Algorithmic Debugging of Missing Answers in Constraint Functional-Logic Programming

Rafael Caballero, Mario Rodríguez Artalejo, and
Rafael del Vado Vírseda *

TECHNICAL REPORT SIC - 2/08

Dpto. de Sistemas Informáticos y Computación,
Universidad Complutense de Madrid
{rafa,mario,rdelvado}@sip.ucm.es

Abstract. We present a declarative method for diagnosing *missing computed answers* in $CFLP(\mathcal{D})$, a generic scheme for lazy *Constraint Functional-Logic Programming* which can be instantiated by any constraint domain $\mathcal{D}$ given as parameter. As far as we know, declarative diagnosis of missing answers in such an expressive framework has not been tackled before. Our approach combines and extends previous work done separately for constraint logic programming and lazy functional programming languages. Diagnosis can be started whenever a user finds that the set of computed answers for a given goal with finite search space misses some expected solution w.r.t. an *intended interpretation* of the program, that provides a declarative descriptions of its expected behavior. Diagnosis proceeds by exploring a *proof tree*, that provides a declarative view of the *answer-collection* process performed by the computation, and it ends up with the detection of some function definition in the program that is incomplete w.r.t. the intended interpretation. We can prove the *logical correctness* of the diagnosis method under the assumption that the recollection of computed answers performed by the goal solving system can be represented as a proof tree. We argue the plausibility of this assumption, and we describe the prototype of a tool which implements the diagnosis method.

1 Introduction

Debuggers are a practical need for helping programmers to understand why their programs do not work as intended. Declarative programming paradigms involving complex operational details, such as constraint solving and lazy evaluation, do not fit well to traditional debugging techniques relying on the inspection of low-level computation traces. For this reason, the design of usable debugging tools becomes a difficult task. As a solution to this problem, and following a seminal idea by Shapiro [28], *declarative diagnosis* (a.k.a. *declarative debugging*

* The authors have been partially supported by the Spanish National Projects MERIT-FORMS (TIN2005-09027-C03-03) and PROMESAS-CAM (S-0505/TIC/0407).

or *algorithmic debugging*) proposes to use *Computation Trees* (shortly, CTs) in place of traces. CTs are built *a posteriori* to represent the structure of a computation whose top-level outcome is regarded as a symptom of the unexpected behavior by the user, with results attached to their nodes representing the computation of some observable result, and such that the result at any internal node follows from the results at the children nodes, using a program fragment also attached to the node. Declarative diagnosis explores a CT looking for a so-called *buggy node* which computes an unexpected result from children whose results are all expected. Each buggy node points to a program fragment responsible for the unexpected behavior. The search for a buggy node can be implemented with the help of an external *oracle* (usually the user with some semiautomatic support) who has a reliable declarative knowledge of the expected program semantics, the so-called *intended interpretation*.

The generic description of declarative diagnosis in the previous paragraph follows [22]. Declarative diagnosis was first proposed in the field of *Logic Programming* (LP) [28, 14, 18], and it has been successfully extended to other declarative programming paradigms, including (lazy) *Functional Programming* (FP) [25, 24, 27, 26], *Constraint Logic Programming* (CLP) [30, 15] and *Functional Logic Programming* (FLP) [23, 6, 7]. The nature of unexpected results differs according to the programming paradigm. Unexpected results in FP are mainly *incorrect values*, while in CLP and FLP an unexpected result can be either a single computed answer regarded as *incorrect*, or a set of computed answers (for one and the same goal with a finite search space) regarded as *incomplete*. These two possibilities give rise to the declarative diagnosis of *wrong* and *missing* computed answers, respectively. The case of unexpected *finite failure* of a goal is a particular symptom of missing answers with special relevance. However, diagnosis methods must consider the more general case, since finite failure of a goal is often caused by non-failing subgoals that do not compute all the expected answers.

In contrast to alternative approaches to error diagnosis based on *abstract interpretation* techniques [17], declarative diagnosis often involves complex queries to the user. This problem has been tackled by means of various techniques, such as user-given partial specifications of the program's semantics [1, 7], safe inference of information from answers previously given by the user [6], or CTs tailored to the needs of a particular debugging problem over a particular computation domain [15]. Another practical problem with declarative diagnosis is that the size of CTs can cause excessive overhead in the case of computations that demand a big amount of computer storage. As a remedy, techniques for piecemeal construction of CTs have been considered; see [26] for a recent proposal in the FP field.

In spite of the above mentioned difficulties, we are confident that declarative diagnosis methods can be useful for detecting programming bugs by observing computations whose demand of computer storage is modest. In this paper, we present a declarative method for diagnosing *missing computed answers* in $CFLP(\mathcal{D})$ [20], a generic scheme for lazy *Constraint Functional-Logic Programming* which can be instantiated by any constraint domain $\mathcal{D}$ given as parameter,

and supports a powerful combination of functional and constraint logic programming over $\mathcal{D}$. Sound and complete goal solving procedures for the $CFLP(\mathcal{D})$ scheme have been obtained [19, 11, 12]. Moreover, useful instances of this scheme have been implemented in the $\mathcal{TOY}$ system [21] and tested in practical applications [13].

The rest of the paper is organized as follows: Section 2 motivates our approach and presents a debugging example, intended to illustrate the main features of our diagnosis method. Section 3 presents the abbreviated proof trees used as CTs in our method, as well as the results ensuring the logical correctness of the diagnosis. Section 4 presents a prototype debugger under development, and Section 5 concludes and gives an overview of planned future work. Full proofs of the main results given in Section 3 are available in [10].

2 Motivation

While methods and tools for the declarative diagnosis of *wrong answers* are known for FLP [23, 6, 7] and $CFLP$ [4, 8] languages, we are not aware of any research concerning the declarative diagnosis of *missing answers* in $CFLP$ languages, except our poster presentation [9]. However, missing answers are a common problem which can arise even in the absence of wrong answers.

We are interested in the declarative diagnosis of missing answers in $CFLP(\mathcal{D})$ [20], a very expressive generic scheme for Functional and Constraint Logic Programming over a constraint domain $\mathcal{D}$ given as parameter. Each constraint domain provides basic values and primitive operations for building domain specific constraints to be used in programs and goals. Useful constraint domains include the Herbrand domain $\mathcal{H}$ for equality (`==`) and disequality (`/=`) constraints over constructed data values; the domain $\mathcal{R}$ for arithmetic constraints over real numbers; and the domain $\mathcal{FD}$ for finite domain constraints over integer values.

The $CFLP(\mathcal{D})$ scheme supports programming with lazy functions that may be non-deterministic and/or higher-order. *Programs* $\mathcal{P}$ include *program rules* of the form $f\, t_1 \ldots t_n \rightarrow r \Leftarrow \Delta$, abbreviated as $f\, \bar{t}_n \rightarrow r \Leftarrow \Delta$, with Δ omitted if empty. Such a rule specifies that f when acting over parameters matching the patterns $\bar{t}_n$ at the left hand side, will return the values resulting from the right hand side expression r, provided that the constraints in Δ can be satisfied. *Goals* G for a given program have the general form $\exists \overline{U}.\ (R \,\square\, S)$, where $\exists \overline{U}$ is an existentially quantified prefix of local variables, $R = (P \,\square\, \Delta)$ is the yet *unsolved part*, including *productions* $e \rightarrow s$ in P and *constraints* in Δ, and $S = (\Pi \,\square\, \sigma)$ is the *constraint store*, consisting of *primitive constraints* Π and an *idempotent substitution* σ. Productions $e \rightarrow s$ are solved by *lazy narrowing*, a combination of unification and lazy evaluation; the expression e must be narrowed to match the pattern s. *Initial goals* have neither productions nor local variables, and *solved goals* have the form $\exists \overline{U}.\ S$. Solved goals are also called *computed answers* and abbreviated as $\hat{S}$.

In this paper we focus mainly in $CFLP(\mathcal{D})$ programming as implemented in $\mathcal{TOY}$ [21]. The interested reader is referred to [20, 19, 11] for formal details on

the declarative and operational semantics of the $CFLP(\mathcal{D})$ scheme. The following small $CFLP(\mathcal{H})$-program $\mathcal{P}_{fD}$, written in $\mathcal{TOY}$ syntax, includes program rules for the non-deterministic functions `(//)` and `fDiff`, and the deterministic functions `gen` and `even`. Note the infix syntax used for `(//)`, as well as the use of the equality symbol `=` in place of the rewrite arrow `-->` for the program rules of those functions viewed as deterministic by the user. This is just meant as user given information, not checked by the $\mathcal{TOY}$ system, which treats all the program defined functions as possibly non-deterministic.

```
infixr 40 //              % non-deterministic choice operator

(//) :: A -> A -> A
X // _ --> X
_ // Y --> Y

fDiff :: [A] -> A
fDiff [X]      --> X
fDiff (X:Y:Zs) --> X // fDiff (Y:Zs) <== X /= Y
fDiff (X:Y:Zs) --> X                 <== X == Y

gen :: A -> A -> [A]        even :: int -> bool
gen X Y = X : Y : gen Y X   even N = true <== (mod N 2) == 0
```

Function `fDiff` is intended to return any element belonging to the longest prefix `Xs` of the list given as parameter such that `Xs` does not include two identical elements in consecutive positions. In general, there will be several such elements, and therefore `fDiff` is non-deterministic. Function `gen` is deterministic and returns a potentially infinite list of the form $[d_1, d_2, d_2, d_1, d_1, d_2, ...]$, where the elements d_1 and d_2 are the given parameters. Therefore, the lazy evaluation of `(fDiff (gen 1 2))` is expected to yield the two possible results `1` and `2` in alternative computations, and the initial goal G_{fD} : `even (fDiff (gen 1 2)) == true` for $\mathcal{P}_{fD}$ is expected to succeed, since `(fDiff (gen 1 2))` is expected to return the even number `2`. However, if the third program rule for function `fDiff` were missing in program $\mathcal{P}_{fD}$, the expression `(fDiff (gen 1 2))` would return only the numeric value `1`, and therefore the goal G_{fD} would fail unexpectedly. At this point, a diagnosis for missing answers could take place, looking for a *buggy node* in a suitable CT in order to detect some incomplete function definition (that of function `fDiff`, in this case) to be blamed for the missing answers.

We propose to use CTs whose nodes have attached so-called *answer collection assertions*, briefly *aca*s. The *aca* at the root node has the form $G_0 \Rightarrow \bigvee_{i \in I} \hat{S}_i$, where G_0 is the initial goal and $\bigvee_{i \in I} \hat{S}_i$ (written as the *failure symbol* $\blacklozenge$ if $I = \emptyset$) is the disjunction of computed answers observed by the user. This root *aca* asserts that the computed answers cover all the solutions of the initial goal, and will be regarded as a false statement in case that the user misses computed answers. For example, the root *aca* corresponding to the initial goal G_{fD} for program $\mathcal{P}_{fD}$ is `even (fDiff (gen 1 2)) == true` $\Rightarrow \blacklozenge$ stating that this goal

has (unexpectedly) failed. The *aca*s at internal nodes in our CTs have the form $f\bar{t}_n \rightarrow t \,\square\, S \Rightarrow \bigvee_{i\in I} \hat{S}_i$, asserting that the disjunction of computed answers $\bigvee_{i\in I} \hat{S}_i$ covers all the solutions for the intermediate goal $G' : f\bar{t}_n \rightarrow t \,\square\, S$. Note that G' asks for the solutions of the production $f\bar{t}_n \rightarrow t$ which satisfy the constraint store S. The *aca*s of this form correspond to the intermediate calls to program defined functions f needed for collecting all the answers computed for the initial goal G_0. Due to *lazy evaluation*, the parameters $\bar{t}_n$ and the result t will appear in the most evaluated form demanded by the topmost computation. When these values are functions, they are represented in terms of partial applications of top-level function names. This is satisfactory under the assumption that no local function definitions are allowed in programs, as it happens in $\mathcal{TOY}$.

We build our CTs as abbreviated *proof trees* w.r.t. a logically sound inference system for deriving *aca*s. For this reason, our CTs are such that the validity of the *aca* at each node follows from the validity of the *aca*s at their children, under the assumption that the function definition relating the parent node to the children nodes is complete w.r.t. the *intended interpretation* of the program. Any CT whose root *aca* is invalid must include at least one *buggy node* labeled with an invalid *aca* and whose children are all labeled with valid *aca*s. Each buggy node N is related to some particular function f whose program rules are responsible for the computation of the *aca* at N from the *aca*s at N's children. Therefore, the program rules for f can be diagnosed as incomplete. The search for a buggy node can be implemented with the help of an external *oracle* who has a reliable declarative knowledge of the valid *aca*s w.r.t. the intended program interpretation. Since the oracle is usually the programmer, she can even experiment with different choices of the intended interpretation in order to obtain different diagnosis of possibly incomplete functions.

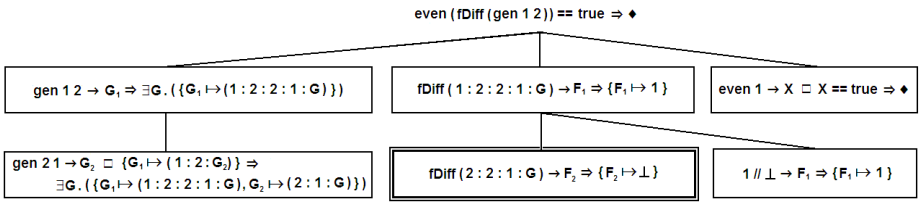


Fig. 1. CT for the declarative diagnosis of missing answers

A CT corresponding to the goal G_{fD} for program $\mathcal{P}_{fD}$ (with the third program rule for function `fDiff` omitted) is displayed in **Fig. 1**. More on its structure and construction will be explained in Section 3. In this case, the programmer will judge the root *aca* as *invalid* because she did not expect finite failure. Moreover, from her knowledge of the intended interpretation, she will decide to consider the *aca*s for the functions `gen`, `even` and `(//)` as valid. However, the *aca* `fDiff (2:2:1:G)` $\rightarrow F_2 \Rightarrow (F_2 \mapsto \bot)$ asserts that the *undefined value* $\bot$ is the only

possible result for the function call `fDiff (2:2:1:G)`, while the user expects also the result `2`. Therefore, the user will judge this *aca* as invalid. The node where it sits (enclosed within a double box in **Fig. 1**) has no children and thus becomes buggy, leading to the diagnosis of `fDiff` as incomplete. This particular incompleteness symptom could be mended by placing the third rule for `fDiff` within the program.

3 Declarative Diagnosis of Missing Answers

As explained in the previous sections, the declarative diagnosis method proposed in this paper relies on building CTs as *abbreviated proof trees* w.r.t. a logically sound inference system for deriving *aca*s. In this section, we present such an inference system, whose *Negative Proof Trees* represent the deduction of *aca*s from the *Negative Theory* $\mathcal{P}^-$ associated to a given $CFLP(\mathcal{D})$-program $\mathcal{P}$. We also present results ensuring the *logical correctness* of the declarative diagnosis method whose CTs are abbreviated representations of NPTs.

