Generalized Generalization Generalizers

Extended Abstract

Halime Büyükyıldız and Pierre Flener

Department of Computer Engineering and Information Science
Faculty of Engineering, Bilkent University, 06533, Bilkent, Ankara, Turkey
Email: \{halime, pf\}@cs.bilkent.edu.tr

1 Introduction

Schema-guided program construction and synthesis were studied in logic programming [4, 5, 10, 16] and in functional programming [14, 15]. Using schemas for logic program transformation was first studied in [9] and extended in [17]. Schema-guided program transformation was also studied in [7, 11]. If the transformation schema embodies some generalization technique, then it is called generalization schema (or: generalization generalizer, as in the title of this paper). This paper results from the research that began by investigating the suggestions in [7]. The contributions of this research are:

- pre-compilation of more generalized generalization schemas (tupling and descending), since the schemas in [7] were restricted to special families of divide-and-conquer programs;
- pre-compilation of new generalization schemas that we call simultaneous-tupling-and-descending generalization schemas;
- validation of the generalization schemas, based on the notions of correctness of a program, steadfastness of a program in a set of specifications, and equivalence of two programs;
- performance tests for evaluation of the generalization schemas.

Throughout the paper, the word program (resp. procedure) is used to mean typed definite program (resp. procedure). The definitions for correctness of a program, steadfastness of a program in a set of specifications, and equivalence of two programs, which is used in proving the transformation schemas, are given in Appendix A. The specification of a predicate \( r \) is written in the format:

\[
\forall X : T_X, \forall Y : T_Y. \quad I_r(X) \Rightarrow [r(X, Y) \Leftrightarrow O_r(X, Y)]
\]

where \( I_r(X) \) denotes the input condition that must be fulfilled before the execution of the procedure, and \( O_r(X, Y) \) denotes the output condition that will be fulfilled after the execution.

1.1 Program Schemas

The notion of program schema was also used in [4, 5, 9, 10, 11, 17], but here we have an additional component, which makes our definition [6, 7, 8] of program schemas different from their definitions.

Definition 1 A program schema contains a template program with a fixed data flow, but without specific indications about the actual computations, except that they must satisfy certain (steadfastness) constraints, which are the second component of a schema.

A program schema thus abstracts a whole family of particular programs that can be obtained by instantiating the place-holders of its template to particular computations, using the specification and the program synthesized so far, so that the constraints of the schema are satisfied [6].

To illustrate our purpose, the example below is used as the initial problem throughout the paper.

Example 1 Let \( \text{flat}(B, F) \) hold iff list \( F \) is the infix representation of binary tree \( B \), where infix representation means the list representation of the infix traversal of the tree. The constant \( \text{void} \) is used to represent the empty binary tree, and the compound term \( \text{bt}(L, E, R) \) is used to represent a binary tree of root \( E \), left subtree \( L \), and right subtree \( R \).

For the problem above, two different programs and the two program schemas that these programs belong to are explained in the remainder of this section. These program schemas abstract sub-families of divide-and-conquer programs. The divide-and-conquer (DC) program schemas explained in this section
are here restricted, for pedagogical reasons only, to binary predicates with \( X \) as the induction parameter and \( Y \) as the result parameter. Another restriction in the schemas is that when \( X \) is non-minimal, then \( X \) is decomposed into one head \( HX \) and two tails \( TX_1 \) and \( TX_2 \), so that \( Y \) is composed from one head \( HY \) (which is the result of processing \( HX \)) and two tails \( TY_1 \) and \( TY_2 \) (which are the results of recursively calling the predicate \( r \) with \( TX_1 \) and \( TX_2 \), respectively) by infix composition (i.e. \( Y \) is composed by putting \( HY \) between \( TY_1 \) and \( TY_2 \)). For the sake of better understandability, the program template and the program that is an instantiation of that template are given side by side in the figures.

\[
\begin{align*}
\text{r}(X, Y) = & \quad \text{minimal}(X), \\
\text{solve}(X, Y) & \quad \text{flat}(B, F) \\
\text{r}(X, Y) = & \quad \text{non-minimal}(X), \\
\text{decompose}(X, HX, TX_1, TX_2), & \quad B = \text{void}, \\
\text{r}(TX_1, TY_1), \text{r}(TX_2, TY_2), & \quad F = []
\end{align*}
\]

The constraints on these schemas (i.e. their semantics) are shown in a companion paper [8]. If we denote the functional version of the \( \text{compose} \) predicate with \( \oplus \), then the composition of \( Y \) in template \( DCRL \) by left-to-right (LR) composition ordering can be written as \( Y = ((e \oplus TY_1) \oplus HY) \oplus TY_2 \). The composition of \( Y \) in \( DCRL \) by right-to-left (RL) composition ordering can be written as \( Y = TY_1 \oplus (HY \oplus (TY_2 \oplus e)) \).

Since \( \text{append} \), which is \( \text{compose} \) in our \( \text{flat} \) example, is associative and has \([] \) as the identity element, Programs 1 and 2 are equivalent. This shows that the problem families that the two program schemas abstract have an intersection family (resulting in equivalent programs for the problem), if \( \text{compose} \) satisfies the constraints above.

The mentioned generalized schemas for any number of heads and tails, and any composition place of the head in the result parameter (infix composition above is only one possibility) are listed elsewhere [3].

### 1.2 Schema-Guided Program Transformation

In schema-guided transformation, transformation techniques are pre-compiled at the schema-level.

**Definition 2** A transformation schema encoding a transformation technique is a triple \( \langle S_1, S_2, C \rangle \), where \( S_1 \) and \( S_2 \) are program schemas and \( C \) is a set of (applicability) conditions. Schema \( S_2 \) is expressed using all the predicates of \( S_1 \), and the steadfastness constraints on \( S_2 \) are thus a superset of those of \( S_1 \).

Thus, a schema-guided program transformation system will have a collection of transformation schemas and the transformation of a program \( P_1 \) to a program \( P_2 \) reduces to: first selecting a transformation schema \( \langle S_1, S_2, C \rangle \), such that \( P_1 \) is an instance of \( S_1 \) under some form of higher-order substitution \( \sigma \), then the conditions \( C\sigma \) must be verified, and finally \( P_2 \) is computed by applying \( \sigma \) to the schema \( S_2 \). If \( P_1 \) is synthesized in a schema-guided fashion (i.e. if \( \sigma \) is known), then transformation can be automated.

Generalization is used to transform a possibly inefficient program into a more efficient one, because the generalization process may provoke a complexity reduction by loop merging and because the output program may be (semi-)tail-recursive (which can be further transformed into an iterative program by an
optimizing interpreter). The problem generalization techniques that are used in this paper are explained in detail in [4], but using these techniques for synthesizing and/or transforming a program in a schema-guided fashion was first proposed in [4, 5], and then extended in [7].

Given a program, the generalization process works as follows: first the specification of the initial program is generalized, then a recursive program for the generalized specification is synthesized, and finally a non-recursive program for the initial problem can be written, since the initial problem is a particular case of the generalized one. The two generalization approaches used here are:

1. **Structural generalization**: The intended relation is generalized by generalizing the structure (or type) of a parameter. If a problem dealing with a term is generalized to a problem dealing with a list of terms, then this generalization is called **tupling generalization**.

2. **Computational generalization**: The intended relation is generalized so as to express the general state of a computation in terms of what has been done and what remains to be done. **Ascending** and **descending** generalizations are two particular cases of computational generalization, where in ascending generalization, information about what has already been done is also needed, but in descending generalization the information about what remains to be done is enough.

If schema $S_2$ is obtained by any method of generalization described above, the transformation schema is called a **generalization schema**.

