A parsimonious module system

TODO

1 Introduction

TODO: BLABLA Intro

We propose an expressive, syntactic and parsimonious module system. By expressive, we mean that we support a rich subset of features that have so far not been formalized together and allow the user for rich manipulation of modules, notably transparent ascriptions, module aliases, enrichment constraints and applicative functors. Unlike work like ?, our rules are defined syntactically, directly on the concrete syntax. This makes the rule easier to grasp, make deriving an inference algorithm almost immediate, and ensure that error messages and typing information are easier to surface to the user. Finally, this system is parsimonious, in that it does the minimum amount of expansion required by using ideas similar to explicit substitutions for module operations. Module types can be quite large, and limiting the expansion can simplify error messages and improve the speed of the typechecker.

2 Motivations

TODO: Explain functor aliasing.

TODO: Explain transparent ascriptions, with constraints

TODO: Explain avoidance problem

TODO: Add Examples

TODO: Motivate incrementality

3 An ML module calculus

We now introduce our module calculus. First, let us defined some conventions: lowercase meta-variables such as x, e or t are used for the core language; capitalized meta-variables such as X, M, S represent modules; calligraphic meta-variables such as X, M, S represent module types.

3.1 Syntax

The syntax is defined in Figure 1. The module expression language contains paths P, which might be variables X or qualified accesses P.X, parameterized modules i.e. functors $((X : \mathcal{M}) \to M)$, module type annotations i.e. ascriptions, both opaque $(M : \mathcal{M})$ and transparent $(M <: \mathcal{M})$, and finally structures $\mathsf{struct}_X S$ end. Structures contains a list of declarations, noted D, which can be value, type, module or module type bindings. Structures are also annotated with a "Self" variable, which represent the object through which other field in the module should be accessed.

In our core calculus, module types are stratified. \mathcal{M} represent module types with operations, while \mathcal{M}° are raw module types without initial operations (i.e., in head normal form). Raw module types can be (qualified) module type variables \mathcal{X} or $\mathcal{P}.\mathcal{X}$, functor types $(\mathcal{X}:\mathcal{M}_1) \to \mathcal{M}_2$, signatures $\operatorname{sig}_X \mathcal{S}$ end, and singletons (= \mathcal{P}). As with structures, signatures are a list of declarations noted \mathcal{D} , and a "Self" variable. Allowed operations strengthening a module by a path, noted \mathcal{M}/\mathcal{P} , enriching a module with additional constraints, noted $\mathcal{M}_{[\mathcal{C}]}$, and locally binding a module name, noted let $X:\mathcal{M}$ in \mathcal{M}' .

As described before, there are two kinds of paths. The first, P, is mostly used for *dynamic* components, and is composed of a list of modules of the form A.B.C. The second, \mathcal{P} , is mostly used for *static* components such as types, and can additionally contain functor applications and transparent ascriptions, for instance F(X).Y or $(X <: \mathcal{M}).Y$

Finally, the core language is left largely undefined in our calculus, except for qualified accesses.

 $\mathcal{M}[X \mapsto Y]$ denotes the substitution of variable X by variable Y in \mathcal{M} . We only substitute variables by other variables.

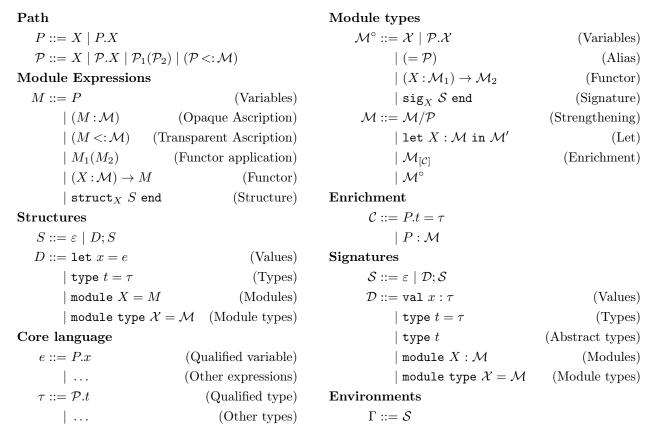


Figure 1: Module language

3.2 Notations

Let us now define a few notations that we will use in the rest of this article. We will then describe how these various construction interact on specific features.

