Lecture 3: Algebraic Effects II

Gordon Plotkin

Laboratory for the Foundations of Computer Science, School of Informatics, University of Edinburgh

NII Shonan Meeting No. 146
Programming and Reasoning with Algebraic Effects and Effect Handlers
Outline

1. Algebra with parameterised operations
2. Algebra with parameters and parametric arguments
3. Algebraic operations and generic effects (cntnd.)
Plan of Lecture

Ordinary algebra → Alg. with par. ops

Alg. with ops. with abstraction → Alg. with par. ops with abstraction
Outline

1. Algebra with parameterised operations
2. Algebra with parameters and parametric arguments
3. Algebraic operations and generic effects (cntnd.)
Parametric finitary equational theories: syntax

- **First-order multi-sorted signature**

  \[ \Sigma_p = (S, \text{Fun}, \text{Pred}, \text{ar}_{\text{fun}} : \text{Fun} \rightarrow S^* \times S, \text{ar}_{\text{pred}} : \text{Pred} \rightarrow S^*) \]

- **Parametric signature**

  \[ \Sigma_e = (\text{Op}, \text{ar}_{\text{op}} : \text{Op} \rightarrow S^* \times \mathbb{N}) \]

- **Terms**

  \[ t ::= x \mid \text{op}_{u_1, \ldots, u_m}(t_1, \ldots, t_n) \ (\text{op} : s_1, \ldots, s_m; n \text{ and } u_i : s_i). \]

- **Equations**

  \[ t = u \ (\varphi) \] where \( \varphi \) is a first-order formula over \( \Sigma_p \).

- **Axiomatisations**

  Sets \( \text{Ax} \) of equations

- **Deduction** (an interesting question, not treated here)
Examples

- **Exceptions** $\Sigma_\rho$ has a single sort $\text{exc}$, and constants $e : 0$ for each $e \in E$.
  $\Sigma_e$ has a single operation symbol $\text{raise} : \text{exc}; 0$. There are no equations.

- **Probability** $\Sigma_\rho$ has a single sort $\text{interval}$, constants $0, 1$, binary function symbols $\times$, a unary function symbol $1-\cdot$, and a relation symbol $\prec$.
  $\Sigma_e$ has a single binary operation symbol $+: \text{interval}; 2$.
  Here is an example equation:

$$
(x +_p y) +_r z = x +_{pr} (y + \frac{r - pr}{1 - pr} z) \quad (r < 1, p < 1)
$$
Addition to $\lambda$-calculus syntax

- **Types**
  \[
  \sigma ::= s \ (s \in S) \ | \ \text{bool}
  \]

- **Terms**
  \[
  M ::= f(M_1, \ldots, M_n) \ (f \in \text{Fun}) \ | \ P(M_1, \ldots, M_n) \ (P \in \text{Pred}) \ | \\
  \text{true} \ | \ \text{false} \ | \ \text{if} \ L \ \text{then} \ M \ \text{else} \ N \ | \\
  \text{op}_{M_1, \ldots, M_m}(N_1, \ldots, N_n)
  \]

- **Example type-checking rule**
  \[
  \Gamma \vdash M_1 : s_1, \ldots, \Gamma \vdash M_m : s_m, \ \Gamma \vdash N_1 : \sigma, \ldots, \Gamma \vdash N_n : \sigma \Rightarrow \\
  \Gamma \vdash \text{op}_{M_1, \ldots, M_m}(N_1, \ldots, N_n) : \sigma
  \]

  where \( \text{op} : s_1, \ldots, s_m ; n \)
Parametric finitary equational theories: semantics

- **Parameter interpretation** We fix an interpretation $\mathcal{M}$ of $\Sigma_p$.
- **Algebras** With that, a $\Sigma_e$-algebra is a structure

$$(A, \text{op}_A : \mathcal{M}[s] \times A^n \to A \ (\text{op} : s; n))$$

where $\mathcal{M}[s_1, \ldots, s_m] = \text{def} \mathcal{M}[s_1] \times \ldots \mathcal{M}[s_m]$.

Homomorphisms are then defined in the evident way.