In the remainder of this paper, we explain in detail how automation of program transformation is achieved by tupling, descending, and simultaneous tupling and descending generalization, with the help of the flat example, in Sections 2, 3, and 4, respectively. In Section 5, we discuss, by using the results of performance tests, how a prototype transformation system chooses one of the transformation schemas, if the applicability conditions of more than one transformation schema are satisfied by the input program. Finally, in Section 6, we conclude.

2 Program Transformation by Tupling Generalization

Let us generalize the initial specification of our example by using tupling generalization:

$\text{flat } \mathcal{J}(B_s, F)$ iff $F$ is the concatenation of the infix representations of the elements in binary tree list $B_s$.

The specification above can be constructed by instantiating the following specification $S_{\text{tupling}}$:

$$\forall X : \text{list of } T_X, \forall Y : T_Y. \ (\forall X : T_X. \ X \in X_s \Rightarrow I_r(X) \Rightarrow [r \text{tupling}(Xs, Y) \Leftrightarrow (Xs = [\ ] \land Y = e) \lor (Xs = [X_1, X_2, \ldots, X_n] \land \bigwedge_{i=1}^{n} O_r(X_i, Y_i) \land I_1 = Y_1 \land \bigwedge_{i=2}^{n} O_r(I_{i-1}, Y_i, I_i) \land Y = I_n)]$$

where $O_e$ is the output condition of compose, and $O_r$ is the output condition of $r$, and $n \geq 1$.

The tupling generalization schemas (one for each input divide-and-conquer program schema given in Section 1.1) (the proof of $T_G^1$ is given in Appendix B) are:

$T_G^1 : (\ DCLR, TG, C_{11} )$ where

$C_{11} : - \ \text{compose} \ \text{is associative}$

- $\forall X : T_X. \ I_r(X) \land \text{minimal}(X) \Rightarrow O_r(X, e)$

- $\forall X : T_X. \ I_r(X) \Rightarrow [\neg \text{minimal}(X) \Rightarrow \neg \text{nonminimal}(X)]$

$T_G^2 : (\ DCRL, TG, C_{12} )$ where

$C_{12} : - \ \text{compose} \ \text{is associative}$

- $\forall X : T_X. \ I_r(X) \land \text{minimal}(X) \Rightarrow O_r(X, e)$

- $\forall X : T_X. \ I_r(X) \Rightarrow [\neg \text{minimal}(X) \Rightarrow \neg \text{nonminimal}(X)]$

where the shared template $T_G$ and the corresponding program for flat are in Figure 3.

Note that the predicates in $T_G$ include all of those of $DCLR$ (resp. $DCRL$), plus $r \text{tupling}$. The steadfastness constraints of $T_G$ are thus those of $DCLR$ (resp. $DCRL$), plus $S_{\text{tupling}}$. The tupling generalization schemas above give rise to full automation of the transformation process. However, this is not the end of an effective transformation, since the transformed program can be simplified by further detection of properties of solve, process, and compose. If solve$(X,Y)$ converts $X$ into a constant ‘size’ $Y$, and process$(HX, HY)$ converts $HX$ into a constant ‘size’ $HY$, and partial evaluation
of \texttt{compose}(HX, TY, Y) is possible when the 'size' of HX is constant, then Program 3 can be further
simplified. The conjunction \( HF = [] \), \texttt{append}(HF, TF, F) in the second clause of \texttt{flat} can be trans-
formed into \( F = TF \), which can be further unfolded into the head of that clause. The conjunction
\( FL = [] \), \( I_0 = [] \), \texttt{append}(I_0, FL, I_1), \( HHF = [E] \), \texttt{append}(I_1, HHF, I_2), \( HF = I_2 \), \texttt{append}(HF, TF, F)
in the third clause of \texttt{flat} can be transformed into \( F = [E]TF \), which can also be unfolded into the head
of that clause. After doing the simplifications above and further doing syntactic transformations on the
transformed program, such as "compiling" \( U = \texttt{void} \) into \( N = \texttt{bt}(U, E, R) \) in the last clause of \texttt{flat}, the
\texttt{flat} procedure is semi-tail-recursive. The resulting \texttt{flat} program is:

\[
\texttt{flat}(B, F) \rightarrow \texttt{flat}(B, F) \\
\texttt{flat}([[], []]) \rightarrow \\
\texttt{flat}([\texttt{void}][T Bs], F) \rightarrow \texttt{flat}(T Bs, F) \\
\texttt{flat}([\texttt{bt}(\texttt{void}, E, R)][T Bs], [E]TF) \rightarrow \texttt{flat}([T Bs], TF) \\
\texttt{flat}([\texttt{bt}(L, E, R)][T Bs], F) \rightarrow L \neq \texttt{void}, \texttt{flat}([L, \texttt{bt}(\texttt{void}, E, R)][T Bs], F)
\]

\section{Program Transformation by Descending Generalization}

In [7], schema-guided program transformation by descending generalization is explained in detail and a
specific generalization schema is given for an example. In this section, descending generalization schemas
are given and we explain how we eliminate the eureka finding step by using the schemas only. Eureka
finding [7] is actually finding the specification of the descendingly generalized problem. Descending
generalization can also be called \textit{accumulation strategy} (in functional programming [2], also used in logic
programming [13]), since it introduces an accumulator parameter, which is progressively extended to the
final result. This can also be seen as transformation towards \textit{difference-structure} manipulation.

Since the conditions of each descending generalization schema are different, the process of choosing the
appropriate generalization schema for the input divide-and-conquer program is done only by checking the
conditions, then the eureka comes for free. Four descending generalization schemas (two for each
divide-and-conquer program schema) are given. The first two descending generalization schemas are:

\[
DG_1 : \{ DCLR, DGLR, C_d \} \quad \text{where}
\]
They have the same formal specification (i.e. eureka) for the output program DGLR:

\[ \forall X : \mathcal{T}_X, \forall Y, A : \mathcal{T}_Y. \ T_r(X) \Rightarrow [r_{\text{descending}}(X, Y, A) \iff \exists S : \mathcal{T}_Y. \ O_r(X, S) \land O_c(A, S, Y)] \]

where \( O_c \) is the output condition of \textit{compose}, and \( O_r \) is the output condition of \( r \).

If we apply \( DG_1 \) or \( DG_2 \) to our \textit{flat} example (since the conditions of both schemas are satisfied by Programs 1 and 2), then the descending generalization of the initial specification is:

\( \text{flat}_{D_1}(B, F, A) \) iff list \( B \) is the concatenation of list \( A \) and the infix representation of binary tree \( F \).

The transformed program and the template it belongs to (i.e. \textit{DCLR}) are in Figure 4.

The other two descending generalization schemas (the proof of \( DG_2 \) is given in Appendix C) are:

\[ DG_2 : (DCLR, DGLR, C_{d_2}) \] where

\[ C_{d_2} : - \textit{compose} \text{ is associative} \]

\[ - \textit{compose} \text{ has } e \text{ as the right identity element, where } e \text{ appears in DCLR} \]

They have the same formal specification (i.e. eureka) for the output program DGLR:

\[ \forall X : \mathcal{T}_X, \forall Y, A : \mathcal{T}_Y. \ T_r(X) \Rightarrow [r_{\text{descending}}(X, Y, A) \iff \exists S : \mathcal{T}_Y. \ O_r(X, S) \land O_c(S, A, Y)] \]

where \( O_c \) is the output condition of \textit{compose}, and \( O_r \) is the output condition of \( r \).

If we apply \( DG_2 \) or \( DG_3 \) to our \textit{flat} example (since the conditions of both schemas are satisfied by Programs 1 and 2), then the descending generalization of the initial specification is:

\( \text{flat}_{D_2}(B, F, A) \) iff list \( B \) is the concatenation of the infix representation of binary tree \( B \) and list \( A \).