3.1 Standardized Programs and Negative Theories

Let $\mathcal{P}$ be a $CFLP(\mathcal{D})$-program. Its associated *Negative Theory* $\mathcal{P}^-$ is obtained in two steps. First, each program rule $f\ \bar{t}_n \to r \Leftarrow \Delta$ is replaced by a *standardized* form $f\ \overline{X}_n \to Y \Leftarrow \hat{R}$, where $\overline{X}_n, Y$ are new variables, $\hat{R} = \exists \overline{U}.\ R$ with $\overline{U} = var(R) \setminus \{\overline{X}_n, Y\}$, and the condition R is $X_1 \to t_1, \ldots, X_n \to t_n, \Delta, r \to Y$. Next, $\mathcal{P}^-$ is built by taking one *axiom* $(f)^-_{\mathcal{P}}$ of the form $\forall \overline{X}_n, Y.\ (f\ \overline{X}_n \to Y \Rightarrow (\bigvee_{i \in I} \hat{R}_i) \vee (\bot \to Y))$ for each function symbol f whose standardized program rules are $\{f\ \overline{X}_n \to Y \Leftarrow \hat{R}_i\}_{i \in I}$. By convention, we may use the notation D_f for the disjunction $(\bigvee_{i \in I} \hat{R}_i) \vee (\bot \to Y)$, and we may leave the universal quantification of the variables $\overline{X}_n, Y$ implicit. Intuitively, the axiom $(f)^-_{\mathcal{P}}$ says that any result computed for f must be obtained by means of some of the rules for f in the program. The last alternative $(\bot \to Y)$ within D_f says that Y is bound to the undefined result $\bot$ in case that no program rule for f succeeds to compute a more defined result. For example, let $\mathcal{P}_{fD}$ be the $CFLP(\mathcal{H})$-program given in Section 2, with the third program rule for `fDiff` omitted. Then $\mathcal{P}^-_{fD}$ includes (among others) the following axiom for the function symbol *fDiff*:

$$
\begin{aligned}
(\mathit{fDiff})^-_{\mathcal{P}_{fD}} : \forall\, L, F.\ (\mathit{fDiff}\ L \to F \Rightarrow\ & \\
& \exists X.\ (L \to [X] \wedge X \to F) \vee \\
& \exists X, Y, Zs.\ (L \to (X : Y : Zs) \wedge X \texttt{ /= } Y \wedge X \mathbin{//} \mathit{fDiff}\ (Y : Zs) \to F) \vee \\
& (\bot \to F))
\end{aligned}
$$

Interpretations $\mathcal{I}$ are formally defined in [20]. Each interpretation represents a certain behavior of the program defined functions. We write $\mathcal{I} \Vdash_{\mathcal{D}} f\bar{t}_n \to t$ to indicate that the statement $f\bar{t}_n \to t$ is *valid* in $\mathcal{I}$. Here, f is a program defined function, $\bar{t}_n$ stand for possibly partially evaluated arguments, and t stands for a possibly partially evaluated result. Knowing the valid assertions $\mathcal{I} \Vdash_{\mathcal{D}} f\bar{t}_n \to t$

suffices for defining the *solution set* $Sol_{\mathcal{I}}(G)$ whose elements are all the *valuations* (i.e., substitutions of domain values for variables) that satisfy the goal G w.r.t. $\mathcal{I}$. We will use similar notations for other solution sets in the rest of the paper, writing $Sol_{\mathcal{D}}$ instead of $Sol_{\mathcal{I}}$ whenever the solutions do not depend on the interpretation $\mathcal{I}$ of program defined functions. The following definition helps to understand the semantics of missing answers:

Definition 1 (Interpretation-Dependent Semantics). *Let $\mathcal{P}$ a $CFLP(\mathcal{D})$-program and $\mathcal{I}$ an interpretation over $\mathcal{D}$.*

1. *$\mathcal{I}$ is a* **model** *of $\mathcal{P}^-$ iff every axiom $(f)^-_{\mathcal{P}} : (f\ \overline{X}_n \to Y \Rightarrow D_f) \in \mathcal{P}^-$ satisfies $Sol_{\mathcal{I}}(f\ \overline{X}_n \to Y) \subseteq Sol_{\mathcal{I}}(D_f)$. When this inclusion holds, we say that $(f)^-_{\mathcal{P}}$ is* **valid** *in $\mathcal{I}$, or also that f's definition as given in $\mathcal{P}$ is* **complete** *w.r.t. $\mathcal{I}$.*

2. *The aca $G \Rightarrow \bigvee_{i\in I} \hat{S}_i$ is a* **logical consequence** *of $\mathcal{P}^-$ iff $Sol_{\mathcal{I}}(G) \subseteq \bigcup_{i\in I} Sol_{\mathcal{D}}(\hat{S}_i)$ for any model $\mathcal{I}$ of $\mathcal{P}^-$. When this happens, we also say that the disjunction of answers $\bigvee_{i\in I} \hat{S}_i$ is* **complete** *for G w.r.t. $\mathcal{P}$.*

3.2 Negative Proof Trees for Answer Collection Assertions

The declarative debugging of missing answers presupposes an *intended interpretation* of the program, starts with the observation of an *incompleteness symptom* and ends with an *incompleteness diagnosis*. A more precise definition of this *debugging scenario* is as follows:

Definition 2 (Debugging Scenario). *For any given $CFLP(\mathcal{D})$-program $\mathcal{P}$:*

1. *The* **intended interpretation** *is some interpretation $\mathcal{I}_{\mathcal{P}}$ over $\mathcal{D}$ which represents the behavior of the functions defined in $\mathcal{P}$ as expected by the programmer.*

2. *An* **incompleteness symptom** *occurs if the goal solving system computes finitely many solved goals $\{\hat{S}_i\}_{i\in I}$ as answers for an admissible initial goal G, and the programmer judges that $Sol_{\mathcal{I}_{\mathcal{P}}}(G) \not\subseteq \bigcup_{i\in I} Sol_{\mathcal{D}}(\hat{S}_i)$, meaning that the aca $G \Rightarrow \bigvee_{i\in I} \hat{S}_i$ is not valid in the intended interpretation $\mathcal{I}_{\mathcal{P}}$, so that some expected answers are missing.*

3. *An* **incompleteness diagnosis** *is given by pointing to some defined function symbol f such that the axiom $(f)^-_{\mathcal{P}}$ for f in $\mathcal{P}^-$ is not valid in $\mathcal{I}_{\mathcal{P}}$, which means $Sol_{\mathcal{I}_{\mathcal{P}}}(f\ \overline{X}_n \to Y) \not\subseteq Sol_{\mathcal{I}_{\mathcal{P}}}(D_f)$, showing that f's definition as given in $\mathcal{P}$ is incomplete w.r.t. $\mathcal{I}_{\mathcal{P}}$.*

Some concrete debugging scenarios have been discussed in Section 2 and [9]. Assume now that an incompleteness symptom has been observed by the programmer. Since the goal solving system has computed the disjunction of answers

SF Solved Form $\dfrac{}{R \sqcap S \Rightarrow D}$ if $Sol_{\mathcal{D}}(S) \subseteq Sol_{\mathcal{D}}(D)$.

CJ Conjunction

$$\frac{R_1 \sqcap S \Rightarrow \bigvee_{i\in I} \exists \overline{Z}_i.\ S_i \quad \ldots\ (\hat{R}_2\ \&\ \hat{S}_i) \Rightarrow \bigvee_{j\in J_i} \exists \overline{Z}_{ij}.\ S_{ij}\ \ldots\ (i \in I)}{(R_1 \wedge R_2) \sqcap S \Rightarrow \bigvee_{i\in I} \bigvee_{j\in J_i} \exists \overline{Z}_i, \overline{Z}_{ij}.\ S_{ij}}$$

if $\overline{Z}_i \notin var((R_1 \wedge R_2) \sqcap S)$, $\overline{Z}_{ij} \notin var((R_1 \wedge R_2) \sqcap S) \cup \overline{Z}_i$, for all $i \in I, j \in J_i$.

TS Trivial Statement $\dfrac{}{\varphi : G \Rightarrow D}$

if φ is a trivial *aca* s.t. $Sol(G) \subseteq Sol_{\mathcal{D}}(D)$.

DC DeComposition $\dfrac{\overline{e_m \to t_m} \sqcap S \Rightarrow D}{h\overline{e}_m \to h\overline{t}_m \sqcap S \Rightarrow D}$ if $h\overline{e}_m$ is not a pattern.

IM IMitation $\dfrac{\overline{e_m \to X_m} \sqcap (S \wedge h\overline{X}_m \to X) \Rightarrow \bigvee_{i\in I} \exists \overline{Z}_i.\ S_i}{h\overline{e}_m \to X \sqcap S \Rightarrow \bigvee_{i\in I} \exists \overline{X}_m, \overline{Z}_i.\ S_i}$

if $h\overline{e}_m$ is not a pattern, $X \in \mathcal{V}$, and $\overline{X}_m \notin var(h\overline{e}_m \to X \sqcap S)$.

$(\mathbf{AR})_p$ Argument Reduction for Primitive Functions

$$\frac{\overline{e_n \to X_n} \sqcap (S \wedge p\overline{X}_n \to!\ t) \Rightarrow \bigvee_{i\in I} \exists \overline{Z}_i.\ S_i}{p\overline{e}_n \to^{?} t \sqcap S \Rightarrow (S \wedge \bot \to t) \vee (\bigvee_{i\in I} \exists \overline{X}_n, \exists \overline{Z}_i.\ S_i)}$$

if $p \in PF^n, \overline{X}_n \notin var(p\overline{e}_n \to^{?} t \sqcap S)$, and $\to^{?} \equiv \to$ (*production*) $\cup \to!$ (*constraint*). For instance, *equality constraints* e_1 `==` e_2 (resp., *disequality constraints* e_1 `/=` e_2). are abbreviations of e_1 `==` $e_2 \to! \mathit{true}$ (resp., e_1 `==` $e_2 \to! \mathit{false}$).

$(\mathbf{AR})_f$ Argument Reduction for Defined Functions

$$\frac{(\overline{e_n \to X_n} \wedge f\overline{X}_n \to t) \sqcap S \Rightarrow \bigvee_{i\in I} \exists \overline{Z}_i.\ S_i}{f\overline{e}_n \to t \sqcap S \Rightarrow \bigvee_{i\in I} \exists \overline{X}_n, \overline{Z}_i.\ S_i}$$

if $f \in DF^n$, and $\overline{X}_n \notin var(f\overline{e}_n \to t \sqcap S)$.

$$\frac{(\overline{e_n \to X_n} \wedge f\overline{X}_n \to Y \wedge Y\overline{a}_k \to t) \sqcap S \Rightarrow \bigvee_{i\in I} \exists \overline{Z}_i.\ S_i}{f\overline{e}_n\overline{a}_k \to t \sqcap S \Rightarrow \bigvee_{i\in I} \exists \overline{X}_n, Y, \overline{Z}_i.\ S_i}$$

if $f \in DF^n$ $(k > 0)$, and $\overline{X}_n, Y \notin var(f\overline{e}_n\overline{a}_k \to t \sqcap S)$.

$(\mathbf{DF})_f$ Defined Function $\dfrac{\ldots\ R_i[\overline{X}_n \mapsto \overline{t}_n, Y \mapsto t] \sqcap S \Rightarrow D_i\ \ldots\ (i \in I)}{\boxed{f\overline{t}_n \to t \sqcap S \Rightarrow (S \wedge \bot \to t) \vee (\bigvee_{i\in I} D_i)}}$

if $f \in DF^n$, $\overline{X}_n, Y \notin var(f\overline{t}_n \to t \sqcap S)$, and $(f\overline{X}_n \to Y \Rightarrow \bigvee_{i\in I} \hat{R}_i) \in \mathcal{P}^-$.

Fig. 2. The Constraint Negative Proof Calculus $CNPC(\mathcal{D})$

$D = \bigvee_{i \in I} \hat{S}_i$, the *aca* $G \Rightarrow D$ asserting that the computed answers cover all the solutions of G should be derivable from $\mathcal{P}^-$. The *Constraint Negative Proof Calculus* $CNPC(\mathcal{D})$ consisting of the inference rules displayed in **Fig. 2** has been designed with the aim of enabling logical proofs $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} G \Rightarrow D$ of *aca*s. We use a special operator & in order to express the result of attaching to a given goal G a solved goal $\hat{S}'$ resulting from a previous computation, so that computation can continue from the new goal G & $\hat{S}'$.

Formally, assuming $G = \exists \overline{U}.\ (R \sqcap (\Pi \sqcap \sigma))$ and $\hat{S}' = \exists \overline{U}'.\ (\Pi' \sqcap \sigma')$ a solved goal such that $\overline{U} \setminus dom(\sigma') \subseteq \overline{U}'$, $\sigma\sigma' = \sigma'$ and $Sol_{\mathcal{D}}(\Pi') \subseteq Sol_{\mathcal{D}}(\Pi\sigma')$, the operation G & $\hat{S}'$ is defined as $\exists \overline{U}'.\ (R\sigma' \sqcap (\Pi' \sqcap \sigma'))$. The inference rule **CJ** infers an *aca* for a goal with composed kernel $(R_1 \wedge R_2) \sqcap S$ from *aca*s for goals with kernels of the form $R_1 \sqcap S$ and $(\hat{R}_2$ & $\hat{S}_i)$, respectively; while other inferences deal with different kinds of atomic goal kernels.

Any $CNPC(\mathcal{D})$-derivation $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} G \Rightarrow D$ can be depicted in the form of a *Negative Proof Tree* over $\mathcal{D}$ (shortly, NPT) with *aca*s at its nodes, such that the *aca* at any node is inferred from the *acas* at its children using some $CNPC(\mathcal{D})$ inference rule. We say that a goal solving system for $CFLP(\mathcal{D})$ is *admissible* iff whenever finitely many solved goals $\{\hat{S}_i\}_{i \in I}$ are computed as answers for an admissible initial goal G, one has $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} G \Rightarrow \bigvee_{i \in I} \hat{S}_i$ with some witnessing NPT. The next theorem is intended to provide some plausibility to the pragmatic assumption that actual $CFLP$ systems such as *Curry* [16] or $\mathcal{TOY}$ [21] are admissible goal solving systems.