- **Denotation** $\mathcal{A}[t](\rho_p, \rho_e)$, where $\rho_e : \text{Var} \to A$. For example

$$\mathcal{A}[\text{op}_{u_1, \ldots, u_m}(t_1, \ldots, t_n)](\rho_p, \rho_e) = \text{op}_A(\mathcal{M}[u_1, \ldots, u_m](\rho_p), \mathcal{A}[t_1](\rho_p, \rho_e), \ldots, \mathcal{A}[t_n](\rho_p, \rho_e))$$

Validity and Models are then defined in the evident way.
Theorem

Let $\mathsf{Ax}$ be a set of parametric $\Delta_e$-axioms. Then there is a free model $F_{\mathsf{Ax}}(X)$ of $\mathsf{Ax}$ over any $X$. That is, there is an $\eta : X \to T_{\mathsf{Ax}}(X)$, where $T_{\mathsf{Ax}}(X)$ is the carrier of $F_{\mathsf{Ax}}(X)$, such that for any model $\mathcal{A}$ of $\mathsf{Ax}$, and any function $f : X \to \mathcal{A}$ there is a unique homomorphism $f^\dagger : F_{\mathsf{Ax}}(X) \to \mathcal{A}$ such that the following diagram commutes:

```
X
  |     \\
  |  \eta \\
T_{\mathsf{Ax}}(X)  f  \rightarrow  A
  |   |   \\
  |   |  f^\dagger \\
  |   |     \\
  |     A
```
Idea of proof

The idea is to reduce to ordinary equational theories.

- For every \( \text{op} : s; n \) and \((a_1, \ldots, a_m) \in M[s]\) we introduce an operation symbol \(f_{a_1,\ldots,a_m} : n\).
- Then from any parametric term \(t\) and \(\rho_p\) we can obtain an ordinary term \(t^{\rho_p}\). For example:

\[
\text{op}_{u_1,\ldots,u_m}(t_1, \ldots, t_n)^{\rho_p} = \text{op}_{M[u_1,\ldots,u_m]}(\rho_p)(t_1^{\rho_p}, \ldots, t_n^{\rho_p})
\]

- Then one obtains a set of ordinary equations from any parametric equation in \(Ax\), taking all \(\rho_p\)'s.
- We know all these ordinary equations have a free model. That immediately gives a parametric model of \(Ax\) with the same carrier, “gluing” the interpretations of all the \(f_{a_1,\ldots,a_m}\) together. Keeping the same unit, we immediately deduce parametric freeness from ordinary freeness.
Outline

1. Algebra with parameterised operations
2. Algebra with parameters and parametric arguments
3. Algebraic operations and generic effects (cntnd.)
State treated algebraically

Suppose we have locations which can store natural numbers. We have natural programming notation for reading and writing:

\[
\begin{align*}
M : \text{loc} & \quad M : \text{loc}, N : \text{nat} \\
!M : \text{nat} & \quad M := N : \text{unit}
\end{align*}
\]

But

\[
\begin{align*}
\text{loc} & \xrightarrow{!} \text{nat} \\
\text{loc} \times \text{nat} & \xrightarrow{=} \text{unit}
\end{align*}
\]

do not seem to have much to do with algebra.

Hint: Read “\(M + N\)” as “choose 0 or 1 and then do whichever continuation \(M\) or \(N\) is appropriate.”

One can read \(M +_p N\) similarly, but in terms of tossing a biased coin with head having probability \(p\).
State treated algebraically (cntnd.)

So for writing we would have an operation, \texttt{update} say, which writes and then carries on (i.e. has a single continuation). This suggests:

\begin{verbatim}
update : loc, nat; 1
\end{verbatim}

which fits within parametric algebra.

For reading we would have an operation, \texttt{lookup} say, which reads a location and then carries on with a continuation depending on the value read. This suggests:

\begin{verbatim}
lookup : loc ; nat
\end{verbatim}

a parameterised \texttt{infinitary} operation!