The transformed program and the template it belongs to (i.e. \textit{DGLR}) are in Figure 5.

The reason why we call the descendingly generalized program schemas \textit{DGLR} and \textit{DGLR} is similar to the reason why we call the divide-and-conquer program schemas \textit{DCLR} and \textit{DCLR}, respectively. In descending generalization, the composition ordering for extending the accumulator parameter in template \textit{DGLR} is from left-to-right and the composition ordering for extending the accumulator parameter in template \textit{DGLR} is from right-to-left.

Since \textit{append}, which is \textit{compose} of Program 1, has \([\ ]\) as left and right identity element, both conditions of descending generalization schemas \( DG_1 \) and \( DG_2 \) are satisfied by Program 1. Similarly, both conditions of descending generalization schemas \( DG_2 \) and \( DG_4 \) are satisfied by Program 2. Syntactic simplifications may be done on Programs 4 and 5, but it can be observed that Program 5 is more efficient than Program 4.
\[
\text{Fig. 5: Template \textit{DGRL} and Program 5}
\]

since we can also eliminate \textit{append} by partial evaluation during simplification. For example, Program 5 may be simplified into:

\[
\text{flat}(B, F) = \text{flat}_{\text{d}2}(B, F, [])
\]

\[
\text{flat}_{\text{d}2}(\text{void}, A, A) = \text{flat}_{\text{d}2}(\text{bt}(L, E, R), F, A, A) = \text{flat}_{\text{d}2}(\text{bt}(L, E, R), F, A)
\]

Our next plan is to extend the descending generalization schemas by adding the ‘simplifiability’ conditions to the schemas, since we want to achieve the objective that the system chooses in one step the transformation that yields the most efficient output program for input problems like \textit{flat}, where the conditions of each schema for the same input program are satisfied.

\section{Program Transformation by Simultaneous-Tupling-and-Descending Generalization}

While working on constructing possible generalized generalization schemas for different input program schemas, we also tried to apply descending generalization to a tupling generalized program, and vice versa. The generalization schemas that we explain in this section are the results of this work. We call them simultaneous-tupling-and-descending generalization schemas, although the reader may notice by looking at the specification of the generalized problem that the process may also be thought of as applying descending generalization to a tupling generalized problem.

Like we did in Section 3 for descending generalization, four simultaneous-tupling-and-descending generalization schemas are given. The first two are:

\[
T_{DG1} : (\text{DCLR, TDGLR, C}_{\text{dird}}) \text{ where } C_{\text{dird}} : - \text{ compose is associative}
- \text{ compose has } \epsilon \text{ as the left and right identity element, where } \epsilon \text{ appears in DCLR}
- \forall X : T_X, L_r(X) \land \text{minimize}(X) \Rightarrow O_r(X, \epsilon)
- \forall X : T_X, L_r(X) \Rightarrow \lnot \text{minimize}(X) \iff \text{nonminimize}(X)
\]

\[
T_{DG4} : (\text{DCRL, TDGLR, C}_{\text{dir}}) \text{ where } C_{\text{dir}} : - \text{ compose is associative}
- \text{ compose has } \epsilon \text{ as the left and right identity element, where } \epsilon \text{ appears in DCRL}
- \forall X : T_X, L_r(X) \land \text{minimize}(X) \Rightarrow O_r(X, \epsilon)
- \forall X : T_X, L_r(X) \Rightarrow \lnot \text{minimize}(X) \iff \text{nonminimize}(X)
\]

If we apply \textit{T}_{DG1} or \textit{T}_{DG4} to our \textit{flat} example (since the conditions of both schemas are satisfied by Programs 1 and 2), then the simultaneous-tupling-and-descending generalization of the initial specification is:

\[
\text{flat}_{\text{d}2}(B, A) \text{ if list } F \text{ is the concatenation of list } A \text{ and the infix representations of the elements in binary tree list } B.
\]

The transformed program and the template it belongs to (i.e. \textit{T}_{DGLR}) are in Figure 6.

The other two simultaneous-tupling-and-descending generalization schemas are: \textit{T}_{DG2} : (\text{DCLR, TDGLR, C}_{\text{dird}}) \text{ and } \textit{T}_{DG3} : (\text{DCRL, TDGLR, C}_{\text{dir}}), \text{ where } C_{\text{dird}} \text{ and } C_{\text{dir}} \text{ are as above.
By looking at Table 1, we can say that the most time efficient programs for our example test problems are the ones that belong to the \textit{DGRL} program schema. It is so because the examples we chose are such that
r(X, Y) ←r
    r_\text{ld}_2([X], Y, \varepsilon)

r_\text{ld}_2(X_s, Y, A) ←
\begin{align*}
    & X_s = [], Y = A \\
    & X_s = [X[T X_s]], Y = A
\end{align*}

\begin{align*}
    & \text{minimal}(X), \text{solve}(X, H Y), \\
    & A = A_3, r_{\text{ld}_2}(T X_s, A_1, A_2), \\
    & \text{compose}(H Y, A_1, A_0), Y = A_0
\end{align*}

r_\text{ld}_2(X_s, Y, A) ←
\begin{align*}
    & X_s = [X[T X_s]], \\
    & \text{compose}(X, H X, T X_1, T X_2), \\
    & \text{minimal}(T X_1), \text{solve}(T X_1, T Y_1), \\
    & A = A_3, r_{\text{ld}_2}(T X_2[T X_s], A_2, A_3), \\
    & \text{compose}(H X, H Y), \text{compose}(H Y, A_2, A_1), \\
    & \text{compose}(T Y_1, A_1, A_0), Y = A_0
\end{align*}

\begin{align*}
    & \text{decompose}(X, H X, T X_1, T X_2), \\
    & \text{minimal}(T X_1), \text{minimal}(U), \\
    & \text{decompose}(N, H X, U, T X_2), \\
    & r_{\text{ld}_2}(T X_1, N[T X_s], Y, A)
\end{align*}

\begin{align*}
    & \text{flatd}(B, F) ←
    & \text{flatd}_2([B], F, [])
\end{align*}

\begin{align*}
    & \text{flatd}_2(B_s, F, A) ←
    & B_s = [], F = A
\end{align*}

\begin{align*}
    & \text{flatd}_2(B_s, F, A) ←
    & B_s = [B[T B_s], \\
    & B = \text{bt}(\_ \_ \_), \\
    & L = \text{void}, F L = [], \\
    & A = A_3, \text{flatd}_3([R[T B_s], A_2, A_3], \\
    & H F = [E], \text{append}(H F, A_2, A_1), \\
    & \text{append}(F L, A_1, A_0), F = A_0
\end{align*}

\begin{align*}
    & \text{flatd}_2(B_s, F, A) ←
    & B_s = [B[T B_s], \\
    & B = \text{bt}(\_ \_ \_), \\
    & \text{bt}(L, E, R), \\
    & L = \text{void}, U = \text{void}, \\
    & N = \text{bt}(U, E, R), \\
    & \text{flatd}_2([L, N[T B_s], F, A])
\end{align*}

