

Efficient algebraic effect handlers for Prolog

AMR HANY SALEH and TOM SCHRIJVERS

KU Leuven, Belgium

Department of Computer Science

(e-mail: {ah.saleh,tom.schrijvers}@cs.kuleuven.be)

submitted 6 May 2016; revised 8 July 2016; accepted 22 August 2016

Abstract

Recent work has provided *delimited control* for Prolog to dynamically manipulate the program control-flow, and to implement a wide range of control-flow and dataflow effects on top of. Unfortunately, delimited control is a rather primitive language feature that is not easy to use.

As a remedy, this work introduces *algebraic effect handlers* for Prolog, as a high-level and structured way of defining new side-effects in a modular fashion. We illustrate the expressive power of the feature and provide an implementation by means of elaboration into the delimited control primitives.

The latter add a non-negligible performance overhead when used extensively. To address this issue, we present an optimised compilation approach that combines partial evaluation with dedicated rewrite rules. The rewrite rules are driven by a lightweight effect inference that analyses what effect operations may be called by a goal. We illustrate the effectiveness of this approach on a range of benchmarks.

KEYWORDS: delimited control, algebraic effect handlers, program transformation

1 Introduction

The work of Schrijvers *et al.* (2013) has introduced *delimited control* constructs for Prolog. Delimited control is a very powerful means to dynamically manipulate the control-flow of programs that was first explored in the setting of functional programming (Felleisen 1988; Danvy and Filinski 1990). Schrijvers *et al.* (2013) show its usefulness in Prolog to concisely define implicit state, DCGs and coroutines. More recently, Desouter *et al.* (2015) have shown that delimited control also concisely captures the control-flow manipulation of tabling.

Unfortunately, there are two prominent downsides to delimited control. Firstly, it is a rather primitive feature that has been likened to the imperative `goto`, which was labeled harmful for high-level programming by Dijkstra (1968). Secondly, the overhead of delimited control for encoding state-passing features is non-negligible. For example, the delimited control implementation of DCGs is 10 times slower for a tight loop than the traditional implementation.

This paper addresses both issues. In order to provide a high-level structured interface to delimited control, we adapt the *algebraic effects and handlers* approach

to Prolog. Algebraic effects and handlers have been said to relate to delimited control the way structured loops relate to `goto`. While the structured approach restricts the expressive power, we still show a range of useful applications. Moreover, in exchange for the restricted expressiveness, we provide two benefits. Firstly, multiple handlers can be combined effortlessly to deal with distinct effects, to deal with one effect in terms of another or to customize the behavior of an effect. Secondly, we provide an automated program transformation that eliminates much of the overhead of delimited control. Indeed, compared to the free form of delimited control, the structured approach of effect handlers simplifies the identification of program patterns that can be optimised.

Our specific contributions are as follows:

- We define the syntax of algebraic effects and handlers for Prolog, and provide semantics in terms of an elaboration to the delimited control primitives.
- We illustrate the feature with a number of examples, including DCGs, implicit state and a writer effect.
- We provide a program transformation to eliminate much of the overhead of the elaboration-based implementation. This transformation is formulated in terms of partial evaluation augmented with rewrite rules. These rewrite rules are driven by an effect analysis that characterises which effects may be generated by a goal.
- We have implemented our program analysis and illustrated its effectiveness on a range of benchmarks.

All code of this paper is available at <http://github.com/ah-saleh/prologhandlers>.

2 Algebraic Effect Handlers

This section introduces our algebraic effect handlers for Prolog.

2.1 Syntax and Informal Semantics

We introduce two new syntactic constructs. The *effect operations* are Prolog predicate symbols op/n that are declared as such with the following syntax.

`:- effect op/n.`

For instance, we declare operation $c/1$ to consume a token, $get/1$ and $put/1$ to respectively retrieve and overwrite an implicit state, and $out/1$ to output a term.

The *handler* is a new Prolog goal form that specifies how to interpret effect operations. Its syntax is as follows:

$$\begin{array}{l} \text{handle } G_0 \text{ with} \\ \quad op_1(\bar{X}) \rightarrow G_1; \\ \quad \dots \\ \quad op_n(\bar{X}) \rightarrow G_m \\ \quad [\text{finally}(G_f)] \\ [\text{for}(P_1 = T_1, \dots, P_n = T_n)] \end{array}$$

Effect handlers can be thought of as a generalisation of exception handlers, where calling an effect operation corresponds to throwing an exception. The handler “catches” the operations that arise in G_0 . Its operation clauses $op_i(\bar{X}) \rightarrow G_i$ stipulate that an occurrence of operation $op_i(\bar{X})$ is to be handled by the goal G_i .

