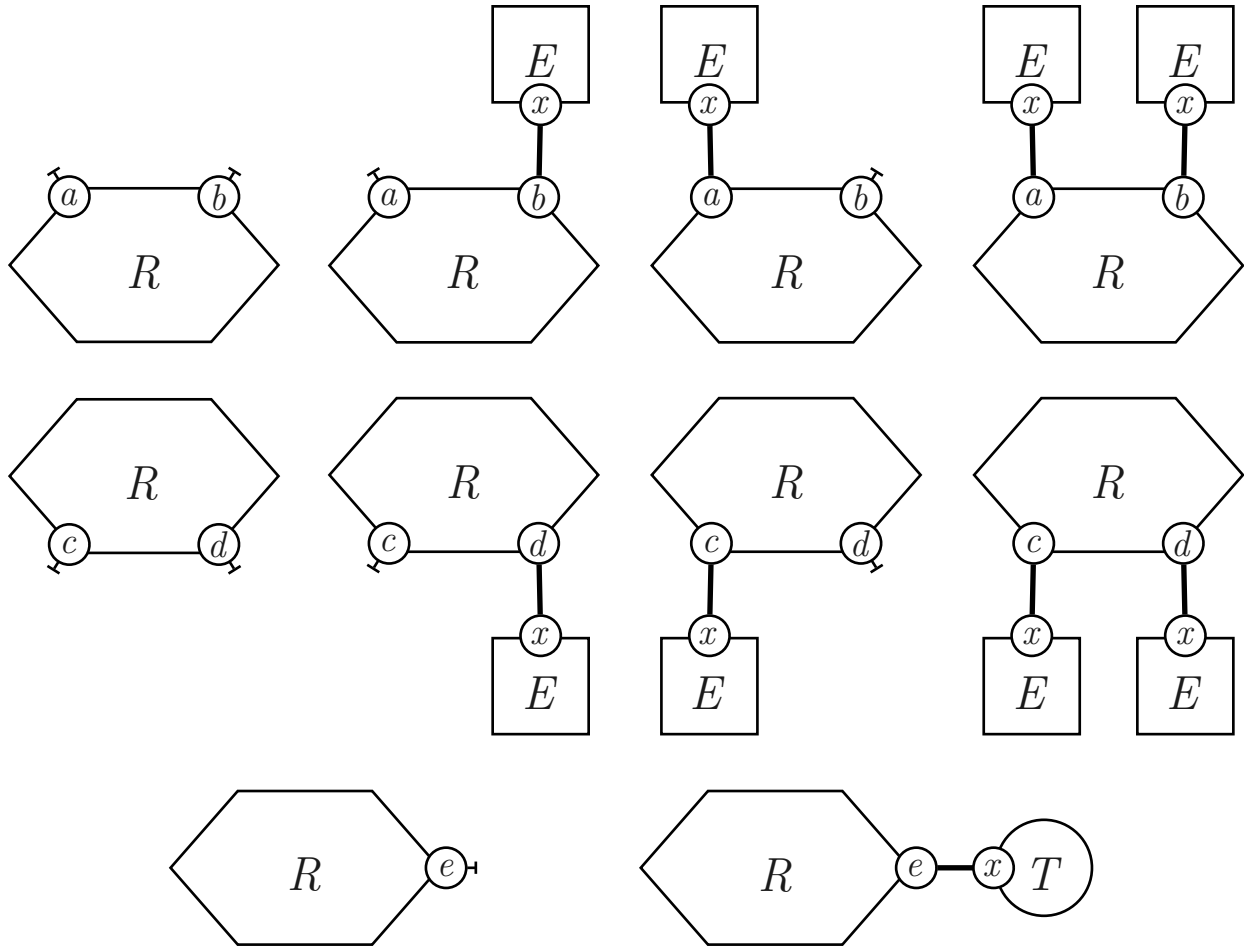


Figure 4: Motifs d'intérêt / patterns of interest.