So we now look at infinitary algebra and a finitary notation for it. We will return later to the status of things like \texttt{!} and \texttt{:=} and see that they form part of the general pattern of \texttt{generic effects}.
Infinitary equational logic: syntax

- **Signature** $\Sigma_e = (\text{Op}, \text{ar} : \text{Op} \to \omega + 1)$. We write $\text{op} : n$ for arities, including $\omega$.
- **Terms** as in finitary case plus: $\text{op}(t_1, t_2, \ldots, t_n, \ldots)$ ($\text{op} : \omega$). We leave open what the set $\text{Var}$ of variables is.
- **Equations** $t = u$ as before
- **Axiomatisations** Sets $\text{Ax}$ of equations
- **Deduction** $\text{Ax} \vdash t = u$ an easy variant of the finitary case
- **Theories** Sets of equations $\text{Th}$ closed under deduction
Algebras are structures $\mathcal{A} = (A, \text{op}_A : A^n \rightarrow A \ (\text{op} : n))$, and recall that here $n$ can be $\omega$.

Homomorphisms $h : \mathcal{A} \rightarrow \mathcal{B}$ are, much as before, functions $h : A \rightarrow B$ such that, for all $\text{op} : n$, and $a \in A^n$:

$$h(\text{op}_A(a)) = \text{op}_B(h(a))$$

Denotation $\mathcal{A}[t](\rho)$ is also defined much as before.

Validity $\mathcal{A} \models t = u$ is defined as before.

Models $\mathcal{A}$ is also defined as before.
The free algebra monad $T_{Ax}$ of an infinitary axiomatic theory $Ax$

All is as before. The free model $F_{Ax}(X)$ over a set $X$ has carrier:

$$T_{Ax}(X) = \text{def } \{ [t]_{Ax} \mid t \text{ is a term with variables in } X \}$$

where $[t]_{Ax} = \text{def } \{ u \mid Ax \vdash u = t \}$; its operations are given by:

$$\text{op}_{F_{Ax}(X)}([t]) = [\text{op}(t)] \quad (\text{op} : n)$$

the unit $\eta : X \to T_{Ax}(X)$ is again $x \mapsto [x]$; for any model $A$ of $Ax$, and any function $f : X \to A$ the unique mediating homomorphism $f^\dagger : F_{Ax}(X) \to A$ is given by:

$$f^\dagger([t]) = A[t](f)$$

and the multiplication is $(\text{id}_{T_{Ax}(X)})^\dagger$. 
Notation and equations for state

\[ t ::= \text{update}_{u_1,u_2}(t) \mid \text{lookup}_u(n : \text{nat.}\ t) \mid x(u_1, \ldots, u_n) \]

Equations for writing and reading a single location:

\[
\begin{align*}
\text{update}_{l,m}(\text{update}_{l,n}(x)) &= \text{update}_{l,n}(x) & (1) \\
\text{lookup}_l(m : \text{nat.}\ \text{lookup}_l(n : \text{nat.}\ x(m, n))) &= \text{lookup}_l(m : \text{nat.}\ x(m, m)) & (2) \\
\text{lookup}_l(n : \text{nat.}\ x) &= x & (3) \\
\text{update}_{l,m}(\text{lookup}_l(n : \text{nat.}\ x(n))) &= \text{update}_{l,m}(x(m)) & (4) \\
\text{lookup}_l(n : \text{nat.}\ \text{update}_{l,n}(x)) &= x & (5)
\end{align*}
\]
Commutation Equations for different locations

\[ \text{update}_{l,m}(\text{update}_{l',n}(x)) = \text{update}_{l',n}(\text{update}_{l,m}(x)) \quad (l \neq l') \quad (7) \]

\[ \text{lookup}_l(m: \text{nat. } \text{lookup}_{l'}(n: \text{nat. } x(m,n))) = \]
\[ \text{lookup}_{l'}(n: \text{nat. } \text{lookup}_l(m: \text{nat. } x(m,n))) \quad (l \neq l') \quad (8) \]

\[ \text{update}_{l,m}(\text{lookup}_{l'}(n: \text{nat. } x(n))) = \]
\[ \text{lookup}_{l'}(n: \text{nat. } \text{update}_{l,m}(x(n))) \quad (9) \]
Equations (3), and (2) and (8) (Mellies) are redundant. For example, for (3) we have:

\[
\text{lookup}_I(n \colon \text{nat}\. \, x) = \text{lookup}_I(n \colon \text{nat}\. \, \text{update}_{I,n}(\text{lookup}_I(n \colon \text{nat}\. \, x))) \quad \text{(by (5))}
\]
\[
= \text{lookup}_I(n \colon \text{nat}\. \, \text{update}_{I,n}(x))) \quad \text{(by (4))}
\]
\[
= x \quad \text{(by (5))}
\]
Parametric axiom. ths. with abstraction: syntax