Before we explain the optional `finally` and `for` clauses, consider a few ways in which the `out/1` operation can be handled in `hw/0`.

```
hw :- out(hello), out(world).
```

In terms of the exception analogy, `hw/0` throws two `out/1` exceptions. Our first handler intercepts the first `out/1` and does nothing.

```
?- handle hw with (out(X) -> true).
true.
```

A more interesting handler prints the argument of `out/1`.

```
?- handle hw with (out(X) -> writeln(X)).
hello
true.
```

Note that only the first `out/1` is handled; this aborts the remainder of `hw` and the second `out/1` is never reached. To handle all operations, effect handlers support a feature akin to *resumable* exceptions: in the lexical scope of G_i , we can call `continue` to resume the part of the computation after the effect operation (i.e., its continuation). For instance, the next handler resumes the computation after handling the first `out/1` operation and intercepts later `out/1` operations in the same way.

```
?- handle hw with (out(X) -> writeln(X), continue).
hello
world
true.
```

Interestingly, we can invoke the same continuation multiple times, for instance both before and after printing the term.

```
?- handle hw with (out(X) -> continue, writeln(X), continue).
world
hello
world
true.
```

The finally Clause The optional `finally` clause is performed when G_0 finishes; if omitted, G_f defaults to `true`.

```
?- handle hw with (out(X) -> writeln(X), continue)
    finally (writeln(done)).
hello
world
done
true.
```

Note that if the goal does not run to completion, the `finally` clause is not invoked.

```

?- handle hw with (out(X) -> writeln(X))
    finally (writeln(done)).
hello
true.

```

The for Clause All variables in the operation and `finally` clauses are local to that clause, except if they are declared in the `for` clause.¹ Every *Var = Term* pair in the `for` clause relates a variable, which we call a *parameter*, that is in scope of all the operations and `finally` clauses with a term whose variables are in scope in the handler context. For instance, the following handler collects all outputs in a list.

```

?- handle hw with (out(X) -> Lin = [X|Lmid],
                  continue(Lmid,Lout))
    finally (Lin=Lout)
    for (Lin = List, Lout=[]).
List = [hello,world].

```

Note that `continue` has one argument for each parameter to indicate which values the parameters take in the continuation.

2.2 Nested Handlers and Forwarding

Nesting algebraic effect handlers is similar to nesting exception handlers. If an operation is not “caught” by the inner handler, it is *forwarded* to the outer handler. Moreover, if the inner handler catches an operation and, in the process of handling it, raises another operation, then this operation is handled by the outer handler. Let us illustrate both scenarios.

We can easily define a non-deterministic choice operator `or/2` in the style of Tor (Schrijvers *et al.* 2014; Schrijvers *et al.* 2014) in terms of the primitive `choice/1` effect which returns either of the two boolean values `t` and `f`.

```

:- effect choice/1.

or(G1,G2) :- choice(B), (B == t -> G1 ; B == f -> G2).

chooseAny(G) :- handle G with (choice(B) -> (B = t ; B = f),
                              continue).

```

The `chooseAny` handler interprets `choice/1` in terms of Prolog’s built-in disjunction `(;)/2`.

```

?- chooseAny(or(X = 1, X = 2)).
X = 1;
X = 2.

```

¹ The `for` clause plays a similar role as that in Schimpf’s *logical loops* (Schimpf 2002).

To obtain more interesting behavior, we can nest this handler with:

```
flip(G) :- handle G with (choice(B) -> choice(B1), not(B1,B),
                        continue).

not(t,f). not(f,t).
```

to flip the branches in a goal without touching the goal's code.

```
?- chooseAny(flip(or(X = 1, X = 2))).
X = 2;
X = 1.
```

What happens is that the inner `flip` handler intercepts the `choice(B)` call of `or/2`. It produces a new `choice(B1)` call that reaches the outer `chooseAny` handler, and unifies `B` with the negation of `B1`, which affects the choice in the continuation of `or/2`.

Thanks to forwarding, we can also easily mix different effects. For instance, with:

```
writeOut(G) :- handle G with (out(T) -> writeln(T), continue).
```

we can combine output and non-determinism.

```
?- chooseAny(writeOut(or(out(hello), out(world)))), fail.
hello
world
false.
```

Note that the inner `writeOut` handler does not know how to interpret the `choice/1` effect. As a consequence, it (implicitly) *forwards* this operation to the next surrounding handler, `chooseAny`, who does know what to do.