Theorem 1 (Existence of Admissible Goal Solving Calculi). *There is an admissible Goal Solving Calculus $GSC(\mathcal{D})$ which formalizes the goal solving methods underlying actual $CFLP$ systems such as Curry or $\mathcal{TOY}$.*

Proof. A more general result can be proved, namely: If $(\widehat{\underline{R} \wedge R'})$ & $\hat{S} \not\sim^{p}_{\mathcal{P},GSC(\mathcal{D})} D$ (with a partially developed search space of finite size p built using the program $\mathcal{P}$, a *Goal Solving Calculus* $GSC(\mathcal{D})$ inspired in [19, 11], and a certain selection strategy that only selects atoms descendants of the part R) then $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} \hat{R}$ & $\hat{S} \Rightarrow D$ with some witnessing NPT. The proof proceeds by induction of p, using an auxiliary lemma to deal with compound goals whose kernel is a conjunction. Details are given in [10]. □

We have also proved in [10] the following theorem, showing that any *aca* which has been derived by means of a NPT is a logical consequence of the negative theory associated to the corresponding program. This result will be used below for proving the correctness of our diagnosis method.

Theorem 2 (Semantic Correctness of the $CNPC(\mathcal{D})$ Calculus). *Let $G \Rightarrow D$ be any aca for a given $CFLP(\mathcal{D})$-program $\mathcal{P}$. If $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} G \Rightarrow D$ then $G \Rightarrow D$ is a logical consequence of $\mathcal{P}^-$ in the sense of Definition 1.*

3.3 Declarative Diagnosis of Missing Answers using Negative Proof Trees

We are now prepared to present a declarative diagnosis method for missing answers which is based on NPTs and leads to correct diagnosis for any admissible goal solving system. First, we show that incompleteness symptoms are caused by incomplete program rules. This is guaranteed by the following theorem:

Theorem 3 (Missing Answers are Caused by Incomplete Program Rules). *Assume that an incompleteness symptom has been observed for a given $CFLP(\mathcal{D})$-program $\mathcal{P}$ as explained in Definition 2, with intended interpretation $\mathcal{I}_\mathcal{P}$, admissible initial goal G, and finite disjunction of computed answers $D = \bigvee_{i \in I} \hat{S}_i$. Assume also that the computation has been performed by an admissible goal solving system. Then there exists some defined function symbol f such that the axiom $(f)^-_\mathcal{P}$ for f in $\mathcal{P}^-$ is not valid in $\mathcal{I}_\mathcal{P}$, so that f's definition as given in $\mathcal{P}$ is incomplete w.r.t. $\mathcal{I}_\mathcal{P}$.*

Proof. Because of the admissibility of the goal solving system, we can assume $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} G \Rightarrow D$. Then the *aca* $G \Rightarrow D$ is a logical consequence of $\mathcal{P}^-$ because of Theorem 2. By Definition 1, we conclude that $Sol_\mathcal{I}(G) \subseteq Sol_\mathcal{D}(D)$ holds for any model $\mathcal{I}$ of $\mathcal{P}^-$. However, we also know that $Sol_{\mathcal{I}_\mathcal{P}}(G) \nsubseteq Sol_\mathcal{D}(D)$, because the disjunction D of computed answers is an incompleteness symptom w.r.t. $\mathcal{I}_\mathcal{P}$. Therefore, we can conclude that $\mathcal{I}_\mathcal{P}$ is not a model of $\mathcal{P}^-$, and therefore the completeness axiom $(f)^-_\mathcal{P}$ of some defined function symbol f must be invalid in $\mathcal{I}_\mathcal{P}$. □

The previous theorem does not yet provide a practical method for finding an incomplete function definition. As explained in Section 2, a declarative diagnosis method is expected to find the incomplete function definition by inspecting a CT. We propose to use abbreviated NPTs as CTs. Note that $\mathbf{(DF)}_f$ is the only inference rule in the $CNPC(\mathcal{D})$ calculus that depends on the program, and all the other inference rules are correct w.r.t. arbitrary interpretations. For this reason, abbreviated proof trees will omit the inference steps related to the $CNPC(\mathcal{D})$ inference rules other than $\mathbf{(DF)}_f$. More precisely, given a NPT $\mathcal{T}$ witnessing a $CNPC(\mathcal{D})$ proof $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} G \Rightarrow D$, its associated *Abbreviated Negative Proof Tree* (shortly, $ANPT$) $\mathcal{AT}$ is constructed as follows:

(1) The root of $\mathcal{AT}$ is the root of $\mathcal{T}$.

(2) The children of any node N in $\mathcal{AT}$ are the closest descendants of N in $\mathcal{T}$ corresponding to *boxed acas* introduced by $\mathbf{(DF)}_f$ inference steps.

As already explained, declarative diagnosis methods search a given CT looking for a *buggy node* whose result is unexpected but whose children's results are all expected. In our present setting, the CTs are $ANPT$s, the "results" attached to nodes are *aca*s, and a given node N is *buggy* iff the *aca* at N is *invalid* (i.e., it represents an incomplete recollection of computed answers in the intended interpretation $\mathcal{I}_\mathcal{P}$) while the *aca* at each children node N_i is *valid* (i.e., it represents a complete recollection of computed answers in the intended interpretation $\mathcal{I}_\mathcal{P}$).

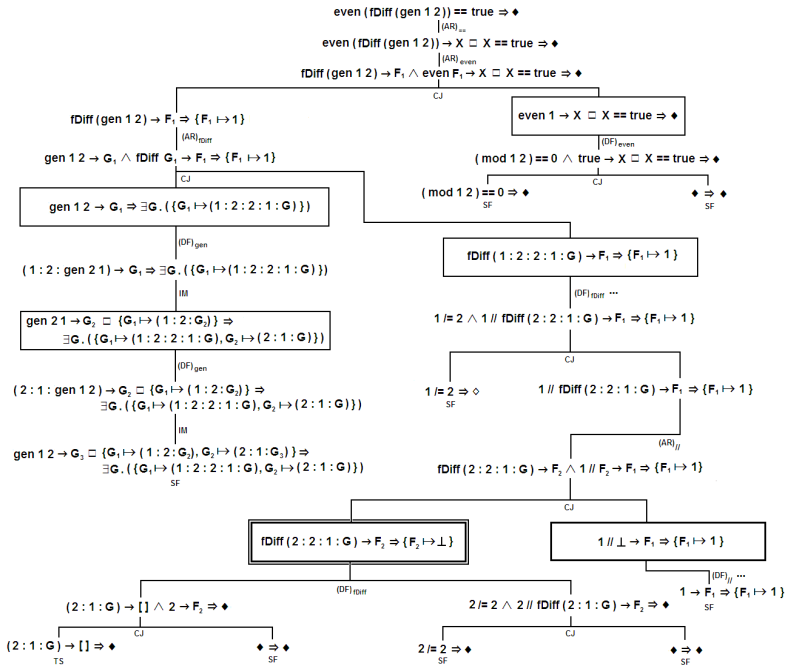


Fig. 3. *NPT* for the declarative diagnosis of missing answers

As a concrete example, **Fig. 3** displays a *NPT* which can be used for the diagnosis of missing answers in the example presented in Section 2. Buggy nodes are highlighted by encircling the *aca*s attached to them within double boxes. The *CT* shown in **Fig. 1** is the *ANPT* constructed from this *NPT*.

Our last result is a refinement of Theorem 3. It guarantees that declarative diagnosis with *ANPT*s used as *CT*s leads to the correct detection of incomplete program functions. A proof can be found in [10].

Theorem 4 (*ANPT*s Lead to the Diagnosis of Incomplete Functions). *As in Theorem 3, assume that an incompleteness symptom has been observed for a given $CFLP(\mathcal{D})$-program $\mathcal{P}$ as explained in Definition 2, with intended interpretation $\mathcal{I}_{\mathcal{P}}$, admissible initial goal G, and finite disjunction of answers $D = \bigvee_{i \in I} \hat{S}_i$, computed by an admissible goal solving system. Then $\mathcal{P}^{-} \vdash_{CNPC(\mathcal{D})} G \Rightarrow D$, and the ANPT constructed from any NPT witnessing this derivation, has some buggy node. Moreover, each such buggy node points to an axiom $(f)^{-}_{\mathcal{P}}$ which is incomplete w.r.t. the user's intended interpretation $\mathcal{I}_{\mathcal{P}}$.*

4 Implementation in the *TOY* System

In this section, we discuss the implementation in the $\mathcal{TOY}$ system of a tool based on the debugging method presented in the previous sections. The current

prototype only supports the Herbrand constraint domain $\mathcal{H}$, although the same principles can be applied to other constraint domains $\mathcal{D}$.

We summarize first the normal process followed by the $\mathcal{TOY}$ system when compiling a source program $\mathcal{P}.toy$ and solving an initial goal G w.r.t. $\mathcal{P}$. During the compilation process the system translates a source program $\mathcal{P}.toy$ into a Prolog program $\mathcal{P}.pl$ including a predicate for each function in $\mathcal{P}$. For instance the function `even` of our running example is transformed into a predicate

```
even(N,R,IC,OC):- ... code for even ... .
```

where the variable `N` corresponds to the input parameter of the function, `R` to the function result, and `IC, OC` represent, respectively, the input and output constraint store. Moreover, each goal G of $\mathcal{P}$ is also translated into a Prolog goal and solved w.r.t. $\mathcal{P}.pl$ by the underlying Prolog system. The result is a collection of answers which are presented to the user in a certain sequence, as a result of Prolog's backtracking.

If the computation of answers for G finishes after having collected finitely many answers, the user may decide that there are some missing answers (*incompleteness symptom*, in the terminology of Definition 2) and type the command `/missing` at the system prompt in order to initiate a debugging session. The debugger proceeds carrying out the following steps:

1. The object program $\mathcal{P}.pl$ is transformed into a new Prolog program $\mathcal{P}^{\mathcal{T}}.pl$. The debugger can safely assume that $\mathcal{P}.pl$ already exists because the tool is always initiated *after* some missing answer has been detected by the user. The transformed program $\mathcal{P}^{\mathcal{T}}$ behaves almost identically to $\mathcal{P}$, the only difference being that it produces a suitable *trace* of the computation in a text file. For instance here is a fragment of the code for the function `even` of our running example in the transformed program:

```
 1 % this clause wraps the original predicate
 2 even(N,R,IC,OC):-
 3     % display the input values for even
 4     write(' begin('), write(' even,'), writeq(N), write(','),
 5     write(R), write(', '), writeq(IC), write(').'), nl,
 6     % evenBis corresponds to the original predicate for even
 7     evenBis(N,R,IC,OC),
 8     % display an output result
 9     write(' output('), write(' even,'), writeq(N), write(','),
10     write(R), write(', '), writeq(OC), write(').'), nl.

11 % when all the possible outputs of the function have been produced
12 even(N,R,IC,OC):-
13     nl, write(' end(even).'), nl,
14     !,
15     fail.
16 evenBis(N,R,IC,OC) :- ... original code for even ... .
```

   As the example shows, the code for each function now displays information about the values of the arguments and the contents of the constraint store at

the moment of using any user defined function (lines 4-5). Then the predicate corresponding to the original function, now renamed with the `Bis` suffix, is called (line 7). After any successful function call the trace displays again the values of the arguments and result, which may have changed, and the contents of the output constraint store (lines 9, 10). A second clause (lines 12-15) displays the value `end` when the function has exhausted its possible output. The clause fails in order to ensure that the program flow is not changed. The original code for each function is kept unaltered in the transformed program except for the renaming (`evenBis` instead of `even` in the example, line 16). This ensures that the program will behave equivalently to the original program, except for the trace produced as a side-effect.

2. In order to obtain the trace file, the debugger repeats the computation of all the answers for the goal G w.r.t. $\mathcal{P}^{\mathcal{T}}$. After each successful computation the debugger enforces a `fail` in order to trigger the backtracking mechanism and produce the next solution for the goal. The program output is redirected to a file, where the trace is stored.
3. The trace file is then analyzed by the *CT builder* module of the tool. The result is the *Computation Tree* (an $ANPT$), which is displayed by a *Java graphical interface*.
4. The tree can be navigated by the user either manually, providing information about the validity of the *aca*s contained in the tree, or using any of the automatic strategies included in the tool which try to minimize the number of nodes that the user must examine (see [29] for a description of some strategies and their efficiency). The process ends when a buggy node is found and the tool points to an incomplete function definition, as explained in Section 3, as responsible for the missing answers. The current implementation of the prototype is available at `http://toy.sourceforge.net`. The generation of trace files works satisfactorily, while the CT builder module and the Java graphical interface do still need more improvements.

Fig. 4 shows how the tool displays the CT corresponding to the debugging scenario discussed in Section 2. The initial goal is not displayed, but the rest of the CT corresponds to **Fig. 1**, whose construction as $ANPT$ has been explained in Section 3. When displaying an *aca* $f\,\bar{t}_n \rightarrow t \,\square\, S \Rightarrow \bigvee_{i\in I} \hat{S}_i$, the tool uses list notation for representing the disjunction $\bigvee_{i\in I} \hat{S}_i$ and performs some simplifications: useless variable bindings within the stores S and S_i are dropped, as in the *aca* displayed as `gen 2 1 -> A ==> [A = 2:1:_]` in **Fig. 4**; and if t happens to be a variable X, the case $\{X \mapsto \bot\}$ is omitted from the disjunction $\bigvee_{i\in I} \hat{S}_i$, so that the user must interpret the *aca* as collecting the possible results for X other than the undefined value $\bot$. The tool also displays the underscore symbol `_` at some places. Within any *aca*, the occurrences of `_` at the right hand side of the implication $\Rightarrow$ must be understood as different existentially quantified variables, while each occurrence of `_` at the left hand side of $\Rightarrow$ must be understood as $\bot$. For instance, `1 // _ -> A ==> [A = 1]` is the *aca* `1 //` $\bot \rightarrow A \Rightarrow \{A \mapsto 1\}$ as displayed by the tool. Understanding the occurrences of `_` at the left hand

side of $\Rightarrow$ as different universally quantified variables would be incorrect. For instance, the *aca* `1 // ` $\bot \rightarrow A \Rightarrow \{A \mapsto 1\}$ is valid w.r.t. the intended interpretation $\mathcal{I}_{\mathcal{P}_{fD}}$ of $\mathcal{P}_{fD}$, while the statement $\forall X.$ (`1 // ` $X \rightarrow A \Rightarrow \{A \mapsto 1\}$) has a different meaning and is not valid in $\mathcal{I}_{\mathcal{P}_{fD}}$.