- First-order multi-sorted signature

\[ \Sigma_p = (S, Ar, Fun, Pred, ar_{fun} : Fun \to S^* \times S, ar_{pred} : Pred \to S^*) \]

with a subcollection \( Ar \subseteq S \) of \textit{arity} sorts

- Parametric signature

\[ \Sigma_e = (Op, ar_{op} : Op \to S^* \times Ar^{**}) \]

- Terms

\[ \Gamma, u : s, \Gamma, x_1 : s_1 \vdash t_1, \ldots, \Gamma, x_n : s_n \vdash t_n \]
\[ \Gamma \vdash op_u(x_1 : s_1. t_1, \ldots, x_n : s_n. t_n) \quad (s_i \in Ar^*, op : s; s_1, \ldots, s_n) \]

- Equations \( t = u \ (\phi) \) and axiomatisations \( Ax \) are as before, and deduction remains an interesting question.
Addition to $\lambda$-calculus syntax

- **Types**
  \[ \sigma ::= s \ (s \in S) \mid \text{bool} \]

- **Terms**
  \[ M ::= \text{op}_M(x_1 : s_1 . N_1, \ldots, x_n : s_n . N_n) \]

- **Example type-checking rule**
  \[
  \begin{align*}
  \Gamma \vdash M : s, & \quad \Gamma, x_1 : s_1 \vdash N_1 : \sigma, \ldots, \Gamma, x_n : s_n \vdash N_n : \sigma \\
  & \quad \therefore \Gamma \vdash \text{op}_M(x_1 : s_1 . N_1, \ldots, x_n : s_n . N_n) : \sigma
  \end{align*}
  \]

  where $\text{op} : s ; s_1, \ldots, s_m$
Parametric axiom. ths. with abstraction: semantics

- **Parameter interpretation** We fix an interpretation $\mathcal{M}$ of $\Sigma_p$, such that $\mathcal{M}[s]$ is countable for all $s \in Ar$.

- **Algebras** With that, a $\Sigma_e$-algebra is a structure

  $$(A, \text{op}_A : \mathcal{M}[s] \times A^{\mathcal{M}[s_1]} \times \ldots \times A^{\mathcal{M}[s_n]} \to A \ (\text{op} : s; s_1, \ldots, s_n))$$

- **Denotation** $\mathcal{A}[t](\rho_p, \rho_e)$, where $\rho_e : \text{Var} \to A$. For example

  $$\mathcal{A}[\text{op}_u(x_1 : s_1. t_1, \ldots, x_n : s_n. t_n)](\rho_p, \rho_e) = \text{op}_A(\mathcal{M}[u](\rho_p), \varphi_1, \ldots, \varphi_n)$$

  where:

  $$\varphi_i(a_i) =_{\text{def}} \mathcal{A}[t_i](\rho_p[a/x_i], \rho_e) \quad (i = 1, n, a_i \in \mathcal{M}[s_i])$$

  Homomorphisms, Validity and Models are defined in the evident way.
As usual, there is a free algebra $F_{\text{Ax}}(X)$ over any set $X$, which induces the corresponding monad $T_{\text{Ax}}(X)$.

The proof is by a (now) evident reduction to (countably) infinitary equational logic.

Restricting the denotations of arity types to be finite still covers many situations, e.g., locations storing bits or words. Thus abstraction can be useful even in the finitary case.
An Example: State

- **First order part** The sorts are $\text{loc}, \text{nat}$, and there is a predicate symbol $=: \text{loc}, \text{loc}$. We assume $\mathcal{M}[=]$ is equality, $\mathcal{M}[	ext{loc}]$ is finite, and $\mathcal{M}[	ext{nat}] = \mathbb{N}$. Set $\text{Loc} = \text{def } \mathcal{M}[	ext{loc}]$.
- **Axioms** $\text{Ax}_S$ is as above.
- **Monad** $T_S(X) = (S \times X)^S$, where $S = \text{def } \mathbb{N}^{\text{Loc}}$
- **Operations**
  