2.3 Elaboration Semantics

There is a straightforward elaboration of handlers into the `shift/1` and `reset/3` delimited control primitives for Prolog (Schrijvers *et al.* 2013). For instance, the last example query of Section 2.1 is elaborated into:

```
?- handler42(hw,List, []).
List = [hello,world].
```

where the declaration of the `out/1` operation is elaborated into:

```
out(X) :- shift(out(X)).
```

which shifts the term representation of the operation. The actual handler code is elaborated into a predicate (with a *fresh* name).

```
handler42(Goal, Lin, Lout) :-
  reset(Goal, Cont, Signal),
  ( Signal == 0 ->
    Lin = Lout
```

```

; Signal = out(X) ->
  Lin = [X|Lmid],
  handler42(Cont,Lmid,Lout)
; shift(Signal),
  handler42(Cont,Lin,Lout)
).

```

This predicate executes the goal in the delimited scope of a `reset/3`, which captures any `shift/1` call. If the goal terminates normally (i.e., `Signal=0`), then the finally code is run. If the goal suspends with a `shift/1`, the predicate checks whether the operation matches the handler's operation clause. If so, the clause's body is run. Note that `continue(Lmid,Lout)` has been expanded into a recursive invocation of the handler with the actual continuation goal `Cont`. If the operation does not match, the handler forwards it to the nearest surrounding handler with `shift/1` and continues with the continuation.

The example above generalizes straightforwardly. Any declaration of an effect operation is elaborated into a predicate definition.

```

:- effect op/n.   ↦   op(X1,...,Xn) :- shift(op(X1,...,Xn)).

```

Also, every handler goal is substituted with a predicate call.

```

handle G0 with
  op1 → G1;
  ...
  opn → Gm
  finally(Gf)
for(P1 = T1, ..., Pn = Tn)
    ↦ h(G0, T1, ..., Tn).

```

where $h/n + 1$ is an auxiliary predicate defined as:

```

h(Goal, P1, ..., Pn) :-
  reset(Goal, Cont, Signal),
  ( Signal == 0 -> Gf
  ; Signal = op1 -> G'1
  ; ...
  ; Signal = opn -> G'n
  ; shift(Signal), h(Cont, P1, ..., Pn)
  ).

```

Here, each G'_i is derived from G_i by replacing all occurrences of `continue(S1, ..., Sn)` with recursive calls `h(Cont, S1, ..., Sn)`.

3 Optimisation

Section 2.3's elaboration of algebraic effects into the delimited control constructs is conveniently straightforward. Unfortunately, capturing the delimited continuation incurs a non-trivial runtime cost. In many simple cases this cost is quite steep compared to more conventional program transformation approaches. For instance,

the implementation of DCGs with delimited control is 10 times slower in a tight loop than the traditional term expansion approach.

Fortunately, the runtime overhead is not inherent in the algebraic effects and handlers approach, and we can obtain competitive performance through optimised compilation. This section presents our optimisation approach, which aims to eliminate most uses of delimited control. The optimisation consists of two collaborating transformation approaches: rewrite rules (Section 3.2) and partial evaluation (Section 3.3). We use term rewrite rules to simplify handler goals and possibly eliminate the handler construct altogether. These rules depend on an effect system (Section 3.1) that infers which effects can or cannot be generated by a goal. Partial evaluation complements the rewrite rules by specialising handled predicate calls. This enables in particular the specialisation of (mutually) recursive predicates.

3.1 Effect System

Driving our optimisation is an *effect system* that associates with each goal G an *effect set* E that denotes which effects the goal may call.

Effect Sets In order to cater for modular programs, effect sets E are not elements of the powerset lattice over the closed set OP of locally known effect operation symbols op/n . Instead, we use the powerset lattice over an open-ended set of effect operations augmented with the additional top element All . This allows us to express the effects of unknown goals and unknown effect operations in an abstract manner.

Hence, we denote effect sets in one of two forms: $\bigcup_i op_i/n_i$ or $All - \bigcup_i op_i/n_i$. The former is an explicit enumeration of effect operations, while the latter expresses the dual: all but the given effect operations.

The \in relation as well as the functions \cup and $-$ are extended in the obvious way from the powerset lattice to our augmented version.

Effect System We use these functions over effect sets in the definition of our effect system judgment $E_c \vdash G : E$. This judgment expresses that goal G calls only effect operations from the effect set E , provided that `continue` calls only effect operations from the effect set E_c . Since `continue` is not defined for a top-level goal, we may assume any value for its E_c . Hence, for convenience, we always take $E_c = \emptyset$ for top-level goals G and just write $\vdash G : E$.