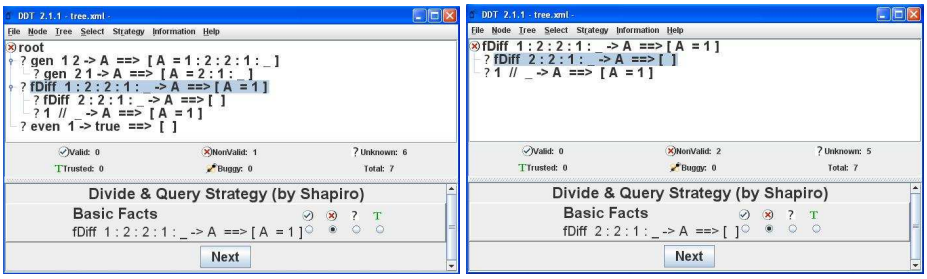


Fig. 4. Snapshots of the prototype

In the debugging session shown in **Fig. 4** the user has selected the *Divide & Query* strategy [29] in order to find a buggy node. The lower part of the left-hand side snapshot shows the first question asked by the tool after selecting this strategy, namely the *aca* `fDiff 1:2:2:1:_ -> A ==> [A=1]`. According to her knowledge of $\mathcal{I}_{\mathcal{P}_{fD}}$ the user marks this *aca* as invalid. The strategy now prunes the CT keeping only the subtree rooted by the invalid *aca* at the previous step (every CT with an invalid root must contain at least one buggy node). The second question, which can be seen at the right-hand side snapshot, asks about the validity of the *aca* `fDiff 2:2:1:_ -> A ==> []` (which in fact represents `fDiff 2:2:1:`$\bot \rightarrow A \Rightarrow \{A \mapsto \bot\}$, as explained above). Again, her knowledge of $\mathcal{I}_{\mathcal{P}_{fD}}$ leads the user to expect that `fDiff 2:2:1:`$\bot$ can return some defined result, and the *aca* is marked as invalid. After this question the debugger points out at `fDiff` as an incomplete function, and the debugging session ends. Regarding the efficiency of this debugging method our preliminary experimental results show that:

1. Producing the transformed $\mathcal{P}^{\mathcal{T}}$. *pl* from $\mathcal{P}$.*pl* is proportional in time to the number of functions of the program, and does require an insignificant amount of system memory since each predicate is transformed separately.
2. The computation of the goal w.r.t. $\mathcal{P}^{\mathcal{T}}$. *pl* requires almost the same system resources as w.r.t. $\mathcal{P}$.*pl* because writing the trace causes no significant overhead in our experiments.
3. Producing the CT from the trace is not straightforward and requires several traverses of the trace. Although more time-consuming due to the algorithmic difficulty, this process only keeps portions of the trace in memory at each moment.

4. The most inefficient phase in our current implementation is the graphical interface. Although it would be possible to keep in memory only the portion of the tree displayed at each moment, our graphical interface loads the whole CT in main memory. We plan to improve this limitation in the future. However the current prototype can cope with CTs containing thousands of nodes, which is enough for medium size computations.
5. As usual in declarative debugging, the efficiency of the tool depends on the computation tree size, which in turn usually depends on the size of the data structures required and not on the program size.

A different issue is the difficulty of answering the questions by the user. Indeed in complicated programs involving constraints the *aca*s can be large and intricate, as it is also the case with other debugging tools for CLP languages. Nevertheless, our prototype works reasonably in cases where the goal's search space is relatively small, and we believe that working with such goals can be useful for detecting many programming bugs in practice. Techniques for simplifying CTs should be worked out in future improvements of the prototype. For instance, asking the user for a concrete missing instance of the initial goal and starting a diagnosis session for the instantiated goal might be helpful.

5 Conclusions and Future Work

We have presented a novel method for the declarative diagnosis of *missing computed answers* in $CFLP(\mathcal{D})$, a declarative programming scheme which combines the expressivity of lazy FP and CLP languages. The method relies on *Computation Trees* (CTs) whose nodes are labeled with *answer collection assertions* (*aca*s). As in declarative diagnosis for FP languages, the values displayed at *aca*s are shown in the most evaluated form demanded by the topmost computation. On the other hand, and following the CLP tradition, we have shown that our CTs are abbreviated proof trees in a suitable inference system, the so-called *Constraint Negative Proof Calculus*. Thanks to this fact, we can prove the correctness of our diagnosis method for any admissible goal solving system whose recollection of computed answers can be represented by means of a proof tree in the Constraint Negative Proof Calculus. As far as we know, no comparable result was previously available for such an expressive framework as $CFLP$.

Intuitively, the notion of *aca* bears some loose relationship to programming techniques related to answer recollection, as e.g., *encapsulated search* [2]. However, *aca*s in our setting are not a programming technique. Rather, they serve as logical statements whose falsity reveals incompleteness of computed answers w.r.t. expected answers. In principle, one could also think of a kind of logical statements somewhat similar to *aca*s, but asserting the *equality* of the observed and expected sets of computed answers for one and the same goal with a finite search space. We have not developed this idea, which could support the declarative diagnosis of a third kind of unexpected results, namely *incorrect answer sets* as done for *Datalog* [5]. In fact, we think that a separate diagnosis of wrong and missing answers is pragmatically more convenient for users of $CFLP$ languages.

On the practical side, our method can be applied to actual $CFLP$ systems such as *Curry* or $\mathcal{TOY}$, leading to correct diagnosis under the pragmatic assumption that they behave as admissible goal solving systems. This assumption is plausible in so far as the systems are based on formal goal solving procedures that can be argued to be admissible. A prototype debugger under development is available, which implements the method in $\mathcal{TOY}$. Although our implementation is based on the ad-hoc trace generated by the transformed program $\mathcal{P}^{\mathcal{T}}$, we think that it could be possible to obtain the CTs from the *redex trail* for functional-logic programming described in [3]. This would allow reasoning about the correctness of the implementation by using the declarative semantics supporting this structure.

Some important pragmatic problems well known for declarative diagnosis tools in FP and CLP languages also arise in our context: both the CTs and the *aca*s at their nodes may be very big in general, causing computation overhead and difficulties for the user in answering the questions posed by the debugging tool. In spite of these difficulties, the prototype works reasonably in cases where the goal's search space is relatively small, and we believe that working with such goals can be useful for detecting many programming bugs in practice. Techniques for simplifying CTs should be worked out in future improvements of the prototype.

Acknowledgments

The authors are grateful to the referees of previous versions of this paper for their constructive comments and suggestions.

References

1. J. Boye, W. Drabent, and J. Maluszynski. Declarative diagnosis of contraint programs: An assertion-based approach. In *Automated and Algorithmic Debugging*, pages 123–140, 1997.
2. B. Brassel, M. Hanus, and F. Huch. Encapsulating non-determinism in functional logic computations. *Journal of Functional and Logic Programming*, 2004.
3. B. Brassel, M. Hanus, F. Huch, and G. Vidal. A semantics for tracing declarative multi-paradigm programs. In *PPDP '04*, pages 179–190. ACM, 2004.
4. R. Caballero. A declarative debugger of incorrect answers for constraint functional-logic programs. In *WCFLP'05*, pages 8–13. ACM, 2005.
5. R. Caballero, Y. García-Ruiz, and F. Sáenz-Pérez. A new proposal for debugging datalog programs. In *WFLP'07*, 2007.
6. R. Caballero and M. Rodríguez-Artalejo. A declarative debugging system for lazy functional logic programs. *Electr. Notes Theor. Comput. Sci.*, 64, 2002.
7. R. Caballero and M. Rodríguez-Artalejo. $\mathcal{DDT}$: A declarative debugging tool for functional-logic languages. In *FLOPS*, volume 2998 of *LNCS*, pages 70–84, 2004.
8. R. Caballero, M. Rodríguez-Artalejo, and R. del Vado-Vírseda. Declarative diagnosis of wrong answers in constraint functional-logic programming. In *ICLP'06*, volume 4079 of *LNCS*, pages 421–422. Springer, 2006.

9. R. Caballero, M. Rodríguez-Artalejo, and R. del Vado-Vírseda. Declarative debugging of missing answers in constraint functional-logic programming. In *ICLP'07*, volume 4670 of *LNCS*, pages 425–427. Springer, 2007.
10. R. Caballero, M. Rodríguez-Artalejo, and R. del Vado-Vírseda. Algorithmic debugging of missing answers in constraint functional-logic programming. Technical Report DSIC 2/08, Universidad Complutense de Madrid, 2008. Available at: `http://gpd.sip.ucm.es/papers.html`.
11. R. del Vado-Vírseda. Declarative constraint programming with definitional trees. In *FroCoS'05*, volume 3717 of *LNCS*, pages 184–199. Springer, 2005.
12. S. Estévez and R. del Vado-Vírseda. Designing an efficient computation strategy in *CFLP*($\mathcal{FD}$) using definitional trees. In *WCFLP'05*, pages 23–31. ACM, 2005.
13. A. J. Fernández, M. T. Hortalá-González, F. Sáenz-Pérez, and R. del Vado-Vírseda. Constraint functional logic programming over finite domains. *Theory and Practice of Logic Programming*, 7(5):537–582, 2007.
14. G. Ferrand. Error diagnosis in logic programming, an adaption of E. Y. Shapiro's method. *J. Log. Program.*, 4(3):177–198, 1987.
15. G. Ferrand, W. Lesaint, and A. Tessier. Towards declarative diagnosis of constraint programs over finite domains. *ArXiv Computer Science e-prints*, 2003.
16. M. Hanus. *Curry*: An integrated functional logic language (version 0.8.2 of march 28, 2006). Available at: `http://www.informatik.uni-kiel.de/~curry`, 2006.
17. M. Hermenegildo, G. Puebla, F. Bueno, and P. López-García. Abstract verification and debugging of constraint logic programs. volume 2627 of *LNCS*, pages 1–14. Springer, 2002.
18. J. W. Lloyd. Declarative error diagnosis. *New Gen. Comput.*, 5(2):133–154, 1987.
19. F. J. López-Fraguas, M. Rodríguez-Artalejo, and R. del Vado-Vírseda. A lazy narrowing calculus for declarative constraint programming. In *PPDP'04*, pages 43–54. ACM, 2004.
20. F. J. López-Fraguas, M. Rodríguez-Artalejo, and R. del Vado-Vírseda. A new generic scheme for functional logic programming with constraints. *Higher-Order and Symbolic Computation*, 20(1-2):73–122, 2007.
21. F. J. López-Fraguas and J. Sánchez-Hernández. $\mathcal{TOY}$: A multiparadigm declarative system. In *RTA'99*, volume 1631 of *LNCS*, pages 244–247. Springer, 1999.
22. L. Naish. A declarative debugging scheme. *Journal of Functional and Logic Programming*, 1997(3), 1997.
23. L. Naish and T. Barbour. A declarative debugger for a logical-functional language. DSTO General Document 5(2):91–99, 1995.
24. H. Nilsson. How to look busy while being as lazy as ever: the implementation of a lazy functional debugger. *J. Funct. Program.*, 11(6):629–671, 2001.
25. H. Nilsson and J. Sparud. The evaluation dependence tree as a basis for lazy functional debugging. *Autom. Softw. Eng.*, 4(2):121–150, 1997.
26. B. Pope. *A Declarative Debugger for Haskell*. PhD thesis, Department of Computer Science and Software Engineering, University of Melbourne, 2006.
27. B. Pope and L. Naish. Practical aspects of declarative debugging in haskell 98. In *PPDP'3*, pages 230–240. ACM, 2003.
28. E. Y. Shapiro. *Algorithmic Program DeBugging*. MIT Press, Cambridge, MA, USA, 1983.
29. J. Silva. A comparative study of algorithmic debugging strategies. In *LOPSTR'06*, volume 4407 of *LNCS*, pages 143–159. Springer, 2006.
30. A. Tessier and G. Ferrand. Declarative diagnosis in the *CLP* scheme. In *Analysis and Visualization Tools for Constraint Programming*, volume 1870 of *LNCS*, pages 151–174. Springer, 2000.

Appendix

5.1 Two Motivating Examples

In the first section of this appendix, intended only as an aid for the reviewers, we give two examples which are used to illustrate our method of declarative diagnosis of missing answers in several instances of the $CFLP(\mathcal{D})$ scheme.

Debugging of missing family relationships in $CFLP(\mathcal{H})$. The following $CFLP(\mathcal{H})$-program fragment shown below is intended to generate family relationships based on the family tree displayed in **Fig. 5**.

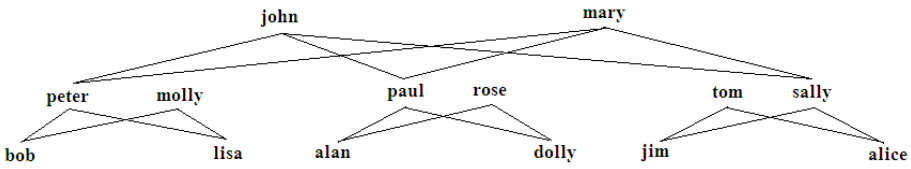


Fig. 5. Missing family relationships

```
data person = john | mary | peter | molly | paul | rose | ... | dolly | jim  | alice

maleChildOf, femaleChildOf :: person -> person -> person
maleChildOf john mary --> peter
% Analogously for other basic family facts

motherOf, fatherOf, sonOf, daughterOf, brotherOf, sisterOf :: person -> person
motherOf  X --> Y  <== maleChildOf Z Y == X // femaleChildOf Z Y == X
brotherOf X --> Y  <== maleChildOf (fatherOf X) (motherOf X) == Y, Y /= X
sonOf     X --> maleChildOf X Y // maleChildOf Y X
% Analogously for other basic family relationships

basicFamilyRelation :: person -> person
basicFamilyRelation --> motherOf // fatherOf // sonOf // ... // brotherOf // sisterOf

familyRelation :: person -> person
familyRelation --> basicFamilyRelation // basicFamilyRelation . basicFamilyRelation

(.) :: (B -> C) -> (A -> B) -> (A -> C)
(F . G) X = F (G X)
```

Note that the program represents family relationships as functions. This is possible because $\mathcal{TOY}$ is a *higher-order* language and functions can be represented as values using partial applications of top-level function names. The goal `familyRelation == R, R alice == alan` fails unexpectedly, since the existence of some family relationship linking `alice` to `alan` can be observed in the family tree underlying the program's intended meaning:

```
Toy> familyRelation == R, R alice == alan
       no
```

Different diagnosis for missing answers are possible. For instance, the answer `R -> sonOf . brotherOf . motherOf` (i.e., `alan` is son of a brother of the mother of `alice`) may be missed due to an incomplete definition of `familyRelation`, which could be extended by adding the new rule `familyRelation --> familyRelation . basicFamilyRelation`; while other answers such as `R -> cousinOf` or `R -> sonOf . uncleOf` may be missed due to an incomplete definition of `basicFamilyRelation`, which could be extended like this: `basicFamilyRelation --> ... // cousinOf // uncleOf`.