  **Lookup** $\text{Loc} \times T_S(X)^\mathbb{N} \xrightarrow{\text{lookup}_{FS(X)}} T_S(X)$ is defined by:
  
  $\text{lookup}_{FS(X)}(l, \varphi) = \sigma \mapsto \varphi(\sigma(l))$

  **Update** $\text{Loc} \times \mathbb{N} \times T_S(X) \xrightarrow{\text{update}_{FS(X)}} T_S(X)$ is defined by:

  $\text{update}_{FS(X)}(l, n, \gamma) = \sigma \mapsto \gamma(\sigma[n/l])$
Another example: interactive I/O

- **First-order part** The sorts are \( \text{in, out} \). The rest, including \( \mathcal{M} \), is as suits the purpose at hand.
- **Operation symbols** input : \( \varepsilon \); \( \text{in} \) and output : \( \text{out} \); \( 1 \)
- **Algebraic Axioms** None!
- **Monad** \( T_{I/O}(X) \) is the least set \( Y \) such that:

\[
Y = Y^{\mathcal{M}[\text{in}]} + (\mathcal{M}[\text{out}] \times Y) + X
\]

and we just write:

\[
T_{I/O}(X) = \mu Y. Y^{\mathcal{M}[\text{in}]} + (\mathcal{M}[\text{out}] \times Y) + X
\]

- \( T_{I/O}(X) \) is a collection of trees. Its internal nodes are either **input** ones, when they have an \( \mathcal{M}[\text{in}] \)-indexed collection of children, or **output** nodes, when they have an \( \mathcal{M}[\text{out}] \) label and one child. Its leaves have an \( X \) label.
I/O cntnd.

- **Operations**

  **Input** \( T_{I/O}(X) \mathcal{M}[\text{in}] \xrightarrow{\text{input}_{F_{I/O}(X)}} T_{I/O}(X) \) is defined by:
  \[
  \text{input}_{F_{I/O}(X)}(\phi) = \text{in}_1(\phi)
  \]

  **Output** \( \mathcal{M}[\text{out}] \times T_{I/O}(X) \xrightarrow{\text{output}_{F_{I/O}(X)}} T_{I/O}(X) \) is defined by:
  \[
  \text{output}_{F_{I/O}(X)}(d, \gamma) = \text{in}_2(d, \gamma)
  \]
The general case, when there are no axioms

We have:

\[ T_{I/O}(X) = \mu Y. \sum_{\text{op: } \mathbf{s}; \mathbf{s}_1, \ldots, \mathbf{s}_n} (\mathcal{M}[\mathbf{s}] \times Y^{\mathcal{M}[\mathbf{s}_1] \times \cdots \times \mathcal{M}[\mathbf{s}_n]}) + X \]

We again have a collection of trees. The internal nodes are \(\mathcal{M}[\mathbf{s}]\)-labelled and have an \(\mathcal{M}[\mathbf{s}_1] \times \cdots \times \mathcal{M}[\mathbf{s}_n]\)-indexed collection of children. As before, the terminal nodes are \(X\)-labelled.
Outline

1. Algebra with parameterised operations
2. Algebra with parameters and parametric arguments
3. Algebraic operations and generic effects (cntnd.)
Algebraic operations, somewhat more generally

Fix a parametric equational axiomatic theory with abstraction $\text{Ax}$, and model $\mathcal{M}$. Then for any set $X$ and operation symbol $\text{op} : \mathbf{s}; \mathbf{s}_1$ we have the function:

$$\mathcal{M}[\mathbf{s}] \times T_{\text{Ax}}(X)^{\mathcal{M}[\mathbf{s}_1]} \xrightarrow{\text{op}_{\text{Ax}}(X)} T_{\text{Ax}}(X)$$