Figure 1 defines this judgment by means of inference rules. Rule (E-VAR) expresses that a variable (i.e., unknown) goal may call All effect operations. Rule (E-OP) states that known effect operation calls itself. Rules (E-CONJ) and (E-DISJ) combine the effects of their subgoals. Rule (E-TRUE) expresses that the goal `true`, as an example for other built-ins, is `op-free`. Rule (E-CONT) captures the invariant that `continue` has the E_c effect. In Rule (E-PRED) the effect of a user-defined predicate is the effect of its body. Finally, the most of the complexity of the inference system is concentrated in Rule (E-HANDLE) that deals with a handler goal. The rule expresses that the handler goal forwards all the effect operations E_0 of the goal G_0 it handles, except for the ones that the handler takes care of, $\bigcup_i op_i/n_i$. In addition, the handler may introduce additional calls to effect operations in its operation and finally

$$\begin{array}{c}
 E_c \vdash X : All \quad (E\text{-VAR}) \qquad \frac{op/n \in OP}{E_c \vdash op(T_1, \dots, T_n) : op/n} \quad (E\text{-OP}) \\
 \\
 \frac{E_c \vdash G_1 : E_1 \quad E_c \vdash G_2 : E_2}{E_c \vdash (G_1, G_2) : E_1 \cup E_2} \quad (E\text{-CONJ}) \qquad \frac{E_c \vdash G_1 : E_1 \quad E_c \vdash G_2 : E_2}{E_c \vdash (G_1; G_2) : E_1 \cup E_2} \quad (E\text{-DISJ}) \\
 \\
 E_c \vdash \text{continue}(\bar{T}) : E_c \quad (E\text{-CONT}) \qquad E_c \vdash \text{true} : \emptyset \quad (E\text{-TRUE}) \\
 \\
 \frac{p(S_1, \dots, S_n) :- G \quad E_c \vdash G : E}{E_c \vdash p(T_1, \dots, T_n) : E} \quad (E\text{-PRED}) \\
 \\
 \frac{E^* = (E_0 - \bigcup_i op_i) \cup E_f \cup \bigcup_i E_i \quad E_c \vdash G_0 : E_0 \quad E_c \vdash G_f : E_f \quad E^* \vdash G_i : E_i \ (\forall i)}{E_c \vdash \left(\begin{array}{l} \text{handle } G_0 \text{ with} \\ \frac{op_i(\bar{X}) \rightarrow G_i}{\text{finally}(G_f) \text{ for}(G_s)} \end{array} \right) : E^*} \quad (E\text{-HANDLE})
 \end{array}$$

Fig. 1. Effect Inference Rules

clauses. Also, note that calls to `continue` in the operation clauses have exactly the same effect as the handler goal itself; they are essentially recursive calls after all.

Here are a few examples:

```

⊢ hw : out/1
⊢ handle hw with (out(X) -> writeln(X)) : ∅
⊢ handle Y with (out(X) -> writeln(X)) : All - out/1
    
```

3.2 Rewrite Rules

We use the information of the effect system and the syntactic structure of goals to perform a number of handler-specific optimisations. We denote these optimisations in terms of semantics-preserving equivalences $G_1 \equiv G_2$ that we use as left-to-right rewrite rules. Figure 2 lists our rewrite rules in the form of inference rules where conditions on the inferred effects are written above the bar.

Rule (O-DISJ) captures the fact that effect handling is orthogonal to disjunction to specialise the branches of a disjunction separately. There are two rules for conjunction. Rule (O-CONJ) pulls the first goal G_1 of a conjunct out of the handler provided that it does not call any of the handler’s operations. This covers both the case were G_1 is an op-free goal and the case where all the operations in G_1 are forwarded by the handler. The second rule for conjunction, Rule (O-OP), statically evaluates the special case where the first goal is an operation dealt with by the handler. This consists of three parts: 1) the unification of the formal and actual parameters, 2) the unification of the formal and actual operation arguments, and 3) calling the operation clause’s goal. Note that we substitute all calls to `continue`(\bar{U}) (for any \bar{U}) in this last goal with the second conjunct wrapped in the handler; note