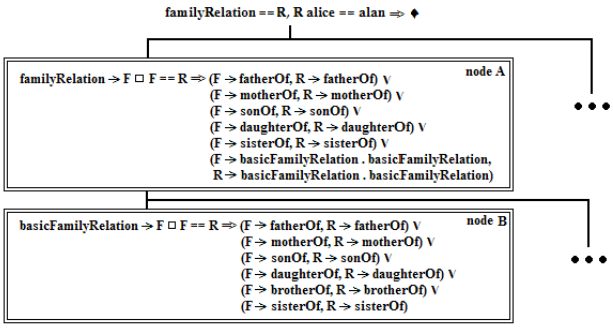


Fig. 6. *CT* for missing family relationships

Consider the *CT* partially displayed in **Fig. 6**. The programmer will judge the *aca* at the root as *invalid*, because he did not expect the finite failure. Moreover, from his knowledge of the family tree underlying intended interpretation, he might miss the answer `R -> sonOf . brotherOf . motherOf` (i.e., `alan` is a son of a brother of the mother of `alice`). If he also decided to consider that the *aca* at node B is *valid*, (i.e., the finite disjunction `(F -> fatherOf, R -> fatherOf)` $\vee \ldots \vee$ `(F -> sisterOf, R -> sisterOf)` covers all the solutions of the intermediate goal `basicFamilyRelation -> F` when satisfies the constraint `F == R`), then node A would become a *buggy node* and the function `familyRelation` would be diagnosed as *incomplete*, suggesting the addition of new program rules for computing family relations as compositions of more than two basic family relations. However, the programmer could also miss other answers such as `R -> cousinOf` or `R -> sonOf . uncleOf` and decide to view the *aca* at node B as *invalid*. In this case node B would become *buggy*, the function `basicFamilyRelation` would be diagnosed as *incomplete*, and the programmer could react by adding new program rules such as `basicFamilyRelation --> cousinOf` and `basicFamilyRelation --> uncleOf`, along with suitable program definitions for `cousinOf` and `uncleOf`.

Debugging of missing geometric transformations in $CFLP(\mathcal{R})$. The following $CFLP(\mathcal{R})$-program is intended to generate geometric transformations based on the squares A, B and C displayed in **Fig. 7**.

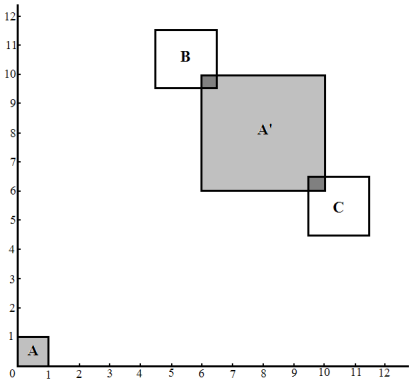


Fig. 7. Missing geometric transformations

```
% Type Definitions
type point  = (real, real)
type figure = point -> bool
type side   = real

% Figure Definitions
square :: point -> side -> figure
square (X0,Y0) S (X,Y) --> true <== X0 <= X, X <= X0 + S, Y0 <= Y, Y <= Y0 + S

isIn :: figure -> point -> bool
isIn F P --> F P

commonPoint :: figure -> figure -> point -> bool
commonPoint F G P --> true <== isIn F P, isIn G P

% Transformations
type transformation = figure -> figure

simpleTrans :: transformation
simpleTrans --> translateTo NewPoint
simpleTrans --> halfSize
simpleTrans --> doubleSize

translateTo :: point -> transformation
translateTo P (square P0 S) --> square P S

halfSize :: transformation
halfSize (square P S) --> square P (S/2)

doubleSize :: transformation
doubleSize (square P S) --> square P (2*S)

% Function composition
(.) :: (B -> C) -> (A -> B) -> (A -> C)
(.) (F . G) X --> F (G X)
```

```
transfor :: transformation
transfor --> simpleTrans
transfor --> simpleTrans . simpleTrans

% Problem Parameters
sqA, sqB, sqC :: figure
sqA --> square (0,0) 1
sqB --> square (4.5,9.5) 2
sqC --> square (9.5,4.5) 2
```

One possible user of this $CFLP(\mathcal{R})$-program should be interested in experiment with different choices of composition of simple transformations `simpleTrans`, in order to learn how is possible to transform the square A into another square A' that intersects simultaneously squares B and C, but using a minimal number of simple transformations for this aim. For example, the user can try to solve the problem using only the composition of two simple transformations `simpleTrans`, by means of the following goal in $\mathcal{TOY}$:

```
% Goal in Toy(R)

Toy(R)> transfor == T, T sqA == F, commonPoint F sqB P, commonPoint F sqC Q

        no
```

The goal fails, and the user can use the declarative diagnosis of missing answers presented in the paper to detect possibly incomplete functions in the program. In this case, it is possible to argue in an analogous way to the example of *missing family relationships*, according to the $ANPT$ displayed in **Fig. 8**.

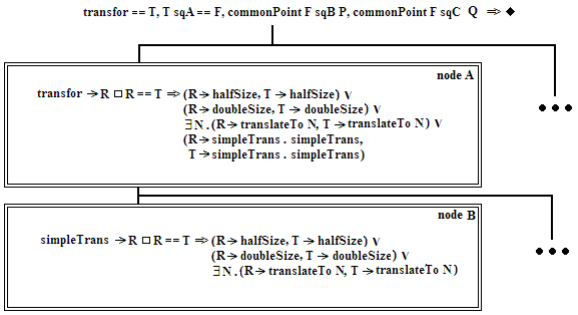


Fig. 8. $ANPT$ for missing geometric transformations

Therefore, the user could complete the `transfor` function adding another `simpleTrans` transformation to the corresponding program rule.

```
transfor :: transformation
transfor --> simpleTrans
transfor --> simpleTrans . simpleTrans . simpleTrans
```

Now, the user obtain the following answer, learning and completing the missing part of knowledge of the intended program interpretation.

```
Toy(R)> transfor == T, T sqA == F, commonPoint F sqB P, commonPoint F sqC Q

        { T -> translateTo (_A, _B) . doubleSize . doubleSize,
          F -> square (_A, _B) 4,
          P -> (_C, _D),
          Q -> (_E, _F) }

        { _C =< 6.5,
          _F =< 6.5,
          _B-_F =< -0.0,
          _D-_B =< 4.0,
          _E >= 9.5,
          _D >= 9.5,
          _A-_E >= -4.0,
          _C-_A >= 0.0 }

        Elapsed time: 15 ms.
```

5.2 Proofs of the Main Results

In this second section of the appendix, intended only as an aid for the reviewers, we give the proofs of the main results omitted in the paper. Within these proofs, we will use the *Goal Solving Calculus* $GSC(\mathcal{D})$ described in **Fig. 9** as a prototypical goal solving system based on goal transformation rules for reducing goals to solved form. In comparison to the $CDNC(\mathcal{D})$ calculus [11], $GSC(\mathcal{D})$ does not use a needed narrowing strategy based on definitional trees; and in comparison to both $CDNC(\mathcal{D})$ and $CLNC(\mathcal{D})$ [19], a rule for guessing bindings of free higher-order variables has been omitted and the invocation of a constraint solver is replaced by the goal transformation rule **SC**, intended to model the ideal behavior of a sound and complete solver. Failing derivations are identified with an inconsistent goal $\blacksquare$ and we use the notation $dvar_{\mathcal{D}}$ to represent the set of *demanded variables* in a goal, formally defined in [19] (other technical details can be found in [20]). Sometimes, for the sake of readability, we will omit the explicit use of the symbol $\widehat{\ }$ in the arguments of the & operator. Moreover, we will use the notation $\mu =_{\mathcal{X}} \mu'$ to indicate that $\mu\restriction_{\mathcal{X}} = \mu'\restriction_{\mathcal{X}}$ for all substitutions μ, μ' and $\mathcal{X} \subseteq \mathcal{V}$, and we abbreviate $\mu =_{\mathcal{V}\setminus\mathcal{X}} \mu'$ as $\mu =_{\setminus\mathcal{X}} \mu'$.

Proposition 1. (*Basic properties of the* & *operator*). *Let $G = \exists\overline{U}.(R\square S)$ be a given goal and $\hat{S}' = \exists\overline{U}'.(\Pi'\square\sigma')$ a solved goal resulting from a computation from G, and let $\mathcal{I}$ be any interpretation over the constraint domain $\mathcal{D}$. Then:*

1. $Sol_{\mathcal{I}}(G \mathbin{\&} \hat{S}') = Sol_{\mathcal{I}}(G) \cap Sol_{\mathcal{D}}(\hat{S}')$.
2. $(G \mathbin{\&} \hat{S}') \mathbin{\&} \hat{S}' = G \mathbin{\&} \hat{S}'$.
3. $G = \hat{R} \mathbin{\&} \hat{S}$.

Proof. (1) ($\subseteq$) Let $\mu \in Sol_{\mathcal{I}}(G \mathbin{\&} \hat{S}')$. By definition, there exists $\mu' =_{\setminus\overline{U}'} \mu$ such that $\mu' \in Sol_{\mathcal{D}}(R\sigma' \square S')$. Therefore, $\mu' \in Sol_{\mathcal{D}}(R\sigma')$ and $\mu' \in Sol_{\mathcal{D}}(S')$. Moreover, $\mu \in Sol_{\mathcal{D}}(\hat{S}')$ because $\mu' =_{\setminus\overline{U}'} \mu$ and $\hat{S}' = \exists\overline{U}'.\, S'$. On the other hand, since $S' = (\Pi' \square \sigma')$, it follows that $\mu' \in Sol_{\mathcal{D}}(\Pi')$ and $\mu' \in Sol(\sigma')$, and then $\sigma'\mu' = \mu'$: for each binding $\{X \to t\} \in \sigma'$, $X\sigma'\mu' = t\mu' = X\mu'$ because $\mu' \in Sol(\sigma')$, and for any $X \notin dom(\sigma')$, $X\sigma'\mu' = X\mu'$. Moreover, since $\mu' \in Sol_{\mathcal{I}}(R\sigma')$,

DC DeComposition

$\exists \overline{U}.\ ((h\overline{e}_m \to h\overline{t}_m \wedge R) \sqcap S) \Vdash_{DC_1} \exists \overline{U}.\ ((\overline{e_m \to t_m} \wedge R) \sqcap S)$ if $h\overline{e}_m$ is passive.

$\exists \overline{U}.\ ((u \to u \wedge R) \sqcap S) \Vdash_{DC_2} \exists \overline{U}.\ (R \sqcap S)$ if $u \in \mathcal{U}$.

SP Simple Production

$\exists X, \overline{U}.\ ((t \to X \wedge R) \sqcap S) \Vdash_{SP_1} \hat{R}\ \&\ \hat{S}'$ where $S' \equiv S \wedge t \to X$.

$\exists \overline{U}.\ ((X \to t \wedge R) \sqcap S) \Vdash_{SP_2} \hat{R}\ \&\ \hat{S}'$ if $t \notin \mathcal{V}$, where $S' \equiv S \wedge X \to t$.

IM Imitation

$\exists \overline{U}.\ ((h\overline{e}_m \to X \wedge R) \sqcap S) \Vdash_{IM} \exists \overline{X}_m, \overline{U}.\ ((\overline{e_m \to X_m} \wedge h\overline{X}_m \to X \wedge R) \sqcap S)$
if $X \in dvar_{\mathcal{D}}(R \sqcap S)$.

EL Elimination

$\exists X, \overline{U}.\ (e \to X \wedge R) \sqcap S \Vdash_{EL} \exists \overline{U}.\ (R \sqcap S)$ if $X \notin var(R \sqcap \Pi)$, with $S \equiv \Pi \sqcap \sigma$.

PF$_p$ Primitive Function

$\exists \overline{U}.\ ((p\overline{e}_n \to^{?} t \wedge R) \sqcap S) \Vdash_{PF_p} (\exists \overline{X}_n, \overline{U}.\ (\overline{e_n \to X_n} \wedge R))\ \&\ \hat{S}'$
if $p \in PF^n$, $t \notin \mathcal{V}$ or $t \in dvar_{\mathcal{D}}(R \sqcap S)$, and $S' \equiv S \wedge p\overline{X}_n \to! t$.

DF$_f$ Defined Function

$\exists \overline{U}.\ ((f\overline{e}_n \to t \wedge R) \sqcap S) \Vdash_{DF_{f,1}}$
$\bigvee_{i \in I} (\exists \overline{X}_n, Y, \overline{Y}_i, \overline{U}.\ ((\overline{e_n \to X_n} \wedge Y \to t \wedge R_i \wedge R) \sqcap S))$

$\exists \overline{U}.\ (f\overline{e}_n\overline{a}_k \to t \wedge R) \sqcap S \Vdash_{DF_{f,2}}$
$\bigvee_{i \in I} (\exists \overline{X}_n, Y, \overline{Y}_i, \overline{U}.\ ((\overline{e_n \to X_n} \wedge Y\overline{a}_k \to t \wedge R_i \wedge R) \sqcap S))$ if $k > 0$

where $f \in DF^n$, $t \notin \mathcal{V}$ or $t \in dvar_{\mathcal{D}}(R \sqcap S)$, $(f\overline{X}_n \to Y \Rightarrow \bigvee_{i \in I} \hat{R}_i) \in_{var} \mathcal{P}^-$, and $\hat{R}_i = \exists \overline{Y}_i.\ R_i$ for each $i \in I$.

SC Simplify Constraints

$\exists \overline{U}.\ (R \sqcap S) \Vdash_{SC} \bigvee_{i \in I} (\hat{R}\ \&\ \hat{S}_i)$ if $Sol_{\mathcal{D}}(\hat{S}) = \bigcup_{i \in I} Sol_{\mathcal{D}}(\hat{S}_i)$.

CF Conflict Failure

$\exists \overline{U}.\ ((h\overline{e}_p \to h'\overline{t}_q \wedge R) \sqcap S) \Vdash_{CF_1} \blacksquare$ if $h \neq h'$ or $p \neq q$.

$\exists \overline{U}.\ ((u \to u' \wedge R) \sqcap S) \Vdash_{CF_2} \blacksquare$ if $u, u' \in \mathcal{U}$ but $u \neq_{\mathcal{U}} u'$.

$\exists \overline{U}.\ ((h\overline{e}_p \to u' \wedge R) \sqcap S) \Vdash_{CF_3} \blacksquare$ if $u' \in \mathcal{U}$.

$\exists \overline{U}.\ ((u \to h\overline{t}_q \wedge R) \sqcap S) \Vdash_{CF_4} \blacksquare$ if $u \in \mathcal{U}$.

FC Failure Constraint $\exists \overline{U}.\ (R \sqcap S) \Vdash_{FC} \blacksquare$ if $Sol_{\mathcal{D}}(\hat{S}) = \emptyset$.