Figure 7: Template \text{T DGRl} and Program 7

the best transformation schemas (which also give rise to the best simplifications) are \text{DG}_2 and \text{DG}_3 for the \text{DCLR} and \text{DCRL} input programs, respectively, since \text{compose} is \text{append} for all the programs. The reason above is also the reason why the \text{DCRL} programs are more efficient than the \text{DCLR} programs. For all programs with left-to-right composition ordering or accumulator extension, we lose efficiency, because the programs can’t be further simplified and we can’t eliminate \text{append}. In the descending and simultaneous-tupling-and-descending generalization schemas, this was difficult, since the percentage of the total running time of the program used by \text{compose} is very high, because of the increase in the ‘size’ of the input parameter. Except for \text{quicksort}, for all the examples, the \text{TG} and \text{T DGRl} programs are more efficient than the initial \text{DCLR} and \text{DCRL} programs, which supports our claim that we can also gain efficiency by tupling and simultaneous-tupling-and-descending generalizations. The reason why we lose efficiency in tupling generalization of \text{quicksort} is that we increase the calls to \text{decompose}, which is \text{partition} in \text{quicksort} and is a very costly operation, although we eliminate \text{compose}. If we try to compare simultaneous-tupling and descending generalization schemas with the other generalization schemas, we may think that we do not gain anything. The reader has to remember that this generalization can also be thought of as applying a descending generalization to a tupling generalized problem, and for infix \text{flat}, we gain some efficiency over tupling generalization. The results also show that, except for \text{quicksort}, the \text{T DGRl} programs are more efficient than the input divide-and-conquer programs. However, since the \text{DGRL} programs seem to be always more efficient than their \text{T DGRl} counterparts, it also appears that the proposed new schemas are useless in practice and that there is no need to perform a descending generalization on a tupling generalized program, i.e. to compose generalization techniques.

6 Conclusion and Future Work

Program generalization can be automated by using generic generalization schemas, since we fully automate the eureka discovery (i.e. finding the specification of the generalized problem). Further detection of properties of the operators in the input template helps us to simplify the program resulting from transformation, which can be done by constructing global plans for each generalization schema. Our next aim is to develop a prototype transformation system supporting this approach.
References


A Correctness and Equivalence Criteria

We will give the correctness and equivalence criteria that we use to prove our generalization schemas. The definitions for correctness of a program and steadfastness of a program in a set of specifications are similar to the definitions in [8, 12]. Our definitions are all model-theoretic.

We give again the format of the specification $S_r$ of a predicate $r$:

$$\forall X : T_X, \forall Y : T_Y. \quad I_r(X) \Rightarrow [r(X, Y) \Leftrightarrow O_r(X, Y)]$$

where $I_r(X)$ denotes the input condition that must be fulfilled before the execution of the procedure, and $O_r(X, Y)$ denotes the output condition that will be fulfilled after the execution.

An open program is a program where some of the predicates appearing in the clause bodies are not appearing in any heads of clauses, and these predicates are called undefined predicates. If all the predicates appearing in the program are defined, then the program is called a closed program. Note that our templates are thus open programs. We do not consider mutually recursive programs.

**Definition 3 (Correctness of a Closed Program)**

Let $P$ be a closed program for relation $r$. We say that $P$ is (totally) correct wrt its specification $S_r$ iff, for any ground term $t$ of $T_X$ such that $I_r(t)$ holds, the following condition holds: $P \models r(t, u)$ iff $O_r(t, u)$, for every ground term $u$ of $T_Y$.

If we replace ‘iff’ by ‘implies’ in the condition above, then $P$ is said to be partially correct wrt $S_r$, and if we replace ‘iff’ by ‘if’, then $P$ is said to be complete wrt $S_r$.

This kind of correctness is not entirely satisfactory, for two reasons. First, it defines the correctness of $P$ in terms of the procedures for the predicates in its clause bodies, rather than in terms of their specifications. Second, $P$ must be a closed program, even though it might be desirable to discuss the correctness of $P$ without having to fully implement it. So, the abstraction achieved through the introduction (and specification) of the predicates in its clause bodies is wasted. This leads us to the notion of steadfastness, which we only define here for the most interesting case, namely where all predicates occurring in the clause bodies are also known by their specifications.

**Definition 4 (Steadfastness of an Open Program in a Set of Specifications)**

Let:

- $P$ be an open program for relation $r$
- $q_1, \ldots, q_m$ be all the undefined predicate names appearing in $P$
- $S_1, \ldots, S_m$ be the specifications of $q_1, \ldots, q_m$.

We say that $P$ is steadfast wrt its specification $S_r$ in $\{S_1, \ldots, S_m\}$ iff, for any closed program $P_S$ such that:

- $P_S$ is correct wrt each specification $S_j$ ($1 \leq j \leq m$)
- $P_S$ contains no occurrences of the predicates defined in $P$,

it is the case that $P \cup P_S$ is correct wrt $S_r$.

Steadfastness has the following interesting property:

**Property:**

Let:

- $P$ be an open program for a relation $r$
- $p_1, \ldots, p_t$ be all the defined predicate names appearing in $P$ (including $r$ thus)
- $q_1, \ldots, q_m$ be all the undefined predicate names appearing in $P$
- $S_1, \ldots, S_m$ be the specifications of $q_1, \ldots, q_m$.

We have that, for $t \geq 2$, $P$ is steadfast wrt $S_r$ in $\{S_1, \ldots, S_m\}$ if every $P_i$ ($1 \leq i \leq t$) is steadfast wrt the specification of $p_i$ in the set of the specifications of all undefined predicates in $P_i$, where $P_i$ is a program for $p_i$, such that $P = \bigcup_{i=1}^{t} P_i$. 

10
For program equivalence, we do not require the two programs to have the same models, because this would not make much sense in a program generalization setting, where the transformed program features predicates that were not in the initially given program. That is why our program equivalence criterion establishes equivalence wrt the specification of a common predicate (usually the root of their call-hierarchies).

Definition 5 (Equivalence of Two Open Programs)
Let \( P \) and \( Q \) be two open programs for a predicate \( r \). We say that \( P \) is equivalent to \( Q \) wrt the specification \( S_r \) iff the following two conditions hold:

\[ \begin{align*}
(a) & \ P \text{ is steadfast wrt } S_r \text{ in } \{S_1, \ldots, S_m\}, \text{ where } S_1, \ldots, S_m \text{ are the specifications of } p_1, \ldots, p_m, \text{ which are all the undefined predicate names appearing in } P \\
(b) & \ Q \text{ is steadfast wrt } S_r \text{ in } \{S'_1, \ldots, S'_t\}, \text{ where } S'_1, \ldots, S'_t \text{ are the specifications of } q_1, \ldots, q_t, \text{ which are all the undefined predicate names appearing in } Q.
\end{align*} \]

Since the ‘is equivalent to’ relation is symmetric, we also say that \( P \) and \( Q \) are equivalent wrt \( S_r \).

B Proof of \( TG_1 \)
As a reminder, we give again the generalization schema \( TG_1 \).

\[ TG_1 : \{ \text{DCLR, } TG, C_{11} \} \text{ where } \]

\[ C_{11} : \]

1. \( \text{compose} \) is associative
2. \( \text{compose} \) has \( e \) as the left and right identity element, where \( e \) appears in \( \text{DCLR} \)
3. \( \forall X : TX . \ I_r(X) \land \text{minimal}(X) \Rightarrow O_r(X, e) \)
4. \( \forall X : TX . \ I_r(X) \Rightarrow [\neg\text{minimal}(X) \Rightarrow \neg\text{nonminimal}(X)] \)

and templates \( \text{DCLR} \) and \( TG \) are in Figure 8.