$$\begin{array}{c}
 \left(\begin{array}{c} \text{handle } (G_1; G_2) \\ \text{with } \overline{op \rightarrow G}; \\ \text{finally}(G_f) \text{ for}(G_s) \end{array} \right) \equiv \left(\begin{array}{c} \text{handle } G_1 \\ \text{with } \overline{op \rightarrow G}; \\ \text{finally}(G_f) \text{ for}(G_s) \end{array} \right); \left(\begin{array}{c} \text{handle } G_2 \\ \text{with } \overline{op \rightarrow G}; \\ \text{finally}(G_f) \text{ for}(G_s) \end{array} \right) \\
 \text{(O-DISJ)} \\
 \\
 \frac{E_c \vdash G_1 : E_1 \quad E_1 \cap \bigcup_i \text{op}_i = \emptyset}{\left(\begin{array}{c} \text{handle } (G_1, G_2) \text{ with} \\ \overline{op \rightarrow G}; \\ \text{finally}(G_f) \text{ for}(G_s) \end{array} \right) \equiv G_1, \left(\begin{array}{c} \text{handle } G_2 \text{ with} \\ \overline{op \rightarrow G}; \\ \text{finally}(G_f) \text{ for}(G_s) \end{array} \right)} \text{(O-CONJ)} \\
 \\
 \frac{(\overline{op(\bar{S}) \rightarrow G_i}) \in \overline{op \rightarrow G} \quad \text{freshen}(\bar{P}_F, \bar{S}, G_i) = (\bar{P}'_F, \bar{S}', G'_i)}{\left(\begin{array}{c} \text{handle } (op(\bar{T}), G_c) \\ \text{with } \overline{op \rightarrow G} \\ \text{finally}(G_f) \\ \text{for}(\bar{P}_F = \bar{P}_A) \end{array} \right) \equiv \bar{P}'_F = \bar{P}_A, \bar{T} = \bar{S}', G'_i \left[\begin{array}{c} \text{handle } G_c \text{ with} \\ \text{with } \overline{op \rightarrow G} \\ \text{finally}(G_f) \\ \text{for}(\bar{P}_F = \bar{U}) \end{array} \right]} \text{(O-OP)} \\
 \\
 \frac{E_c \vdash G : E \quad \overline{op' \rightarrow G'} = (\overline{op \rightarrow G}) \cap E}{\left(\begin{array}{c} \text{handle } G \\ \text{with } \overline{op \rightarrow G} \\ \text{finally}(G_f) \text{ for}(G_s) \end{array} \right) \equiv \left(\begin{array}{c} \text{handle } G \\ \text{with } \overline{op' \rightarrow G'} \\ \text{finally}(G_f) \text{ for}(G_s) \end{array} \right)} \text{(O-DROP)} \\
 \\
 \left(\begin{array}{c} \text{handle}(G) \text{ with } \emptyset \\ \text{finally}(G_f) \text{ for}(G_s) \end{array} \right) \equiv G, G_s, G_f \quad \text{(O-TRIV)} \\
 \\
 \frac{\text{see text} \quad \overline{op'_2 \rightarrow G'_2} = (\overline{op_2 \rightarrow G_2}) - \overline{op_1}}{\left(\begin{array}{c} \text{handle} \left(\begin{array}{c} \text{handle } G \\ \text{with } \overline{op_1 \rightarrow G_1} \\ \text{finally}(G_{1,f}) \\ \text{for}(G_{1,s}) \end{array} \right) \\ \text{with } \overline{op_2 \rightarrow G_2} \\ \text{finally}(G_{2,f}) \\ \text{for}(G_{2,s}) \end{array} \right) \equiv \left(\begin{array}{c} \text{handle } G \text{ with} \\ \overline{op_1 \rightarrow G'_1} \\ \overline{op'_2 \rightarrow G'_2} \\ \text{finally } G'_{1,f} \\ \text{for}(G_{1,s}, G_{2,s}) \end{array} \right)} \text{(O-MERGE)}
 \end{array}$$

Fig. 2. Optimisation Rules for effect handlers

that the arguments \bar{U} become the new actual parameters. In the process we are careful to *freshen* all the local logical variables that are used.

Rule (O-DROP) removes spurious operation clauses from the handler; it only retains those that correspond to operations that the goal may actually call. In the case that no operation clauses remain, Rule (O-TRIV) dispenses with the handler altogether. This amounts to unifying the formal and actual parameters and calling the finally goal.

Finally, the most complex rule of all, Rule (O-MERGE), merges two nested handlers into one single handler and thereby eliminates expensive forwarding of operations. At first, it might seem trivial to merge two handlers: We simply merge all the components of the two handlers pairwise. There is an obvious simplification to perform in the process: we can drop all outer handler’s operation clauses that overlap

with any of the inner handler’s clauses, as the inner handler takes precedence over the outer one.

Yet, there is a further subtle issue that have to be taken account in order to preserve the original semantics. The finally goal $G_{1,f}$ and the operation clause goals $\overline{op_1}$ may call operations that are originally intercepted by the outer handler. We have to make sure that this remains the case. For that reason we adjust those goals to $G'_{1,f}$ and \bar{G}'_1 in the merged handler. Let us explain these adjustments for the different forms of operation clause goals $G_{1,i}$ that we consider.