Fig. 9. The Goal Solving Calculus $GSC(\mathcal{D})$

we also have $\sigma'\mu' \in Sol_{\mathcal{I}}(R)$, and then $\mu' \in Sol_{\mathcal{I}}(R)$. Moreover, since $Sol_{\mathcal{D}}(\Pi')$ $\subseteq Sol_{\mathcal{D}}(\Pi\sigma')$ by definition and $\mu' \in Sol_{\mathcal{D}}(\Pi')$, we also have $\mu' \in Sol_{\mathcal{D}}(\Pi\sigma')$, and then $\sigma'\mu' \in Sol_{\mathcal{D}}(\Pi)$, or equivalently, $\mu' \in Sol_{\mathcal{D}}(\Pi)$. Since by definition σ' $= \sigma\sigma'$ and $\mu' \in Sol(\sigma')$ then $\mu' \in Sol_{\mathcal{D}}(\sigma)$. We deduce that $\mu' \in Sol_{\mathcal{D}}(\Pi \sqcap \sigma)$, and then $\mu' \in Sol_{\mathcal{D}}(S)$. Therefore, $\mu' \in Sol_{\mathcal{I}}(R \sqcap S)$. Finally, since $\mu' =_{\setminus \overline{U}'} \mu$,

where $\mu' = \sigma'\mu'$ and $\overline{U} \setminus dom(\sigma') \subseteq \overline{U}'$, we learn that $\mu =_{\setminus \overline{U}} \mu'$ and we conclude $\mu \in Sol_{\mathcal{I}}(G)$ because $G = \exists \overline{U}.\ (R \sqcap S)$.

($\supseteq$) Let $\mu \in Sol_{\mathcal{I}}(G) \cap Sol_{\mathcal{D}}(\hat{S}')$. Then, $\mu \in Sol_{\mathcal{I}}(G)$ and $\mu \in Sol_{\mathcal{D}}(\hat{S}')$. By definition, there exists $\mu' =_{\setminus \overline{U}'} \mu$ such that $\mu' \in Sol_{\mathcal{D}}(S')$, and then $\mu' \in Sol(\sigma')$. Moreover, there exists $\mu'' =_{\setminus \overline{U}} \mu$ such that $\mu'' \in Sol_{\mathcal{I}}(R \sqcap S)$, and then $\mu'' \in Sol_{\mathcal{I}}(R)$. However, since $\overline{U} \setminus dom(\sigma') \subseteq \overline{U}'$, we deduce that $\mu'' =_{\setminus \overline{U}'} \mu$, and therefore, $\mu'' =_{\setminus \overline{U}'} \mu'$. Then, $\mu' \in Sol_{\mathcal{I}}(R)$ and $\mu' \in Sol(\sigma')$. It follows that $\sigma'\mu' = \mu'$ (for each binding $\{X \to t\} \in \sigma'$, $X\sigma'\mu' = t\mu' = X\mu'$ because $\mu' \in Sol(\sigma')$, and for any $X \notin dom(\sigma')$, $X\sigma'\mu' = X\mu'$), and then $\mu' \in Sol_{\mathcal{I}}(R\sigma')$. Moreover, since $\mu' \in Sol_{\mathcal{D}}(S')$, we obtain $\mu' \in Sol_{\mathcal{I}}(R\sigma' \sqcap S')$. We conclude that $\mu \in Sol_{\mathcal{I}}(G \,\&\, \hat{S}')$ by definition because $\mu =_{\setminus \overline{U}'} \mu'$.

(2) By definition of the & operator, $G \,\&\, \hat{S}' = \exists \overline{U}'.\ (R\sigma' \sqcap S')$. Since trivially $\overline{U}' \setminus dom(\sigma') \subseteq \overline{U}'$, $\sigma'\sigma' = \sigma'$ because σ' is an idempotent substitution, and $Sol_{\mathcal{D}}(\Pi') \subseteq Sol_{\mathcal{D}}(\Pi'\sigma') = Sol_{\mathcal{D}}(\Pi')$ because $\Pi'\sigma' = \Pi'$ by the admissibility conditions, we can apply the definition of the & operator again to obtain $(G \,\&\, \hat{S}') \,\&\, \hat{S}' = \exists \overline{U}'.\ (R\sigma'\sigma' \sqcap S')$. Since σ' is an idempotent substitution, $\sigma'\sigma' = \sigma'$, and then $(G \,\&\, \hat{S}') \,\&\, \hat{S}' = \exists \overline{U}'.\ (R\sigma' \sqcap S') = G \,\&\, \hat{S}'$.

(3) Let $G = \exists \overline{U}.\ (R \sqcap S)$ an admissible goal, where $\hat{R} = \exists \overline{U}.\ R$ (or equivalently, $\hat{R} = \exists \overline{U}.\ (R \sqcap (\Diamond \sqcap \{\}))$, because $\hat{R}$ is the non-solved part of the goal G with an empty solved part $(\Diamond \sqcap \{\})$ such that $dom(\{\}) \cap \overline{U} = \emptyset$, $R\{\} = R$, and $\Diamond\{\} = \Diamond$), and $\hat{S} = \exists \overline{U}.\ S$ with $S = (\Pi \sqcap \sigma)$. Clearly, $\overline{U} \setminus dom(\sigma) \subseteq \overline{U}$, $\{\}\sigma = \sigma$, and $Sol_{\mathcal{D}}(\Pi) \subseteq Sol_{\mathcal{D}}(\Diamond\sigma) = Sol_{\mathcal{D}}(\Diamond) = Val_{\perp}(\mathcal{D})$. Then, we can apply the definition of the & operator to obtain $\hat{R} \,\&\, \hat{S} = \exists \overline{U}.\ (R\sigma \sqcap S)$. Since $R\sigma = R$ by the admissibility conditions of the goal G, we obtain $\hat{R} \,\&\, \hat{S} = \exists \overline{U}.\ (R \sqcap S) = G$. □

The following auxiliary lemma is a key tool for the proof of Theorem 1. Roughly, the lemma says that computing answers for a conjunction $(R_1 \wedge R_2)$ can be decomposed in two stages: first computing answers for R_1, and then proceeding with the computation of answers for R_2 as an extension of each particular answer for R_1. Formally:

Lemma 1. (*Conjunction lemma*). *Assume* $\widehat{(R_1 \wedge R_2)} \,\&\, \hat{S} \Vdash^{p}_{\mathcal{P},GSC(\mathcal{D})} D$ *(with a partially developed search space of size p). Then it is possible to find partially developed finite search spaces as follows:*

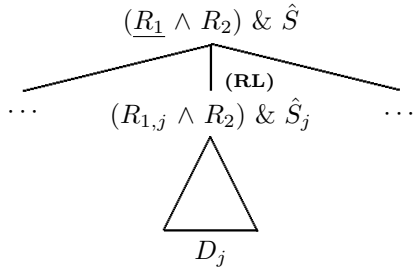
1. $\widehat{(\underline{R_1} \wedge R_2)} \,\&\, \hat{S} \Vdash^{q}_{\mathcal{P},GSC(\mathcal{D})} \bigvee_{i \in I} \hat{S}_i$ *(partially developed search space of size* $q \leq p$*).*
2. $\forall i \in I : \widehat{(R_1 \wedge \underline{R_2})} \,\&\, \hat{S}_i \Vdash^{q_i}_{\mathcal{P},GSC(\mathcal{D})} D_i$ *(partially developed search spaces of sizes* $q_i \leq p$*).*
3. $\bigvee_{i \in I} D_i = D$.

Proof. We reason by *induction on the size p* of the partial search space $(R_1 \wedge R_2) \,\&\, \hat{S} \Vdash^{p}_{\mathcal{P},GSC(\mathcal{D})} D$.

Base case: $p = 0$. In this case, $D \equiv S$, and then $(\underline{R_1} \wedge R_2)\ \&\ \hat{S} \mid\!\sim^0_{\mathcal{P},GSC(\mathcal{D})} S$ with $q = 0 \leq p$, $(R_1 \wedge \underline{R_2})\ \&\ \hat{S} \mid\!\sim^0_{\mathcal{P},GSC(\mathcal{D})} S$ with $q_1 \leq p$, and $\bigvee_{i\in I} D_i = S \equiv D$.

Inductive step: We distinguish three cases, according to the rule **(RL)** used at the root of the partial search space $(R_1 \wedge R_2)\ \&\ \hat{S} \mid\!\sim^p_{\mathcal{P},GSC(\mathcal{D})} D$.

– The rule **(RL)** is applied to R_1. Then, we have the following situation:



corresponding to the computation $(\underline{R_1} \wedge R_2)\ \&\ \hat{S} \mid\!\sim_{\mathbf{RL}} \bigvee_{j\in J} (R_{1,j} \wedge R_2)\ \&\ \hat{S}_j \mid\!\sim^{p_j}_{\mathcal{P},GSC(\mathcal{D})} \bigvee_{j\in J} D_j$ with $p_j < p$ for each $j \in J$, and $D = \bigvee_{j\in J} D_j$. Since $(R_{1,j} \wedge R_2)\ \&\ \hat{S}_j \mid\!\sim^{p_j}_{\mathcal{P},GSC(\mathcal{D})} D_j$ for all $j \in J$, by applying the *induction hypothesis*, we obtain:

$(1)_j$ $(\underline{R_{1,j}} \wedge R_2)\ \&\ \hat{S}_j \mid\!\sim^{q'_j}_{\mathcal{P},GSC(\mathcal{D})} \bigvee_{i\in I_j} \hat{S}_{j,i}$, with $q'_j \leq p_j$.

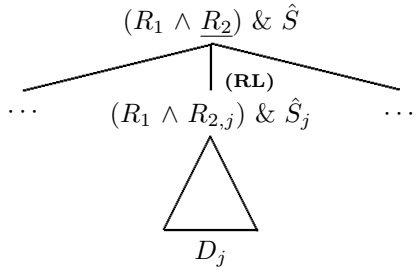
$(2)_j$ $\forall\, i \in I_j : (R_{1,j} \wedge \underline{R_2})\ \&\ \hat{S}_{j,i} \mid\!\sim^{q'_{j,i}}_{\mathcal{P},GSC(\mathcal{D})} D_{j,i}$, with $q'_{j,i} \leq p_j$ and $\bigvee_{i\in I_j} D_{j,i} = D_j$.

Then, we can compound the following computations:

(1) $(\underline{R_1} \wedge R_2)\ \&\ \hat{S} \mid\!\sim_{\mathbf{RL}} \bigvee_{j\in J} (\underline{R_{1,j}} \wedge R_2)\ \&\ \hat{S}_j \mid\!\sim^{q'_j}_{\mathcal{P},GSC(\mathcal{D})} \bigvee_{j\in J} (\bigvee_{i\in I_j} \hat{S}_{j,i})$ applying $(1)_j$, with size $1 + \sum_{j\in J} q'_j \leq 1 + \sum_{j\in J} p_j = p$.

(2) $\forall\, i \in I_j$ and $\forall\, j \in J : (R_1 \wedge \underline{R_2})\ \&\ \hat{S}_{j,i} \mid\!\sim^{q'_{j,i}}_{\mathcal{P},GSC(\mathcal{D})} D_{j,i}$ applying $(2)_j$ with size $q'_{j,i} \leq p_j < p$ and $\bigvee_{j\in J} (\bigvee_{i\in I_j} D_{j,i}) = \bigvee_{j\in J} D_j = D$.

– The rule **(RL)** is applied to R_2. Then, we have the following situation:



corresponding to the computation $(R_1 \wedge \underline{R_2})\ \&\ \hat{S} \mid\!\sim_{\mathbf{RL}} \bigvee_{j\in J} (R_1 \wedge R_{2,j})\ \&\ \hat{S}_j \mid\!\sim^{p_j}_{\mathcal{P},GSC(\mathcal{D})} \bigvee_{j\in J} D_j$ with $p_j < p$ for each $j \in J$, and $D = \bigvee_{j\in J} D_j$. Since

$(R_1 \wedge R_{2,j})$ & $\hat{S}_j \mid\!\sim^{p_j}_{\mathcal{P},GSC(\mathcal{D})} D_j$ for all $j \in J$, by applying the *induction hypothesis*, we obtain:

$(1)_j$ $(\underline{R_1} \wedge R_{2,j})$ & $\hat{S}_j \mid\!\sim^{q'_j}_{\mathcal{P},GSC(\mathcal{D})} \bigvee_{i\in I_j} \hat{S}_{j,i}$, with $q'_j \leq p_j$.

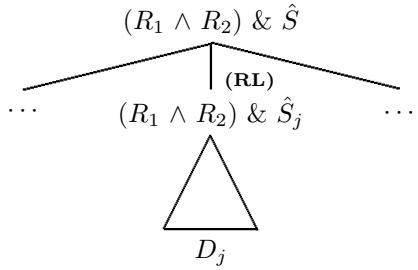
$(2)_j$ $\forall\, i \in I_j : (R_1 \wedge \underline{R_{2,j}})$ & $\hat{S}_{j,i} \mid\!\sim^{q'_{j,i}}_{\mathcal{P},GSC(\mathcal{D})} D_{j,i}$, with $q'_{j,i} \leq p_j$ and $\bigvee_{i\in I_j} D_{j,i} = D_j$.

Then, we can compound the following computations:

(1) $(\underline{R_1} \wedge R_2)$ & $\hat{S} \mid\!\sim_{\mathbf{SC}} \bigvee_{j\in J} (\underline{R_1} \wedge R_2)$ & $\hat{S}_j$ because $Sol_{\mathcal{D}}(S) = \bigcup_{j\in J} Sol_{\mathcal{D}}(S_j)$ thanks to the correctness of the rule **(RL)**. Moreover, $(\underline{R_1} \wedge R_2)$ & $\hat{S}_j \mid\!\sim^{q'_j}_{\mathcal{P},GSC(\mathcal{D})} \bigvee_{i\in I_j} \hat{S}_{j,i}$ for each $j \in J$, applying $(1)_j$ with size $1 + \sum_{j\in J} q'_j \leq 1 + \sum_{j\in J} p_j \leq p$.

(2) $\forall\, i \in I_j$ and $\forall\, j \in J : (R_1 \wedge \underline{R_2})$ & $\hat{S}_{j,i} \mid\!\sim_{\mathcal{P},GSC(\mathcal{D})} (R_1 \wedge \underline{R_{2,j}})$ & $\hat{S}_{j,i}$ for all $j \in J$, because $(R_1 \wedge \underline{R_2})$ & $\hat{S} \mid\!\sim_{\mathbf{RL}} \bigvee_{j\in J} (R_1 \wedge R_{2,j})$ & $\hat{S}_j$. Moreover, $(R_1 \wedge \underline{R_{2,j}})$ & $\hat{S}_{j,i} \mid\!\sim^{q'_{j,i}}_{\mathcal{P},GSC(\mathcal{D})} D_{j,i}$ applying $(2)_j$, with size $1 + q'_{j,i} \leq 1 + p_j \leq p$ for all $j \in J$ and for all $i \in I_j$.