\[ \begin{align*}
\text{r}(X, Y) & \rightarrow \\
\text{minimal}(X), & \text{r}(X, Y) \\
\text{solve}(X, Y) & \rightarrow \\
\text{nonminimal}(X), & \text{r}(X, Y) \\
\text{decompose}(X, HX, TX_1, TX_2), & \text{r}(X, Y) \\
r(TX_1, TY_1), r(TX_2, TY_2), & \text{r}(X, Y) \\
I_0 = e, \text{compose}(I_0, TY_1, I_1), & \text{r}(X, Y) \\
\text{process}(HX, HY), \text{compose}(I_1, HY, I_2), & \text{r}(X, Y) \\
\text{compose}(I_2, TY_2, I_3), Y = I_3 & \text{r}(X, Y)
\end{align*} \]

\[ \begin{align*}
r(X, Y) & \rightarrow \\
r \text{bulting}([X], Y) & \text{r}(X, Y) \\
r \text{bulting}(Xs, Y) & \rightarrow \\
Xs = [], Y = e & \text{r}(X, Y) \\
r \text{bulting}(Xs, Y) & \rightarrow \\
Xs = [X[TXs], \text{minimal}(X), & \text{r}(X, Y) \\
r \text{bulting}(TXs, TY), & \text{r}(X, Y) \\
solve(X, HY), \text{compose}(HY, TY, Y) & \text{r}(X, Y) \\
r \text{bulting}(Xs, Y) & \rightarrow \\
Xs = [X[TXs], \text{nonminimal}(X), & \text{r}(X, Y) \\
\text{decompose}(X, HX, TX_1, TX_2), & \text{r}(X, Y) \\
\text{minimal}(TX_1), & \text{r}(X, Y) \\
r \text{bulting}([TX_2[TXs], TY), & \text{r}(X, Y) \\
solve(TX_1, TY_1), & \text{r}(X, Y) \\
I_0 = e, \text{compose}(I_0, TY_1, I_1), & \text{r}(X, Y) \\
\text{process}(HX, HY), \text{compose}(I_1, HY, I_2), & \text{r}(X, Y) \\
HY = I_2, \text{compose}(HY, TY, Y) & \text{r}(X, Y)
\end{align*} \]

Figure 8: Templates \( \text{DCLR} \) and \( TG \)
The specification \( S_r \) of predicate \( r \) is:

\[
\forall X : TX, \forall Y : TY. \quad I_r(X) \implies [r(X, Y) \Leftrightarrow O_r(X, Y)]
\]

The specification \( S_{\text{tupling}} \) of predicate \( \text{tupling} \) is:

\[
\forall Xs : \text{list of } TX, \forall Y : TY. \quad (\forall X : TX. \ X \in Xs \implies I_r(X)) \implies [\text{tupling}(Xs, Y) \Leftrightarrow (Xs = \epsilon) \wedge Y = \epsilon) \\
\forall (Xs = \left[ X_1, \ldots, X_q \right] \wedge \bigwedge_{i=1}^{q} O_r(X_i, Y_i) \wedge I_1 = Y_1 \wedge \bigwedge_{i=2}^{q} O_r(I_{i-1}, Y_i, I_i) \wedge Y = I_q)
\]

where \( O_r \) is the output condition of \( \text{compose} \), \( O_r \) is the output condition of \( r \), and \( q \geq 1 \).

To prove the generalization schema \( TG_1 \), we have to prove that templates \( DCLR \) and \( TG \) are equivalent wrt \( S_r \). By Definition 5, this holds if the following two conditions hold:

1. \( DCLR \) is steadfast wrt \( S_r \) in \( S = \{ S_{\text{minimal}}, S_{\text{nonminimal}}, S_{\text{solve}}, S_{\text{compose}}, S_{\text{compose}} \} \), where \( S_{\text{minimal}} \), \( S_{\text{nonminimal}} \), \( S_{\text{solve}} \), \( S_{\text{compose}} \) are the specifications of \( \text{minimal} \), \( \text{nonminimal} \), \( \text{solve} \), \( \text{compose} \), \( \text{compose} \), which are all the undefined predicate names appearing in \( DCLR \).

2. \( TG \) is also steadfast wrt \( S_r \) in \( S \).

Note that the sets \( \{ S_1, \ldots, S_m \} \) and \( \{ S'_1, \ldots, S'_q \} \) in Definition 5 are equal when \( Q \) is obtained by generalization of \( P \).

In program transformation, we assume that the input program, here template \( DCLR \), is steadfast wrt \( S_r \) in \( S \), so condition (1) always holds.

To prove equivalence, we have to prove condition (b). We will use the property of steadfastness: \( TG \) is steadfast wrt \( S_r \) in \( S \) if \( P_{r_{\text{tupling}}} \) is steadfast wrt \( S_{r_{\text{tupling}}} \) in \( S \), where \( P_{r_{\text{tupling}}} \) is the procedure for \( \text{tupling} \), and \( P_r \) is steadfast wrt \( S_r \) in \( \{ S_{\text{minimal}}, S_{\text{solve}}, S_{\text{compose}} \} \), where \( P_r \) is the procedure for \( r \).

To prove that \( P_{r_{\text{tupling}}} \) is steadfast wrt \( S_{r_{\text{tupling}}} \) in \( S \), we do a constructive forward proof that we begin with \( S_{r_{\text{tupling}}} \) and from which we try to obtain \( P_{r_{\text{tupling}}} \).

If we separate the cases of \( q \geq 1 \) by \( q = 1 \) or \( q \geq 2 \), then \( S_{r_{\text{tupling}}} \) becomes:

\[
\forall Xs : \text{list of } TX, \forall Y : TY. \quad (\forall X : TX. \ X \in Xs \implies I_r(X)) \implies [\text{tupling}(Xs, Y) \Leftrightarrow (Xs = \epsilon) \wedge Y = \epsilon) \\
\forall (Xs = \left[ X_1 \right] \wedge O_r(X_1, Y_1) \wedge Y_1 = I_1 \wedge Y = I_1) \\
\forall (Xs = \left[ X_1, X_2, \ldots, X_q \right] \wedge \bigwedge_{i=1}^{q} O_r(X_i, Y_i) \wedge I_1 = Y_1 \wedge \bigwedge_{i=2}^{q} O_r(I_{i-1}, Y_i, I_i) \wedge Y = I_q)
\]

where \( q \geq 2 \).

By using applicability conditions (1) and (2):

\[
\forall Xs : \text{list of } TX, \forall Y : TY. \quad (\forall X : TX. \ X \in Xs \implies I_r(X)) \implies [\text{tupling}(Xs, Y) \Leftrightarrow (Xs = \epsilon) \wedge Y = \epsilon) \\
\forall (Xs = \left[ X_1 | TXs \right] \wedge TXs = \left[ X_2, \ldots, X_q \right] \wedge \bigwedge_{i=1}^{q} O_r(X_i, Y_i) \wedge I_1 = Y_1 \wedge \bigwedge_{i=2}^{q} O_r(I_{i-1}, Y_i, I_i) \wedge Y = I_q)
\]

where \( q \geq 2 \).