1. The operation goal $G_{1,i}$ is of the form $G_{1,i,a}, \text{continue}(\bar{V})$ where $G_{1,i,a}$ does not contain any call to `continue`. We wrap the initial part of the goal in the outer handler and finally proceed with `continue`.

$$G'_{1,i} = \left(\begin{array}{c} \text{handle } G_{1,i,a} \text{ with} \\ \overline{op_2 \rightarrow G_2} \\ \text{finally } \text{continue}(\bar{V}, \bar{P}_{2,F}) \\ \text{for } (\bar{P}_{2,F}, \bar{P}'_{2,F}) \end{array} \right)$$

2. The operation goal $G_{1,i}$ does not contain a call to `continue`. In this case we wrap the entire goal in the outer handler and make sure to call the outer handler’s final goal.

$$G'_{1,i} = \left(\begin{array}{c} \text{handle } G_{1,i} \text{ with} \\ \overline{op_2 \rightarrow G_2} \\ \text{finally } G_{2,f} \\ \text{for } (\bar{P}_{2,F}, \bar{P}'_{2,F}) \end{array} \right)$$

Similarly, we adapt the final goal $G_{1,f}$ to

$$G'_{1,f} = \left(\begin{array}{c} \text{handle } G_{1,f} \text{ with} \\ \overline{op_2 \rightarrow G_2} \\ \text{finally } G_{2,f} \\ \text{for } (\bar{P}_{2,F}, \bar{P}'_{2,F}) \end{array} \right)$$

3.3 Partial Evaluation

We use a custom partial evaluation approach to expose more optimisation opportunities for the rewrite rules and to deal with recursive predicates. Our partial evaluation is targeted at predicate calls that are handled. Consider the following simple DCG example that checks if a phrase is a succession of the terminals `ab`:

```
:- effect c/1.
ab.
ab :- c(a), c(b), ab.
query(Lin) :-
    handle ab with
        (c(X) -> Lin1=[X|Lmid], continue(Lmid,Lout1))
    finally (Lin1 = Lout1)
    for (Lin1=Lin,Lout1=[]).
```

Here we abstract the goal `handle ab with ...` into a fresh predicate (say `ab0/2`), which makes abstraction of the actual handler parameters. This yields the new definition of `query/1`:

```
query(Lin) :- ab0(Lin, []).
```

At the same time we unfold the definition of `ab/0` in the newly created predicate `ab0/2`. Because `ab/0` has two clauses, this means that `ab0/2` bifurcates similarly.

```
ab0(Lin,Lout) :-                               ab0(Lin,Lout) :-
  handle true with                               handle (c(a), c(b), ab) with
  (c(X) -> Lin1=[X|Lmid],                         (c(X) -> Lin1=[X|Lmid],
    continue(Lmid,Lout1))                          continue(Lmid,Lout1))
  finally (Lin1 = Lout1)                          finally (Lin1 = Lout1)
  for (Lin1=Lin,Lout1=Lout).                       for (Lin1=Lin,Lout1=Lout).
```

This unfolding exposes new rewriting opportunities. Using the Rules (O-DROP) and (O-TRIV), the first clause specialises to `Lin1=Lin`, `Lout1=Lout`, `Lin1=Lout1`. In the second clause, a double use of Rule (O-OP) deals with the `c/1` operations. This leaves a recursive invocation of `ab/0`, wrapped in the handler. Now the partial evaluation kicks in again, realises that this is a variant of the earlier specialization and ties the knot with a recursive call to `ab0/2`. After further clean-up of the unifications, we get:

```
ab0(L,L).
ab0([a,b|Lmid],Lout) :- ab0(Lmid,Lout).
```

There is no trace of delimited control left. Moreover, this is precisely the tight code that the traditional DCG yields.

4 Evaluation

We evaluate the usefulness of our optimisation approach experimentally on a set of benchmarks. All results were obtained on an Intel Core i7 with 8 GB RAM running hProlog 3.2.38 on Ubuntu 14.04.

The first experiment concerns the `ab` program of Section 3.3. Table 1 lists the timings (in ms) for different input sizes obtained with three different versions of the program: the traditional DCG implementation (based on SICStus), the elaborated handler implementation and the optimised handler implementation. Clearly, the naive use of delimited control slows the program down by more than an order of magnitude. Fortunately, our optimisation eliminates all uses of delimited control and matches the traditional implementation's performance.²

The second experiment considers three scenarios with nested handlers. Table 2 lists the runtime results (in ms) for different input sizes of different versions: the plain elaborated program, the program optimised with only the rewrite rules and the program optimised with both rewrite rules and partial evaluation.