– The rule **(RL)** is not applied to R_1 or R_2. Then, we have the following situation:



corresponding to the computation $(R_1 \wedge R_2)$ & $\hat{S} \mid\!\sim_{\mathbf{RL}} \bigvee_{j\in J} (R_1 \wedge R_2)$ & $\hat{S}_j \mid\!\sim^{p_j}_{\mathcal{P},GSC(\mathcal{D})} \bigvee_{j\in J} D_j$ with $p_j < p$ for each $j \in J$, and $D = \bigvee_{j\in J} D_j$. Since $(R_1 \wedge R_2)$ & $\hat{S}_j \mid\!\sim^{p_j}_{\mathcal{P},GSC(\mathcal{D})} D_j$ for all $j \in J$, by applying the *induction hypothesis*, we obtain:

$(1)_j$ $(\underline{R_1} \wedge R_2)$ & $\hat{S}_j \mid\!\sim^{q'_j}_{\mathcal{P},GSC(\mathcal{D})} \bigvee_{i\in I_j} \hat{S}_{j,i}$, with $q'_j \leq p_j$.

$(2)_j$ $\forall\, i \in I_j : (R_1 \wedge \underline{R_2})$ & $\hat{S}_{j,i} \mid\!\sim^{q'_{j,i}}_{\mathcal{P},GSC(\mathcal{D})} D_{j,i}$, with $q'_{j,i} \leq p_j$ and $\bigvee_{i\in I_j} D_{j,i} = D_j$.

Then, we can compound the following computations:

(1) $(\underline{R_1} \wedge R_2)$ & $\hat{S} \mid\!\sim_{\mathbf{RL}} \bigvee_{j\in J} (\underline{R_1} \wedge R_2)$ & $\hat{S}_j \mid\!\sim^{q'_j}_{\mathcal{P},GSC(\mathcal{D})} \bigvee_{j\in J} (\bigvee_{i\in I_j} \hat{S}_{j,i})$ applying $(1)_j$, with size $1 + \sum_{j\in J} q'_j \leq 1 + \sum_{j\in J} p_j = p$.

(2) $\forall\, i \in I_j$ and $\forall\, j \in J : (R_1 \wedge \underline{R_2})$ & $\hat{S}_{j,i} \mid\!\sim^{q'_{j,i}}_{\mathcal{P},GSC(\mathcal{D})} D_{j,i}$ applying $(2)_j$, with size $q'_{j,i} \leq p_j < p$ and $\bigvee_{j\in J} (\bigvee_{i\in I_j} D_{j,i}) = \bigvee_{j\in J} D_j = D$.

Proof of Theorem 1 (*Admissibility of the* $GSC(\mathcal{D})$ *calculus*). According to item 3 in Proposition 1 over admissible goals, we prove that if $(\underline{R} \wedge R') \square S \mid\!\sim^{p}_{\mathcal{P},GSC(\mathcal{D})} D$, or equivalently $(\underline{R} \wedge R')$ & $\hat{S} \mid\!\sim^{p}_{\mathcal{P},GSC(\mathcal{D})} D$ (with certain partial search space of size p), then $\mathcal{P}^{-} \vdash_{CNPC(\mathcal{D})} R \square S \Rightarrow D$, or equivalently $\mathcal{P}^{-} \vdash_{CNPC(\mathcal{D})} R$ & $\hat{S} \Rightarrow D$ (with a certain NPT). We reason by *induction on the size* p of the partial search space.

Base case: $p = 0$. In this case, the partial search space is a leaf $(\underline{R} \wedge R') \square S \mid\!\sim_{\mathcal{P},GSC(\mathcal{D})} D$, where $D \equiv S$, and $\mathcal{P}^{-} \vdash_{CNPC(\mathcal{D})} R \square S \Rightarrow D$ with a **SF** step, because trivially $Sol_{\mathcal{D}}(S) \subseteq Sol_{\mathcal{D}}(D)$.

Induction step: $p > 0$. In this case, we distinguish several subcases according to the goal transformation rule of the $GSC(\mathcal{D})$ calculus which is used in the root of the computation $(\underline{R} \wedge R') \square S \mid\!\sim^{p}_{\mathcal{P},GSC(\mathcal{D})} D$.

(DC) $(\underline{h\bar{e}_m \rightarrow h\bar{t}_m \wedge R_1} \wedge R') \square S \mid\!\sim_{DC} (\underline{\overline{e_m \rightarrow t_m} \wedge R_1} \wedge R') \square S \mid\!\sim^{q}_{\mathcal{P},GSC(\mathcal{D})} D$, with $p = 1 + q$ (i.e., $q < p$). Since $(\underline{\overline{e_m \rightarrow t_m} \wedge R_1} \wedge R') \square S \mid\!\sim^{q}_{\mathcal{P},GSC(\mathcal{D})} D$, by applying the *Conjunction lemma*, we obtain:

(1) $(\underline{\overline{e_m \rightarrow t_m}} \wedge R_1 \wedge R') \square S \mid\!\sim^{q'}_{\mathcal{P},GSC(\mathcal{D})} \bigvee_{i \in I} \hat{S}_i$, with $q' \leq q < p$. By *induction hypothesis*, $\mathcal{P}^{-} \vdash_{CNPC(\mathcal{D})} \overline{e_m \rightarrow t_m} \square S \Rightarrow \bigvee_{i \in I} \hat{S}_i$. By applying the **DC** rule of the $CNPC(\mathcal{D})$ calculus: $\mathcal{P}^{-} \vdash_{CNPC(\mathcal{D})} h\bar{e}_m \rightarrow \bar{t}_m \square S \Rightarrow \bigvee_{i \in I} \hat{S}_i$.

(2) $\forall i \in I : (\overline{e_m \rightarrow t_m} \wedge \underline{R_1} \wedge R')$ & $\hat{S}_i \mid\!\sim^{q'_i}_{\mathcal{P},GSC(\mathcal{D})} D_i$, with $q'_i \leq q < p$ and $\bigvee_{i \in I} D_i = D$. By *induction hypothesis*, $\mathcal{P}^{-} \vdash_{CNPC(\mathcal{D})} R_1$ & $\hat{S}_i \Rightarrow D_i$ for each $i \in I$.

Finally, by applying the rule **CJ** of the $CNPC(\mathcal{D})$ calculus:

$$\frac{h\bar{e}_m \rightarrow\ h\bar{t}_m\ \square\ S \Rightarrow \bigvee_{i \in I} \hat{S}_i\ \cdots\ (R_1 \& \hat{S}_i) \Rightarrow D_i\ \cdots}{(h\bar{e}_m \rightarrow\ h\bar{t}_m\ \wedge\ R_1)\ \square\ S \Rightarrow \bigvee_{i \in I} D_i}\ \textbf{(CJ)}$$

Therefore, $\mathcal{P}^{-} \vdash_{CNPC(\mathcal{D})} (h\bar{e}_m \rightarrow h\bar{t}_m \wedge R_1) \square S \Rightarrow D$.

(SP)$_1$ In this case, $(\underline{t \rightarrow X \wedge R_1} \wedge R') \square S \mid\!\sim_{SP_1} (\underline{R_1} \wedge R')$ & $S' \mid\!\sim^{q}_{\mathcal{P},GSC(\mathcal{D})} D$, with $q < p$ and $S' \equiv S \wedge t \rightarrow X$. By *induction hypothesis*, $\mathcal{P}^{-} \vdash_{CNPC(\mathcal{D})} R_1$ & $S' \Rightarrow D$. Finally, by applying the rules **(SF)** and **(CJ)** of the $CNPC(\mathcal{D})$ calculus:

$$\frac{\dfrac{\square\ S\ \wedge\ t\ \rightarrow\ X\ \Rightarrow\ S'}{t\ \rightarrow\ X\ \square\ S\ \Rightarrow\ S'}\ \textbf{(SF)} \qquad R_1\ \&\ S'\ \Rightarrow\ D}{(t\ \rightarrow\ X\ \wedge\ R_1)\ \square\ S\ \Rightarrow\ D}\ \textbf{(CJ)}$$

Therefore, $\mathcal{P}^{-} \vdash_{CNPC(\mathcal{D})} (t \rightarrow X \wedge R_1) \square S \Rightarrow D$.

(SP)$_2$ Analogous to **(SP)**$_1$.

(IM) Now, $(\underline{h\bar{e}_m \rightarrow X \wedge R_1} \wedge R') \square S \mid\!\sim_{IM} (\underline{\overline{e_m \rightarrow X_m}} \triangleq \underline{h\overline{X}_m \rightarrow X \wedge R_1} \wedge R') \square S \mid\!\sim^{q}_{\mathcal{P},GSC(\mathcal{D})} D$, with $q < p$. Since $(\underline{\overline{e_m \rightarrow X_m} \wedge h\overline{X}_m \rightarrow X \wedge R_1} \wedge R') \square S \mid\!\sim^{q}_{\mathcal{P},GSC(\mathcal{D})} D$, by applying the *Conjunction lemma*, we obtain:

(1) $(\underline{\overline{e_m \rightarrow X_m}} \wedge \underline{h\overline{X}_m \rightarrow X} \wedge R_1 \wedge R') \square S \mid\!\sim^{q'}_{\mathcal{P},GSC(\mathcal{D})} \bigvee_{i \in I} \hat{S}_i$, with $q' \leq q < p$. By *induction hypothesis*, $\mathcal{P}^{-} \vdash_{CNPC(\mathcal{D})} \overline{e_m \rightarrow X_m} \wedge h\overline{X}_m \rightarrow X \square S \Rightarrow \bigvee_{i \in I} \hat{S}_i$. Equivalently, $\mathcal{P}^{-} \vdash_{CNPC(\mathcal{D})} \overline{e_m \rightarrow X_m} \square (S \wedge h\overline{X}_m \rightarrow X) \Rightarrow \bigvee_{i \in I}$

$\hat{S}_i$. By applying the **IM** rule of the $CNPC(\mathcal{D})$ calculus: $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} h\overline{e}_m \rightarrow X \sqcap S \Rightarrow \bigvee_{i\in I} \hat{S}_i$.

(2) $\forall i \in I : (\overline{e_m \rightarrow X_m} \wedge h\overline{X}_m \rightarrow X \wedge \underline{R_1} \wedge R') \;\&\; \hat{S}_i \Vdash^{q'_i}_{\mathcal{P},GSC(\mathcal{D})} D_i$, with $q'_i \leq q < p$ and $\bigvee_{i\in I} D_i = D$. By *induction hypothesis*, $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} R_1 \;\&\; \hat{S}_i \Rightarrow D_i$ for each $i \in I$.

Finally, by applying the rule **CJ** of the $CNPC(\mathcal{D})$ calculus:

$$\frac{h\overline{e}_m \rightarrow\ X\ \sqcap\ S \Rightarrow \bigvee_{i\in I} \hat{S}_i\ \cdots\ (R_1 \& \hat{S}_i) \Rightarrow D_i\ \cdots}{(h\overline{e}_m \rightarrow\ X\ \wedge\ R_1)\ \sqcap\ S \Rightarrow \bigvee_{i\in I} D_i}\ \textbf{(CJ)}$$

Therefore, $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} (h\overline{e}_m \rightarrow X \wedge R_1) \sqcap S \Rightarrow D$.

(EL) In this case, $(\underline{e \rightarrow X \wedge R_1} \wedge R') \sqcap S \Vdash_{EL} (\underline{R_1} \wedge R') \sqcap S \Vdash^{q}_{\mathcal{P},GSC(\mathcal{D})} D$, with $q < p$ and $X \notin var((R_1 \wedge R') \sqcap S)$. By *induction hypothesis*, $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} R_1 \;\&\; \hat{S} \Rightarrow D$. Finally, by applying the rules **(SF)** and **(CJ)** of the $CNPC(\mathcal{D})$ calculus:

$$\frac{\overline{e\ \rightarrow\ X\ \sqcap\ S\ \Rightarrow\ \hat{S}}\ \textbf{(SF)} \qquad R_1\ \&\ \hat{S}\ \Rightarrow\ D}{(e\ \rightarrow\ X\ \wedge\ R_1)\ \sqcap\ S\ \Rightarrow\ D}\ \textbf{(CJ)}$$

Therefore, $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} (e \rightarrow X \wedge R_1) \sqcap S \Rightarrow D$.

(PF) $(\underline{p\overline{e}_n \rightarrow t \wedge R_1} \wedge R') \sqcap S \Vdash_{PF} (\underline{\overline{e_n \rightarrow X_n} \wedge R_1} \wedge R') \;\&\; S' \Vdash^{q}_{\mathcal{P},GSC(\mathcal{D})} D$, with $q < p$ and $S' \equiv S \wedge p\overline{X}_n \rightarrow! \; t$.

Since $(\underline{\overline{e_n \rightarrow X_n} \wedge R_1} \wedge R') \;\&\; S' \Vdash^{q}_{\mathcal{P},GSC(\mathcal{D})} D$, by applying the *Conjunction lemma*, we obtain:

(1) $(\underline{\overline{e_n \rightarrow X_n}} \wedge R_1 \wedge R') \;\&\; S' \Vdash^{q'}_{\mathcal{P},GSC(\mathcal{D})} \bigvee_{i\in I} \hat{S}_i$, with $q' \leq q < p$. By *induction hypothesis*, $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} \overline{e_n \rightarrow X_n} \;\&\; S' \Rightarrow \bigvee_{i\in I} \hat{S}_i$. Equivalently, $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} \overline{e_n \rightarrow X_n} \sqcap (S \wedge p\overline{X}_n \rightarrow! \; t) \Rightarrow \bigvee_{i\in I} \hat{S}_i$. By applying the $\mathbf{AR}_p$ rule of the $CNPC(\mathcal{D})$ calculus: $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} p\overline{e}_n \rightarrow t \sqcap S \Rightarrow \bigvee_{i\in I} \hat{S}_i$.

(2) $\forall i \in I : (\overline{e_n \rightarrow X_n} \wedge \underline{R_1} \wedge R') \;\&\; \hat{S}_i \Vdash^{q'_i}_{\mathcal{P},GSC(\mathcal{D})} D_i$, with $q'_i \leq q < p$ and $\bigvee_{i\in I} D_i = D$. By *induction hypothesis*, $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} R_1 \;\&\; \hat{S}_i \Rightarrow D_i$ for each $i \in I$.

Finally, by applying the rule **CJ** of the $CNPC(\mathcal{D})$ calculus:

$$\frac{p\overline{e}_n \rightarrow\ t\ \sqcap\ S\ \Rightarrow\ \bigvee_{i\in I} \hat{S}_i\ \cdots\ (R_1 \& \hat{S}_i)\ \Rightarrow\ D_i\ \cdots}{(p\overline{e}_n \rightarrow\ t\ \wedge\ R_1)\ \sqcap\ S\ \Rightarrow\ \bigvee_{i\in I} D_i}\ \textbf{(CJ)}$$

Therefore, $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} (p\overline{e}_n \rightarrow t \wedge R_1) \sqcap S \Rightarrow D$.