By folding using \( S_{r_{\text{tupling}}} \), and renaming:

\[
\forall Xs : \text{list of } TX, \forall Y : TY. \quad (\forall X : TX. \ X \in Xs \implies I_r(X)) \implies [\text{tupling}(Xs, Y) \Leftrightarrow (Xs = \epsilon) \wedge Y = \epsilon) \\
\forall (Xs = \left[ X | TXs \right] \wedge O_r(X, HY) \wedge \text{tupling}(TXs, TY) \wedge O_r(HY, TY, Y))
\]

By taking the ‘decomposition’:

- **clause 1**: \( \text{tupling}(Xs, Y) \leftarrow Xs = \epsilon \)

- **clause 2**: \( \text{tupling}(Xs, Y) \leftarrow Xs = \left[ X | TXs \right], r(X, HY), \text{tupling}(TXs, TY), \text{compose}(HY, TY, Y) \)

By unfolding clause 2 wrt \( r(X, HY) \) using \( DCLR \), and using the assumption that \( DCLR \) is steadfast wrt \( S_r \) in \( S \).
clause 3: \( r \uparrow \text{tupling}(Xs, Y) \leftarrow \\
\quad Xs = [Xs, TXs], \\
\quad \text{minimal}(X), \\
\quad r \uparrow \text{tupling}(TXs, TY), \\
\quad \text{solve}(X, HY), \text{compose}(HY, TY, Y) \\
\)

clause 4: \( r \uparrow \text{tupling}(Xs, Y) \leftarrow \\
\quad Xs = [Xs, TXs], \\
\quad \text{nonminimal}(X), \text{decompose}(X,HX,TX1,TX2), \\
\quad r(TX1,TY1), r(TX2,TY2), \\
\quad I_0 = e, \text{compose}(I_0, TY1, I_1), \\
\quad \text{process}(HX, HY), \text{compose}(I_1, HY, I_2), \\
\quad \text{compose}(I_2, TY2, I_3), HY = I_3, \\
\quad r \uparrow \text{tupling}(TXs, TY), \text{compose}(HY, TY, Y) \\
\)

By introducing \( \text{minimal}(TX1) \lor \text{nonminimal}(TX1) \) in clause 4, using applicability condition (4):

clause 5: \( r \uparrow \text{tupling}(Xs, Y) \leftarrow \\
\quad Xs = [Xs, TXs], \\
\quad \text{nonminimal}(X), \text{decompose}(X,HX,TX1,TX2), \\
\quad \text{minimal}(TX1), r(TX1, TY1), r(TX2, TY2), \\
\quad I_0 = e, \text{compose}(I_0, TY1, I_1), \\
\quad \text{process}(HX, HY), \text{compose}(I_1, HY, I_2), \\
\quad \text{compose}(I_2, TY2, I_3), HY = I_3, \\
\quad r \uparrow \text{tupling}(TXs, TY), \text{compose}(HY, TY, Y) \\
\)

clause 6: \( r \uparrow \text{tupling}(Xs, Y) \leftarrow \\
\quad Xs = [Xs, TXs], \\
\quad \text{nonminimal}(X), \text{decompose}(X,HX,TX1,TX2), \\
\quad \text{minimal}(TX1), r(TX1, TY1), r(TX2, TY2), \\
\quad I_0 = e, \text{compose}(I_0, TY1, I_1), \\
\quad \text{process}(HX, HY), \text{compose}(I_1, HY, I_2), \\
\quad \text{compose}(I_2, TY2, I_3), HY = I_3, \\
\quad r \uparrow \text{tupling}(TXs, TY), \text{compose}(HY, TY, Y) \\
\)

By unfolding clause 5 wrt \( r(TX1,TY1) \) using \( DCLR \), and simplifying using condition (4):

clause 7: \( r \uparrow \text{tupling}(Xs, Y) \leftarrow \\
\quad Xs = [Xs, TXs], \\
\quad \text{nonminimal}(X), \text{decompose}(X,HX,TX1,TX2), \\
\quad \text{minimal}(TX1), \text{minimal}(TX1), \text{solve}(TX1, TY1), \\
\quad r(TX2, TY2), \\
\quad I_0 = e, \text{compose}(I_0, TY1, I_1), \\
\quad \text{process}(HX, HY), \text{compose}(I_1, HY, I_2), \\
\quad \text{compose}(I_2, TY2, I_3), HY = I_3, \\
\quad r \uparrow \text{tupling}(TXs, TY), \text{compose}(HY, TY, Y) \\
\)

By deleting one of the minimal(TX1) atoms in clause 7:

clause 8: \( r \uparrow \text{tupling}(Xs, Y) \leftarrow \\
\quad Xs = [Xs, TXs], \\
\quad \text{nonminimal}(X), \text{decompose}(X,HX,TX1,TX2), \\
\quad \text{minimal}(TX1), \text{solve}(TX1, TY1), \\
\quad r(TX2, TY2), \\
\quad I_0 = e, \text{compose}(I_0, TY1, I_1), \\
\quad \text{process}(HX, HY), \text{compose}(I_1, HY, I_2), \\
\quad \text{compose}(I_2, TY2, I_3), HY = I_3, \\
\quad r \uparrow \text{tupling}(TXs, TY), \text{compose}(HY, TY, Y) \\
\)

By rewriting clause 8 using applicability condition (1):

13
clause 9: \[ r_{\mathcal{D}upling}(Xs, Y) \rightarrow \]
\[
Xs = [X[T Xs],
nonminimal(X), decompose(X, H X, T X_1, T X_2),
minimal(T X_1), solve(T X_1, T Y_1),
r(T X_2, T Y_2),
I_0 = e, compose(I_0, T Y_1, I_1),
process(H X, H HY), compose(I_1, H HY, I_2), HY = I_2,
r_{\mathcal{D}upling}(T Xs, TTY), compose(T Y_2, TTY, T Y),
compose(HY, T Y, Y) \]

By folding clause 9 using clauses 1 and 2:

clause 10: \[ r_{\mathcal{D}upling}(Xs, Y) \rightarrow \]
\[
Xs = [X[T Xs],
nonminimal(X), decompose(X, H X, T X_1, T X_2),
minimal(T X_1), solve(T X_1, T Y_1),
r_{\mathcal{D}upling}(T Xs, T Y),
I_0 = e, compose(I_0, T Y_1, I_1),
process(H X, H HY), compose(I_1, H HY, I_2), HY = I_2,
compose(I_2, T Y_2, I_3), HY = I_3,
r_{\mathcal{D}upling}(T Xs, T Y), compose(HY, T Y, Y) \]

By introducing atom minimal(U) (with new, i.e. existentially quantified, variable U) in clause 6:

clause 11: \[ r_{\mathcal{D}upling}(Xs, Y) \rightarrow \]
\[
Xs = [X[T Xs],
nonminimal(X), decompose(X, H X, T X_1, T X_2),
minimal(U), r(T X_1, T Y_1), r(T X_2, T Y_2),
I_0 = e, compose(I_0, T Y_1, I_1),
process(H X, H HY), compose(I_1, H HY, I_2),
compose(I_2, T Y_2, I_3), HY = I_3,
r_{\mathcal{D}upling}(T Xs, T Y), compose(HY, T Y, Y) \]

By using applicability condition (3):

clause 12: \[ r_{\mathcal{D}upling}(Xs, Y) \rightarrow \]
\[
Xs = [X[T Xs],
nonminimal(X), decompose(X, H X, T X_1, T X_2),
minimal(U), r(U, e), r(T X_1, T Y_1), r(T X_2, T Y_2),
I_0 = e, compose(I_0, T Y_1, I_1),
process(H X, H HY), compose(I_1, H HY, I_2),
compose(I_2, T Y_2, I_3), HY = I_3,
r_{\mathcal{D}upling}(T Xs, T Y), compose(HY, T Y, Y) \]

By using applicability condition (2):

clause 13: \[ r_{\mathcal{D}upling}(Xs, Y) \rightarrow \]
\[
Xs = [X[T Xs],
nonminimal(X), decompose(X, H X, T X_1, T X_2),
minimal(U), r(U, e), r(T X_1, T Y_1), r(T X_2, T Y_2),
I_0 = e, compose(I_0, T Y_1, I_1),
process(H X, H HY), compose(I_1, H HY, I_2),
compose(I_2, T Y_2, I_3), HY = I_3,
r_{\mathcal{D}upling}(T Xs, T Y), compose(HY, T Y, Y) \]