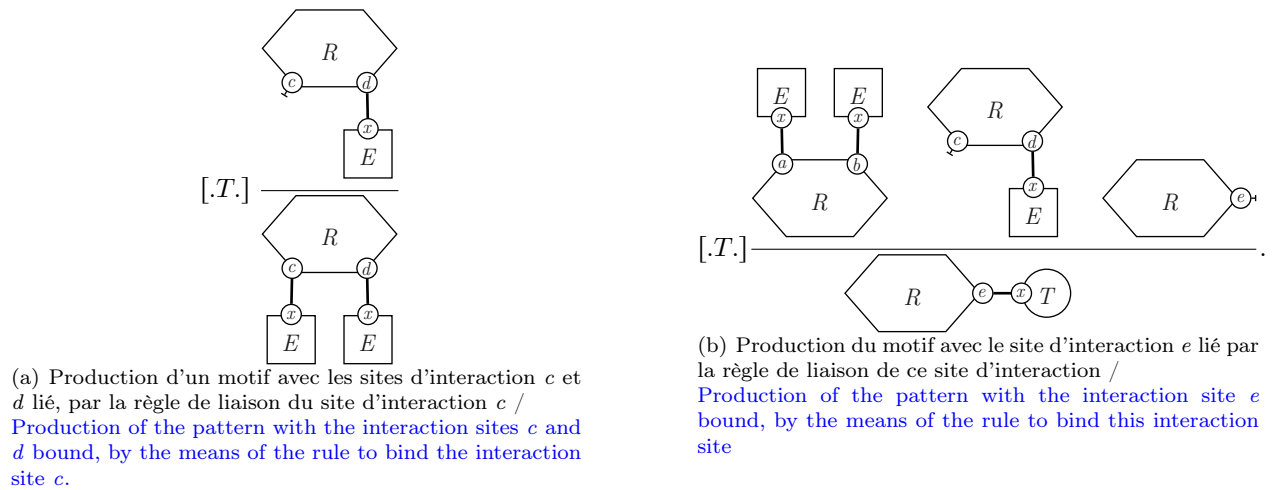


Figure 5: Deux exemples de règles d'inférence / two examples of inference rules.

de la protéine R avec le site d'interaction c libre et le site d'interaction d lié (c) d'une instance de la protéine R avec le site d'interaction e libre, il est alors possible de produire le motif formé d'une instance de la protéine R avec le site e lié. Cette règle d'inférence découle de la règle de liaison du site d'interaction e d'une instance de la protéine R . Comme cette règle d'interaction est présente dans tous les variants du modèle (i.e. quelque soit la valeur des paramètres booléens), cette règle d'inférence est préfixée de la tautologie $.T.$.

The inference rule that is given in Fig. 5(b) states that if it is possible to build the three patterns made of (a) one instance of the protein R with its two interaction sites a and b bound (b) one instance of the protein R with its interaction site c free and its interaction site d bound (c) one instance of the protein R with the interaction site e free, then it is possible to build the pattern made of an instance of the protein R and the interaction site e bound. This inference rule comes from the binding rule for the interaction site e of the instances of the protein R . As this interaction rule is valid in every variant of the model (i.e. whatever the value of the Boolean parameters is), this inference rule is prefixed by the tautology $.T.$.

Question 11 ()** Donner un ensemble de règles d'inférence pour le cas d'étude. Cet ensemble de règles devra être tel que pour tout sous-ensemble $X \subseteq \mathcal{M}$ de motifs d'intérêt, pour tout état x qui ne contient pas de motifs dans l'ensemble $\mathcal{M} \setminus X$, et toute étape de calcul passant de l'état x à un état x' , alors s'il existe un motif d'intérêt M dans l'ensemble $\mathcal{M} \setminus X$ qui apparaît dans l'état x' , alors il existe une règle d'inférence dont la prémisse est un sous-ensemble de l'ensemble X et la conclusion est le motif M .

Give a set of inference rules for the case study. This set of rules should be such that for every subset $X \subseteq \mathcal{P}$ of patterns of interest, for every state x that does not contain any pattern in the set $\mathcal{P} \setminus X$, and for every computation step from the state x to a state x' , then, if there exists a pattern of interest P in the set $\mathcal{P} \setminus X$ that occurs in the state x' , then there exists an inference rule the premise of which is a subset of the set X and the conclusion is the pattern P .

Question 12 ()** Étant donné un sous-ensemble initial de motifs d'intérêt accessibles et un ensemble de règles d'inférence préfixées par des préconditions sur la valeur des paramètres booléens, proposer un algorithme itératif qui associe à chaque motif d'intérêt M une formule booléenne ϕ_M , qui indique pour quelles valeurs des paramètres, il est impossible de prouver, à ce niveau d'abstraction, que le motif M est inaccessible.

Given an initial subset of reachable patterns of interest and a set of inference rules prefixed by some Boolean preconditions over the Boolean parameters of the parametric model, design an iterative algorithm that maps every pattern of interest P to a Boolean formula ϕ_M , that specifies for which values of the Boolean parameters, it is impossible to prove, at our level of abstraction, that the pattern P is unreachable.

Question 13 ()** Appliquer l'analyse proposée à la question précédente au cas d'étude paramétré.
Apply the static analysis proposed in the previous question to the (parametric) case study.

4 Détails d'implantation / Implementation issues

Question 14 ()** Proposer des structures de données pour rendre l'analyse plus efficace. Il est possible d'imposer des contraintes à l'ensemble des motifs d'intérêt considéré.

Suggest some data-structures to improve the efficiency of the static analysis.

It is allowed to set up some constraints about the structure on the set of patterns of interest.