Further for any function $f : X \rightarrow T_{\text{Ax}}(Y)$, $f^\dagger$ is a homomorphism:

$$\mathcal{M}[\mathbf{s}] \times T_{\text{Ax}}(X)^{\mathcal{M}[\mathbf{s}_1]} \xrightarrow{\text{op}_{\text{Ax}}(X)} T_{\text{Ax}}(X)$$

$$\text{id}_{\mathcal{M}[\mathbf{s}]} \times (f^\dagger)^{\mathcal{M}[\mathbf{s}_1]} = f^\dagger$$

$$\mathcal{M}[\mathbf{s}] \times T_{\text{Ax}}(Y)^{\mathcal{M}[\mathbf{s}_1]} \xrightarrow{\text{op}_{\text{Ax}}(Y)} T_{\text{Ax}}(Y)$$

We again call such a family of functions $\varphi_X$ algebraic.
Generic effects, somewhat more generally

- Given an algebraic family
  \[ M[s] \times T_{Ax}(X)^{M[s_1]} \xrightarrow{\varphi_X} T_{Ax}(X) \]
  we obtain the generic effect:
  \[ M[s] \xrightarrow{e} T_{Ax}(M[s_1]) = \varphi_{M[s_1]}(\cdot, \eta_{M[s_1]}) \]

- Given such an \( e \) we obtain such an algebraic family:
  \[
  \begin{align*}
  & M[s] \times T_{Ax}(X)^{M[s_1]} \xrightarrow{id_M[s] \times (\cdot)^\dagger} M[s] \times T_{Ax}(X)^{T_{Ax}(M[s_1])} \\
  & \quad \xrightarrow{e \times id} T_{Ax}(M[s_1]) \times T_{Ax}(X)^{T_{Ax}(M[s_1])} \\
  & \quad \xrightarrow{ev} T_{Ax}(X)
  \end{align*}
  \]

- This correspondence is a bijection between algebraic families and generic effects.
An example: side-effects

- **Lookup** The generic effect corresponding to

\[
\text{Loc} \times T_S(X)^{\mathbb{N}} \xrightarrow{\text{lookup}_{FS(X)}} T_S(X)
\]

is

\[
\text{Loc} \xrightarrow{!} T_S(\mathbb{N}) = (S \times \mathbb{N})^S
\]

where

\[
!(l) = \sigma \mapsto (\sigma, \sigma(l))
\]

- **Update** The generic effect corresponding to

\[
\text{Loc} \times \mathbb{N} \times T_S(X)^{\mathbb{I}} \xrightarrow{\text{update}_{FS(X)}} T_S(X)
\]

is

\[
\text{Loc} \times \mathbb{N} \xrightarrow{:=} T_S(\mathbb{I})
\]

where

\[
:= (l, v) = \sigma \mapsto (\sigma[l/n], *)
\]
Another example: interactive I/O

- **Input** The generic effect corresponding to

  \[
  T_{I/O}(X) \overset{\mathcal{M}[\text{in}]}{\longrightarrow} T_{I/O}(X)
  \]

  is

  \[
  \text{myread} \in T_{I/O}(\mathcal{M}[\text{in}])
  \]

  where

  \[
  \text{myread} = \text{in}_1(d \in \mathcal{M}[\text{in}] \mapsto \text{in}_3(d))
  \]

- **Output** The generic effect corresponding to

  \[
  \mathcal{M}[\text{out}] \times T_{I/O}(X) \overset{\text{output}_{\mathcal{F}_{I/O}(X)}}{\longrightarrow} T_{I/O}(X)
  \]

  is

  \[
  \mathcal{M}[\text{out}] \overset{\text{write}}{\longrightarrow} T_{I/O}((1))
  \]

  where

  \[
  \text{write}(d) = \text{in}_2(d, \text{in}_3(\ast))
  \]
Programming counterpart of being algebraic

- **Evaluation contexts** are given by:
  \[ \mathcal{E} ::= [ \cdot ] \mid \mathcal{E} \mathcal{N} \mid (\lambda x : \sigma. M)\mathcal{E} \]

- For any operation symbol \( \text{op} : s; s_1, \ldots, s_m \) we have:
  \[
  \models \mathcal{E}[\text{op}_M(x_1 : s_1. N_1, \ldots, x_n : s_n. N_n)] = \text{op}_M(x_1 : s_1. \mathcal{E}[N_1], \ldots, x_n : s_n. \mathcal{E}[N_n])
  \]
  assuming variable clashes are avoided.