By using applicability conditions (1) and (2):
clause 14: $r_d\uplus(X_s, Y) =$
\[X_s = [X][T X_s],\]
nonminimal($X$), decompose($X, H X, T X_1, T X_2$),
nonminimal($T X_1$),
minimal($U$), $r(U, e)$,
$r(T X_1, TY_1), r(T X_2, TY_2)$,
\[l_0 = e, \text{ compose}(l_0, e, l_1),\]
process($H X, H H Y$), compose($l_1, H H Y, l_2$),
compose($l_2, TY_2, l_3$), $H Y = l_3$,
$r_d\uplus(X T X_s, TY)$, compose($H Y, TY, T I$),
compose($TY_1, T I, Y$)

By introducing a new, i.e. existentially quantified, variable $Y U$ in place of some occurrences of $e$.

clause 15: $r_d\uplus(X_s, Y) =$
\[X_s = [X][T X_s],\]
nonminimal($X$), decompose($X, H X, T X_1, T X_2$),
nonminimal($T X_1$),
minimal($U$), $r(U, Y U)$,
$r(T X_1, TY_1), r(T X_2, TY_2)$,
\[l_0 = e, \text{ compose}(l_0, Y U, l_1),\]
process($H X, H H Y$), compose($l_1, H H Y, l_2$),
compose($l_2, TY_2, l_3$), $H Y = l_3$,
$r_d\uplus(X T X_s, TY)$, compose($H Y, TY, T I$),
compose($TY_1, T I, Y$)

By introducing nonminimal($N$) and decompose($N, H X, U, T X_2$), since $\exists N: T X, \text{ nonminimal}(N) \land$ decompose($N, H X, U, T X_2$) always holds (because $N$ is existentially quantified):

clause 16: $r_d\uplus(X_s, Y) =$
\[X_s = [X][T X_s],\]
nonminimal($X$), decompose($X, H X, T X_1, T X_2$),
nonminimal($T X_1$),
minimal($U$), $r(U, Y U)$,
nonminimal($N$), decompose($N, H X, U, T X_2$),
$r(T X_1, TY_1), r(T X_2, TY_2)$,
\[l_0 = e, \text{ compose}(l_0, Y U, l_1),\]
process($H X, H H Y$), compose($l_1, H H Y, l_2$),
compose($l_2, TY_2, l_3$), $H Y = l_3$,
$r_d\uplus(X T X_s, TY)$, compose($H Y, TY, T I$),
compose($TY_1, T I, Y$)

By duplicating goal decompose($N, H X, U, T X_2$):

clause 17: $r_d\uplus(X_s, Y) =$
\[X_s = [X][T X_s],\]
nonminimal($X$), decompose($X, H X, T X_1, T X_2$),
nonminimal($T X_1$),
minimal($U$), $r(U, Y U)$,
nonminimal($N$), decompose($N, H X, U, T X_2$),
decompose($N, H X, U, T X_2$),
$r(T X_1, TY_1), r(T X_2, TY_2)$,
\[l_0 = e, \text{ compose}(l_0, Y U, l_1),\]
process($H X, H H Y$), compose($l_1, H H Y, l_2$),
compose($l_2, TY_2, l_3$), $H Y = l_3$,
$r_d\uplus(X T X_s, TY)$, compose($H Y, TY, T I$),
compose($TY_1, T I, Y$)

By folding clause 17 using DCLR:
By folding clause /1/8 using clauses /1 and /2/:

\[
\text{clause } 18: \quad r\text{\textunderscore}\text{tupling}(Xs, Y) \leftarrow \\
Xs = [X[TXs], \\
\text{nonminimal}(X), \text{decompose}(X, HX, TX1, TX2), \\
\text{nonminimal}(TX1), \\
\text{minimal}(U), \\
\text{decompose}(N, HX, U, TX2), \\
r(TX1, TY1), r(X, HY), \\
r\text{\textunderscore}\text{tupling}(TXs, TY), \text{compose}(HY, TY, TI), \\
\text{compose}(TY1, TI, Y)
\]

By folding clause 18 using clauses 1 and 2:

\[
\text{clause } 19: \quad r\text{\textunderscore}\text{tupling}(Xs, Y) \leftarrow \\
Xs = [X[TXs], \\
\text{nonminimal}(X), \text{decompose}(X, HX, TX1, TX2), \\
\text{nonminimal}(TX1), \\
\text{minimal}(U), \\
\text{decompose}(N, HX, U, TX2), \\
r(TX1, TY1), r\text{\textunderscore}\text{tupling}([N[TXs], TI), \\
\text{compose}(TY1, TI, Y)
\]

Clauses 1, 3, 10, and 20 are the clauses of \(P_{\text{\textunderscore}\text{tupling}}\). Therefore \(P_{\text{\textunderscore}\text{tupling}}\) is steadfast wrt \(S_{\text{\textunderscore}\text{tupling}}\) in \(S\).

To prove that \(P_r\) is steadfast wrt \(S_r\) in \(\{S_{\text{\textunderscore}\text{tupling}}\}\), we do a backward proof that we begin with \(P_r\) in \(TG\) and from which we try to obtain \(S_r\).

The procedure \(P_r\) for \(r\) in \(TG\) is:

\[
r(X, Y) \leftarrow r\text{\textunderscore}\text{tupling}([X], Y)
\]

By taking the ‘completion’:

\[
\forall X : TX, \forall Y : TY. \quad I_r(X) \Rightarrow [r(X, Y) \Leftrightarrow r\text{\textunderscore}\text{tupling}([X], Y)]
\]

By unfolding the ‘completion’ above wrt \(r\text{\textunderscore}\text{tupling}([X], Y)\) using \(S_{\text{\textunderscore}\text{tupling}}\):

\[
\forall X : TX, \forall Y : TY. \quad I_r(X) \Rightarrow [r(X, Y) \Leftrightarrow \exists Y1, l1 : TY. \quad O_r(X, Y1) \land l1 = Y1 \land Y = l1]
\]

By simplification:

\[
\forall X : TX, \forall Y : TY. \quad I_r(X) \Rightarrow [r(X, Y) \Leftrightarrow O_r(X, Y)]
\]

We obtain \(S_r\), so \(P_r\) is steadfast wrt \(S_r\) in \(\{S_{\text{\textunderscore}\text{tupling}}\}\).

Therefore, \(TG\) is also steadfast wrt \(S_r\) in \(S\).
C Proof of $DG_2$

As a reminder, we give again the generalization schema $DG_2$.

$DG_2 : \{\text{DCLR, DGRL, C}_{21}\}$ where

(1) $\text{compose}$ is associative

(2) $\text{compose}$ has $e$ as the right identity element, where $e$ appears in $\text{DCLR}$

and templates $\text{DCLR}$ and $\text{DGRL}$ are in Figure 9.

\[
\begin{align*}
\text{r}(X, Y) &\rightleftharpoons \\
\text{minimal}(X), \\
\text{solve}(X, Y) \\
\text{r}(X, Y) &\rightleftharpoons \\
\text{nonminimal}(X), \\
\text{decompose}(X, HX, TX_1, TX_2), \\
\text{r}(TX_1, TY_1), \text{r}(TX_2, TY_2), \\
I_0 = e, \text{compose}(I_0, TY_1, I_1), \\
\text{process}(HX, HY), \text{compose}(I_1, HY, I_2), \\
\text{compose}(I_2, TY_2, I_3), Y = I_3 \\
\text{r}_{\text{descending2}}(X, Y, c) &\rightleftharpoons \\
\text{minimal}(X), \\
\text{solve}(S, A, Y) \\
\text{r}_{\text{descending2}}(X, Y, A) &\rightleftharpoons \\
\text{nonminimal}(X), \\
\text{decompose}(X, HX, TX_1, TX_2), \\
A = A_3, \text{r}_{\text{descending2}}(TX_2, A_2, A_3), \\
\text{process}(HX, HY), \text{compose}(HY, A_2, A_1), \\
\text{r}_{\text{descending2}}(TX_1, A_0, A_1), \\
\text{compose}(e, A_0, Y)
\end{align*}
\]

Figure 9: Templates $\text{DCLR}$ and $\text{DGRL}$

The specification $S_r$ of predicate $r$ is:

\[
\forall X : TX, \forall Y : TY. \quad T_r(X) \iff [r(X, Y) \leftrightarrow \text{O}_r(X, Y)]
\]

The specification $S_{r_{\text{descending2}}}$ of predicate $r_{\text{descending2}}$ is:

\[
\forall X : TX, \forall Y, A : TY. \quad T_r(X) \iff [r_{\text{descending2}}(X, Y, A) \leftrightarrow \exists S : T_Y. \quad \text{O}_r(X, S) \land \text{O}_r(S, A, Y)]
\]

where $O_r$ is the output condition of $\text{compose}$ and $O_r$ is the output condition of $r$.