(SC) In this case, $(\underline{R} \wedge R') \sqcap S \Vdash_{SC} \bigvee_{i\in I} (\underline{R} \wedge R') \;\&\; \hat{S}_i \Vdash^{q_i}_{\mathcal{P},GSC(\mathcal{D})} \bigvee_{i\in I} D_i$, with $q_i < p$ for each $i \in I$, $D = \bigvee_{i\in I} D_i$, and $Sol_{\mathcal{D}}(S) = \bigcup_{i\in I} Sol_{\mathcal{D}}(\hat{S}_i)$. By *induction hypothesis*, $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} R \;\&\; \hat{S}_i \Rightarrow D_i$ for each $i \in I$. Finally, by applying the rules **(SF)** and **(CJ)** of the $CNPC(\mathcal{D})$ calculus:

$$\frac{\overline{\sqcap\ S\ \Rightarrow\ \bigvee_{i\in I}\ \hat{S}_i}\ \textbf{(SF)}\ \cdots\ R\ \&\ \hat{S}_i\ \Rightarrow\ D_i\ \cdots}{R\ \sqcap\ S\ \Rightarrow\ \bigvee_{i\in I} D_i}\ \textbf{(CJ)}$$

Therefore, $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} R \sqcap S \Rightarrow D$.

(CF) In this case, $(\underline{h\overline{e}_p \rightarrow h'\overline{t}_q \wedge R_1 \wedge R'}) \sqcap S \Vdash_{CF} \blacksquare$, with $h \neq h'$ or $p \neq q$. By applying the rules **(TS)**, **(SF)**, and **(CJ)** of the $CNPC(\mathcal{D})$ calculus:

$$\frac{\dfrac{}{(h\overline{e}_p \rightarrow h'\overline{t}_q) \sqcap S \Rightarrow \blacklozenge}\ \textbf{(TS)} \quad \dfrac{}{R_1 \& \blacklozenge \Rightarrow \blacklozenge}\ \textbf{(SF)}}{(h\overline{e}_p \rightarrow h'\overline{t}_q \wedge R_1) \sqcap S \Rightarrow \blacklozenge}\ \textbf{(CJ)}$$

Therefore, $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} (h\overline{e}_p \rightarrow h'\overline{t}_q \wedge R_1) \sqcap S \Rightarrow \blacklozenge$.

(FS) In this case, $(\underline{R} \wedge R') \sqcap S \Vdash_{FS} \blacksquare$, with $Sol_{\mathcal{D}}(S) = \emptyset$. Since $Sol_{\mathcal{D}}(\blacklozenge) = \emptyset$, by applying the rules **(SF)** and **(CJ)** of the $CNPC(\mathcal{D})$ calculus:

$$\frac{\dfrac{}{\sqcap\ S \Rightarrow \blacklozenge}\ \textbf{(SF)} \quad \dfrac{}{R \& \blacklozenge \Rightarrow \blacklozenge}\ \textbf{(SF)}}{R \sqcap S \Rightarrow \blacklozenge}\ \textbf{(CJ)}$$

Therefore, $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} R \sqcap S \Rightarrow \blacklozenge$.

(DF) Now, in this case we have, $(\underline{f\overline{e}_n\overline{a}_k \rightarrow t \wedge R_1} \wedge R') \sqcap S \Vdash_{DF} \bigvee_{i\in I} ((\underline{\overline{e}_n \rightarrow \overline{X}_n} \wedge Y\overline{a}_k \rightarrow t \wedge R'_i \wedge \underline{R_1} \wedge R') \sqcap S) \Vdash^{q_i}_{\mathcal{P},GSC(\mathcal{D})} \bigvee_{i\in I} D_i$, with $q_i < p$ for each $i \in I$, $D = \bigvee_{i\in I} D_i$, and $(f\overline{X}_n \rightarrow Y \Rightarrow \bigvee_{i\in I} \hat{R}'_i) \in_{var} \mathcal{P}^-$. Since we have the finite partial search space $((\underline{\overline{e}_n \rightarrow \overline{X}_n} \wedge Y\overline{a}_k \rightarrow t \wedge R'_i \wedge R_1 \wedge R') \sqcap S) \Vdash^{q_i}_{\mathcal{P},GSC(\mathcal{D})} D_i$ for each $i \in I$, by applying the *Conjunction lemma*, we obtain:

$(1)_i$ $(\underline{\overline{e_n \rightarrow X_n} \wedge Y\overline{a}_k \rightarrow t \wedge R'_i} \wedge R_1 \wedge R') \sqcap S \Vdash^{q'_i}_{\mathcal{P},GSC(\mathcal{D})} \bigvee_l \hat{S}_{il}$, with $q'_i \leq q_i < p$ for all $i \in I$.

$(2)_i$ $\forall\ l : (\overline{e_n \rightarrow X_n} \wedge Y\overline{a}_k \rightarrow t \wedge R'_i \wedge \underline{R_1} \wedge R') \ \&\ \hat{S}_{il} \Vdash^{q'_{ij}}_{\mathcal{P},GSC(\mathcal{D})} D_{il}$, with $q'_{ij} \leq q < p$, and $D_i = \bigvee_l D_{il}$ for all $i \in I$. By *induction hypothesis*, $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} R_1 \ \&\ \hat{S}_{il} \Rightarrow D_{il}$ for all $i \in I$ and for all l.

We can apply again the *Conjunction lemma* to the finite partial search space $(1)_i$ for each $i \in I$:

$(1.1)_i$ $(\overline{e_n \rightarrow X_n} \wedge Y\overline{a}_k \rightarrow t \wedge \underline{R'_i} \wedge R_1 \wedge R') \sqcap S \Vdash^{q''_i}_{\mathcal{P},GSC(\mathcal{D})} \bigvee_j \hat{S}_{ij}$, with $q''_i \leq q'_i \leq q_i < p$ for all $i \in I$. By *induction hypothesis*, $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} R'_i \sqcap S \Rightarrow \bigvee_j \hat{S}_{ij}$ for all $i \in I$.

$(1.2)_i$ $\forall\ j : (\underline{\overline{e_n \rightarrow X_n} \wedge Y\overline{a}_k \rightarrow t} \wedge R'_i \wedge R_1 \wedge R') \ \&\ \hat{S}_{ij} \Vdash^{q''_{ij}}_{\mathcal{P},GSC(\mathcal{D})} D_{ij}$, with $q''_{ij} \leq q'_i \leq q_i < p$ and $\bigvee_j D_{ij} = \bigvee_l \hat{S}_{il}$ for all $i \in I$. By *induction hypothesis*, $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} (\overline{e_n \rightarrow X_n} \wedge Y\overline{a}_k \rightarrow t) \ \&\ \hat{S}_{ij} \Rightarrow D_{ij}$.

Combining the derivations obtained from $(2)_i$, $(1.1)_i$ and $(1.2)_i$ by means of the inference rules **(CJ)**, $\mathbf{(AR)}_f$ and $\mathbf{(DF)}_f$ of the $CNPC(\mathcal{D})$ calculus, we obtain the following derivations:

$$\frac{\cdots\ R'_i \sqcap S \Rightarrow \bigvee_j \hat{S}_{ij}\ \cdots}{f\overline{X}_n \rightarrow Y \sqcap S \Rightarrow (S \wedge \bot \rightarrow Y) \vee (\bigvee_{ij} \hat{S}_{ij})}\ \mathbf{(DF)}_f$$

$\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} (\overline{e_n \rightarrow X_n} \wedge Y\overline{a}_k \rightarrow t) \ \&\ (S \wedge \bot \rightarrow Y) \Rightarrow \blacklozenge$ because $\bot\overline{a}_k \rightarrow t$ is not derivable in $CNPC(\mathcal{D})$ if $k > 0$.

$\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} (\overline{e_n \rightarrow X_n} \wedge Y\overline{a}_k \rightarrow t) \ \&\ \hat{S}_{ij} \Rightarrow D_{ij}$

By applying **(CJ)**, we obtain:

$$\frac{(\overline{e_n \rightarrow X_n} \wedge f\overline{X}_n \rightarrow Y \wedge Y\overline{a}_k \rightarrow t) \sqcap S \Rightarrow \bigvee_{ij} D_{ij} = \bigvee_{il} \hat{S}_{il}}{f\overline{e}_n\overline{a}_k \rightarrow t \sqcap S \Rightarrow \bigvee_{il} D_{il}} \ \textbf{(AR)}_f$$

$\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} R_1 \ \& \ \hat{S}_{il} \Rightarrow D_{il}$ for all $i \in I$ and for all l.
Finally, by applying **(CJ)** again:
$\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} (f\overline{e}_n\overline{a}_k \rightarrow t \wedge R_1) \sqcap S \Rightarrow \bigvee_{il} D_{il}$
Since $\bigvee_{il} D_{il} = \bigvee_{i \in I} D_i = D$, we conclude that
$\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} (f\overline{e}_n\overline{a}_k \rightarrow t \wedge R_1) \sqcap S \Rightarrow D$.

Proof of Theorem 2 (*Semantic correctness of the* $CNPC(\mathcal{D})$ *calculus*). We note that each inference rule in $CNPC(\mathcal{D})$ has the form:

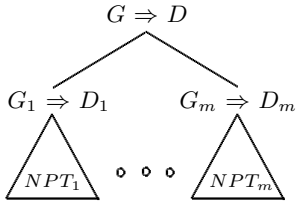
$$\frac{G_1 \Rightarrow D_1 \quad \ldots \quad G_m \Rightarrow D_m}{G \Rightarrow D} \quad \textbf{(IR)}$$

where $G_i \Rightarrow D_i$ $(1 \leq i \leq m)$ are the *premises* of the rule and $G \Rightarrow D$ is the *conclusion* of the rule. Each of this inference rules separately is semantically correct in the following sense: any model $\mathcal{I} \models_{\mathcal{D}} \mathcal{P}^-$ which satisfies the premises (i.e., such that $Sol_{\mathcal{I}}(G_i) \subseteq Sol_{\mathcal{D}}(D_i)$ holds for all $1 \leq i \leq m$) also satisfies the conclusion (i.e., $Sol_{\mathcal{I}}(G) \subseteq Sol_{\mathcal{D}}(D)$). Moreover, the assumption that the interpretation $\mathcal{I}$ is a model of $\mathcal{P}^-$ is necessary only for the inference rule $\textbf{(DF)}_f$. Any other inference rule **(IR)** of the $CNPC(\mathcal{D})$ calculus is correct w.r.t. any arbitrary interpretation $\mathcal{I}$ over $\mathcal{D}$.

Assume now that $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} G \Rightarrow D$ has been proved with an NPT of a certain depth p. In order to conclude that $\mathcal{P}^- \models_{\mathcal{D}} G \Rightarrow D$ as asserted by the theorem, and because of Definition 1, we must prove that $G \Rightarrow D$ is satisfied by $\mathcal{I}$ (i.e., the inclusion $Sol_{\mathcal{I}}(G) \subseteq Sol_{\mathcal{D}}(D)$ holds) for any arbitrarily fixed interpretation $\mathcal{I}$ such that $\mathcal{I} \models_{\mathcal{D}} \mathcal{P}^-$. This can be done by induction on p, using the already proved correctness of each individual inference rule **(IR)** of $CNPC(\mathcal{D})$.

Base case: $p = 0$. In this case, the *aca* $G \Rightarrow D$ has been inferred by means of an inference rule **(IR)** with zero premises. Since this rule is correct and all its (zero) premises are trivially satisfied in $\mathcal{I}$, we can conclude that $G \Rightarrow D$ is also satisfied in $\mathcal{I}$.

Induction step: $p > 0$. In this case, the NPT that witnesses $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} G \Rightarrow D$ has the following form:



where the depth of each NPT_i is $p_i < p$ and the *aca* $G \Rightarrow D$ has been inferred from the m *acas* $G_i \Rightarrow D_i$ by means of an inference rule **(IR)** with m premises Then, $G_i \Rightarrow D_i$ is satisfied in $\mathcal{I}$ for all $1 \leq i \leq m$ by induction hypothesis, and since **(IR)** is correct, we can conclude that $G \Rightarrow D$ is also satisfied in $\mathcal{I}$.

Proof of Theorem 4 (*$ANPT$s lead to the diagnosis of incomplete functions*). We start as in the proof of Theorem 3: the admissibility of the goal solving system guarantees a derivation $\mathcal{P}^- \vdash_{CNPC(\mathcal{D})} G \Rightarrow D$ with a certain NPT, say $\mathcal{T}$. Let $\mathcal{AT}$ be the $ANPT$ built from $\mathcal{T}$. The root of $\mathcal{AT}$ is the same as that of $\mathcal{T}$, and its attached *aca* is invalid in the intended interpretation $\mathcal{I}_\mathcal{P}$ because $\mathcal{T}$ witnesses an incompleteness symptom. By induction on the depth of $\mathcal{AT}$ it is easy to prove that $\mathcal{AT}$ has some buggy node N. Because of the construction of $\mathcal{AT}$ from $\mathcal{T}$, N is also a node of $\mathcal{T}$, and the children of N in $\mathcal{AT}$ are the closest descendants of N in $\mathcal{T}$ corresponding to *boxed acas* introduced by $\mathbf{(DF)}_f$ inference steps. Therefore, each child of N in $\mathcal{T}$ can be inferred from some children of N in $\mathcal{AT}$ by means of $CNPC(\mathcal{D})$ inference rules other than $\mathbf{(DF)}_f$. Such inferences are correct w.r.t. every interpretation, as we have seen in the proof of Theorem 2. On the other hand, since N is a buggy node in $\mathcal{AT}$, the *aca* attached to any child of N in $\mathcal{AT}$ is valid in $\mathcal{I}_\mathcal{P}$. From all this we can conclude that N is also a buggy node in $\mathcal{T}$, and the $CNPC(\mathcal{D})$ rule which relates N to its children in $\mathcal{T}$ must be $\mathbf{(DF)}_f$ for some defined function symbol $f \in DF^n$ (any other $CNPC(\mathcal{D})$ inference rule **(IR)** is correct w.r.t. $\mathcal{I}_\mathcal{P}$, and could not infer the invalid *aca* at N from the valid *aca*s at N's children). Moreover, N cannot be the root node of $\mathcal{T}$, which is never the result of a $\mathbf{(DF)}_f$ inference.

To conclude the proof, let us note that the $\mathbf{(DF)}_f$ step linking N to its children in $\mathcal{T}$ infers an *aca* which is invalid in the intended interpretation $\mathcal{I}_\mathcal{P}$ from *aca*s which are valid in $\mathcal{I}_\mathcal{P}$. The proof of Theorem 2 shows that this incorrect inference would not be possible if the completeness axiom $(f)^-_\mathcal{P}$ were valid in $\mathcal{I}_\mathcal{P}$. Therefore, $(f)^-_\mathcal{P}$ is not valid in $\mathcal{I}_\mathcal{P}$; in other words, the defined function f associated to the buggy node N is incomplete w.r.t. the intended interpretation.