² Thanks to more aggressive inlining it is even slightly faster.

Table 1. DCG benchmark results in ms

Input Size	Traditional	Elaboration	Optimised
1×10^3	0	2	0
1×10^4	1	4	1
1×10^5	8	37	5
1×10^6	32	321	29
2×10^6	67	635	58
5×10^6	150	1821	146
1×10^7	300	4757	297
1×10^8	2953	47632	2922

The first benchmark, `state_dcg`, extends the `ab` example with an implicit state that is incremented with every occurrence of `ab` in the input. Because the rewrite rules merge the two handlers in this benchmark, they generate an almost two-fold speed-up. With partial evaluation, the speed-up of around two orders of magnitude is much more dramatic. The main reason is that delimited control is again eliminated.

The second benchmark adds an inner-most dummy handler for an unused `foo/0`. The aim of this benchmark is to assess the cost of forwarding. In the plain elaborated version, we can see there is a significant overhead. Thanks to the rewriting, the three handlers are again merged and most of the overhead of the spurious handler disappears – the only remaining cost is the spurious `foo/0` operation clause. Finally, with partial evaluation, all trace of the `foo/0` is eliminated.

The third benchmark re-implements the calculator example of Dragan *et al.* (2009) with two handlers, one to manage an implicit stack and one to one for an implicit register. The behavior is similar to the other two benchmarks: merging the handlers roughly halves the runtime and partially evaluating them speeds up the code by two orders of magnitude.

5 Related Work

Language Extensions Various Prolog language extensions have been proposed in terms of program transformations. Van Roy has proposed *Extended DCGs* (Roy 1989) to thread multiple named accumulators. Similarly, Ciao Prolog's structured state threading (Ivanovic *et al.* 2009) enables different implicit states. Algebraic effects and handlers can easily provide similar functionality.

Schimpf's *logical loops* (Schimpf 2002) approach has been very influential on our handler design, in particular regarding the elaboration into recursive predicates and the notions of locally fresh variables and parameters. Of course, both features originate in distinct paradigms: logical loops are inspired by imperative loops, while handlers originate in the functional programming paradigm.

Control Primitives Various works have considered extensions of Prolog that enable control-flow manipulation. Before the work of Schrijvers *et al.* (Schrijvers *et al.* 2013), Tarau and Dahl already allowed the users of BinProlog to access and manipulate the program's continuation.

Table 2. Runtimes of nested-handler benchmarks in ms

Program Name	Input Size	Elaborated	Rewriting	Rewriting + PE
<i>state_dcg</i>	1×10^3	3	2	0
<i>state_dcg</i>	1×10^4	20	11	0
<i>state_dcg</i>	1×10^5	151	63	3
<i>state_dcg</i>	1×10^6	1879	604	37
<i>state_dcg</i>	2×10^6	2814	1208	75
<i>state_dcg</i>	5×10^6	7919	4348	186
<i>state_dcg</i>	1×10^7	29695	18094	375
<i>state_dcg_foo</i>	1×10^3	4	3	0
<i>state_dcg_foo</i>	1×10^4	23	11	0
<i>state_dcg_foo</i>	1×10^5	358	61	3
<i>state_dcg_foo</i>	1×10^6	4666	670	37
<i>state_dcg_foo</i>	2×10^6	8777	1350	75
<i>state_dcg_foo</i>	5×10^6	30026	4551	186
<i>calculator</i>	1×10^3	4	3	1
<i>calculator</i>	1×10^4	30	16	1
<i>calculator</i>	1×10^5	307	78	10
<i>calculator</i>	1×10^6	1195	761	57
<i>calculator</i>	2×10^6	3015	1525	110
<i>calculator</i>	5×10^6	12326	6114	247

Various coroutine-like features have been proposed in the context of Prolog for implementing alternative execution mechanisms, such as constraint logic programming and delay. Nowadays most of these are based on a single primitive concept: attributed variables (Holzbaur 1992). Like delimited control, attributed variables are a very low-level feature that is meant to be used directly, but is often used by library writers as the target for much higher-level declarative features.

Algebraic Effects and Handlers The work in this paper adapts the existing work in the functional programming community on algebraic effects and handlers to Prolog. Both algebraic effects (Plotkin and Power 2002) and handlers (Plotkin and Matija 2013) were first explored at a theoretical level, before giving rise to a whole range of implementations in functional programming languages, such as Eff (Bauer and Pretnar 2015), Multicore OCaml (Dolan *et al.* 2015) and Haskell (Kammar *et al.* 2013; Kiselyov *et al.* 2013; Wu and Schrijvers 2015) to name a few.