To prove the generalization schema $DG_2$, we have to prove that templates $\text{DCLR}$ and $\text{DGRL}$ are equivalent wrt $S_r$. By Definition 5, this holds iff the following two conditions hold:

(a) $\text{DCLR}$ is steadfast wrt $S_r$ in $\mathcal{S} = \{S_{\text{minimal}}, S_{\text{nonminimal}}, S_{\text{solve}}, S_{\text{decompose}}, S_{\text{compose}}\}$, where $S_{\text{minimal}}$, $S_{\text{nonminimal}}$, $S_{\text{solve}}$, $S_{\text{decompose}}$, $S_{\text{compose}}$ are the specifications of $\text{minimal}$, $\text{nonminimal}$, $\text{solve}$, $\text{decompose}$, $\text{compose}$, which are all the undefined predicate names appearing in $\text{DCLR}$.

(b) $\text{DGRL}$ is also steadfast wrt $S_r$ in $\mathcal{S}$.

In program transformation, we assume that the input program, here template $\text{DCLR}$, is steadfast wrt $S_r$ in $\mathcal{S}$, so condition (a) always holds.

To prove equivalence, we have to prove condition (b). We will use the property of steadfastness: $\text{DGRL}$ is steadfast wrt $S_r$ in $\mathcal{S}$ if $P_{r_{\text{descending2}}}$ is steadfast wrt $S_{r_{\text{descending2}}}$ in $\mathcal{S}$, where $P_{r_{\text{descending2}}}$ is the procedure for $r_{\text{descending2}}$, and $P_r$ is steadfast wrt $S_r$ in $\{S_{r_{\text{descending2}}}\}$, where $P_r$ is the procedure for $r$.

To prove that $P_{r_{\text{descending2}}}$ is steadfast wrt $S_{r_{\text{descending2}}}$ in $\mathcal{S}$, we do a constructive forward proof that we begin with $S_{r_{\text{descending2}}}$ and from which we try to obtain $P_{r_{\text{descending2}}}$.

By taking the ‘decompletion’ of $S_{r_{\text{descending2}}}$:

\[
\text{clause 1 : } r_{\text{descending2}}(X, Y, A) \rightleftharpoons r(X, S), \text{compose}(S, A, Y)
\]

By unfolding clause 1 wrt $(X, S)$ using $\text{DCLR}$, and using the assumption that $\text{DCLR}$ is steadfast wrt $S_r$ in $\mathcal{S}$.
clause 2: \[ r_{\text{descending}}(X, Y, A) \gets \]

\[ \text{minimal}(X), \]

\[ \text{solve}(X, S), \text{compose}(S, A, Y) \]

clause 3: \[ r_{\text{descending}}(X, Y, A) \gets \]

\[ \text{nonminimal}(X), \text{decompose}(X, HX, TX_1, TX_2), \]

\[ r(TX_1, TS_1), r(TX_2, TS_2), \]

\[ I_0 = e, \text{compose}(I_0, TS_1, I_1), \]

\[ \text{process}(HX, HS), \text{compose}(I_1, HS, I_2), \]

\[ \text{compose}(I_2, TS_2, I_3), S = I_3, \]

\[ \text{compose}(S, A, Y) \]

By using applicability condition (1) on clause 3:

clause 4: \[ r_{\text{descending}}(X, Y, A) \gets \]

\[ \text{nonminimal}(X), \text{decompose}(X, HX, TX_1, TX_2), \]

\[ r(TX_1, TS_1), r(TX_2, TS_2), \]

\[ \text{compose}(TS_2, A, A_2), \]

\[ \text{process}(HX, HY), \text{compose}(HY, A_2, A_1), \]

\[ r_{\text{descending}}(TX_1, A_0, A_1), \]

\[ \text{compose}(e, A_0, Y) \]

By twice folding clause 4 using clause 1:

clause 5: \[ r_{\text{descending}}(X, Y, A) \gets \]

\[ \text{nonminimal}(X), \text{decompose}(X, HX, TX_1, TX_2), \]

\[ r_{\text{descending}}(TX_2, A_2, A), \]

\[ \text{process}(HX, HY), \text{compose}(HY, A_2, A_1), \]

\[ r_{\text{descending}}(TX_1, A_0, A_1), \]

\[ \text{compose}(e, A_0, Y) \]

By introducing a new, i.e. existentially quantified, variable \( A_3 \):

clause 6: \[ r_{\text{descending}}(X, Y, A) \gets \]

\[ \text{nonminimal}(X), \text{decompose}(X, HX, TX_1, TX_2), \]

\[ A = A_3, r_{\text{descending}}(TX_2, A_2, A_3), \]

\[ \text{process}(HX, HY), \text{compose}(HY, A_2, A_1), \]

\[ r_{\text{descending}}(TX_1, A_0, A_1), \]

\[ \text{compose}(e, A_0, Y) \]

Clauses 2 and 6 are the clauses of \( P_{r_{\text{descending}}} \). Therefore \( P_{r_{\text{descending}}} \) is steadfast wrt \( S_{r_{\text{descending}}} \) in \( S \).

To prove that \( P_r \) is steadfast wrt \( S_r \) in \( \{ S_{r_{\text{descending}}} \} \), we do a backward proof that we begin with \( P_r \) in \( \text{DGRL} \) and from which we try to obtain \( S_r \).

The procedure \( P_r \) for \( r \) in \( \text{DGRL} \) is:

\[ r(X, Y) \gets r_{\text{descending}}(X, Y, e) \]

By taking the ‘completion’:

\[ \forall X: TX, \forall Y: TY. \quad I_r(X) \Rightarrow [r(X, Y) \iff r_{\text{descending}}(X, Y, e)] \]

By unfolding the ‘completion’ above wrt \( r_{\text{descending}}(X, Y, e) \) using \( S_{r_{\text{descending}}} \):

\[ \forall X: TX, \forall Y: TY. \quad I_r(X) \Rightarrow [r(X, Y) \iff \exists S: TY. O_r(X, S) \land O_e(S, e, Y)] \]

By using applicability condition (2):

\[ \forall X: TX, \forall Y: TY. \quad I_r(X) \Rightarrow [r(X, Y) \iff \exists S: TY. O_r(X, S) \land S = Y] \]

By simplification:

\[ \forall X: TX, \forall Y: TY. \quad I_r(X) \Rightarrow [r(X, Y) \iff O_r(X, Y)] \]

We obtain \( S_r \), so \( P_r \) is steadfast wrt \( S_r \) in \( \{ S_{r_{\text{descending}}} \} \).

Therefore, \( \text{DGRL} \) is also steadfast wrt \( S_r \) in \( S \).