Schrijvers *et al.* (2014) have previously appealed to a functional model of algebraic effects and handlers to derive a Prolog implementation of search heuristics (2014). This paper enables a direct Prolog implementation that avoids this detour.

6 Conclusion

This paper has defined algebraic effects and handlers for Prolog as a high-level alternative to delimited control for implementing custom control-flow and dataflow effects. In order to avoid undue runtime overhead of capturing delimited continuations, we provide an optimised compilation approach based on partial

evaluation and rewrite rules. Our experimental evaluation shows that this approach greatly reduces the runtime overhead.

Acknowledgments

This work is partly funded by the Flemish Fund for Scientific Research (FWO). We are grateful to Bart Demoen for his support of hProlog.

Supplementary material

For supplementary material for this article, please visit <http://dx.doi.org/10.1017/S147106841600034X>

References

- BAUER, A. AND PRETNAR, M. 2015. Programming with algebraic effects and handlers. *Journal of Logical and Algebraic Methods in Programming* 84, 1, 108–123.
- DANVY, O. AND FILINSKI, A. 1990. Abstracting control. LFP '90. 151–160.
- DESOUTER, B., VAN DOOREN, M. AND SCHRIJVERS, T. 2015. Tabling as a library with delimited control. *Theory and Practice of Logic Programming* 15, 4-5, 419–433.
- DIJKSTRA, E. W. 1968. Letters to the editor: Go to statement considered harmful. *Commun. ACM* 11, 3 (Mar.), 147–148.
- DOLAN, S., WHITE, L., SIVARAMAKRISHNAN, K., YALLOP, J. AND MADHAVAPEDDY, A. 2015. Effective concurrency through algebraic effects. In *OCaml Users and Developers Workshop*.
- FELLEISEN, M. 1988. The theory and practice of first-class prompts. POPL '88. 180–190.
- HOLZBAUR, C. 1992. Meta-structures vs. Attributed Variables in the context of extensible unification. In *Proceedings of the 4th International Symposium on Programming Language Implementation and Logic Programming*, LNCS, vol. 631. 260–268.
- IVANOVIC, D., MORALES CABALLERO, J. F., CARRO, M. AND HERMENEGILDO, M. 2009. Towards structured state threading in Prolog. In *CICLOPS 2009*.
- KAMMAR, O., LINDLEY, S. AND OURY, N. 2013. Handlers in action. In *Proceedings of the 18th ACM SIGPLAN International Conference on Functional programming*. ICFP '14. ACM, 145–158.
- KISELYOV, O., SABRY, A. AND SWORDS, C. 2013. Extensible effects: An alternative to monad transformers. In *Proceedings of the 2013 ACM SIGPLAN Symposium on Haskell*. Haskell '13. ACM, New York, NY, USA, 59–70.
- PLOTKIN, G. D. AND MATIJA, P. 2013. Handling algebraic effects. *Logical Methods in Computer Science* 9, 4.
- PLOTKIN, G. D. AND POWER, J. 2002. Notions of computation determine monads. In *Foundations of Software Science and Computation Structures*, M. Nielsen and U. Engberg, Eds. LNCS, vol. 2303. Springer, 342–356.
- ROY, P. V. 1989. A useful extension to Prolog's definite clause grammar notation. In *ACM SIGPLAN Notices*, Vol. 24, 11, Nov. 1989, 132–134.
- SCHIMPF, J. 2002. Logical loops. In *International Conference on Logic Programming*. Springer, 224–238.
- SCHRIJVERS, T., DEMOEN, B., DESOUTER, B. AND WIELEMAKER, J. 2013. Delimited continuations for Prolog. *Theory and Practice of Logic Programming* 13, 4-5, 533–546.

- SCHRIJVERS, T., DEMOEN, B., TRISKA, M. AND DESOUTER, B. 2014. Tor: Modular search with hookable disjunction. *Sci. Comput. Program.* 84, 101–120.
- SCHRIJVERS, T., WU, N., DESOUTER, B. AND DEMOEN, B. 2014. Heuristics entwined with handlers combined. In *Proceedings of the 16th International Symposium on Principles and Practice of Declarative Programming*, PPDP '14. ACM, 259–270.
- WU, N. AND SCHRIJVERS, T. 2015. Fusion for free - efficient algebraic effect handlers. In *Mathematics of Program Construction*, LNCS, vol. 9129. Springer, 302–322.