We define the following type judgements on modules. Note that the subtyping judgement also *returns* a module type, which is the result of the transparent ascription between the two modules.

```
\Gamma \vdash M : \mathcal{M} : The module M is of type \mathcal{M} in \Gamma. See Figure 2.

\Gamma \vdash \mathcal{M} <: \mathcal{M}' \leadsto \mathcal{M}_r : The module type \mathcal{M} is a subtype of \mathcal{M}' in environment \Gamma, and their ascription is \mathcal{M}_r. See Figure 3.

(\mathcal{M})_{\Gamma} = \mathcal{M}^{\circ} : The module type with operations \mathcal{M} can be reduced to a module type without operations \mathcal{M}^{\circ} in environment \Gamma. See Figures 4 to 6. TODO: Change bracket style
```

$$\frac{P \in \Gamma}{\Gamma \vdash P : (=P)} \qquad \frac{\Gamma \vdash M : \mathcal{M}' \quad \Gamma \vdash \mathcal{M}' < : \mathcal{M} \leadsto}{\Gamma \vdash (M : \mathcal{M}) : \mathcal{M}} = \qquad \frac{\Gamma \vdash M : \mathcal{M}' \quad \Gamma \vdash \mathcal{M}' < : \mathcal{M} \leadsto \mathcal{M}_r}{\Gamma \vdash (M < : \mathcal{M}) : \mathcal{M}_r}$$

$$\frac{\Gamma \vdash M_f : (X : \mathcal{M}_a) \to \mathcal{M}_r \quad Y \text{ fresh} \quad \Gamma \vdash M : \mathcal{M} \quad \Gamma \vdash \mathcal{M} < : \mathcal{M}_a \leadsto \mathcal{M}_c}{\Gamma \vdash M_f (M) : \text{let } Y : \mathcal{M}_c \text{ in } \mathcal{M}_r [X \mapsto Y]}$$

$$\frac{\Gamma \vdash \mathcal{M} \quad X \notin \Gamma \quad \Gamma; \text{module } X : \mathcal{M} \vdash \mathcal{M} : \mathcal{M}' \quad \Gamma \vdash X : \mathcal{S} : \mathcal{S}}{\Gamma \vdash (X : \mathcal{M}) \to \mathcal{M} : (X : \mathcal{M}) \to \mathcal{M}'} \qquad \frac{\Gamma \vdash_X S : \mathcal{S}}{\Gamma \vdash \text{struct}_X S \text{ end} : \text{sig}_X S \text{ end}} \qquad \frac{\Gamma \vdash_Y \varepsilon : \varepsilon}{\Gamma \vdash_Y \varepsilon : \varepsilon}$$

$$\frac{\Gamma \vdash e : \tau \quad x \notin \Gamma(Y) \quad \Gamma; \text{val } Y : x : \tau \vdash_Y S : \mathcal{S}}{\Gamma \vdash_Y (\text{let } x = e; S) : (\text{val } x : \tau \vdash_X S)} \qquad \frac{\Gamma \vdash \mathcal{M} : \mathcal{M} \quad X \notin \Gamma(Y) \quad \Gamma; \text{module } Y : \mathcal{M} \vdash_Y S : \mathcal{S}}{\Gamma \vdash_Y (\text{module } X = \mathcal{M}; S) : (\text{module } X : \mathcal{M} \vdash_Y S : \mathcal{S}}$$

$$\frac{\Gamma \vdash \tau \quad t \notin \Gamma(Y) \quad \Gamma; \text{type } Y : t = \tau \vdash_Y S : \mathcal{S}}{\Gamma \vdash_Y (\text{type } t = \tau; S) : (\text{type } t = \tau; S)} \qquad \frac{\Gamma \vdash \mathcal{M} \quad \mathcal{X} \notin \Gamma(Y) \quad \Gamma; \text{module type } Y : \mathcal{X} = \mathcal{M} \vdash_Y S : \mathcal{S}}{\Gamma \vdash_Y (\text{module type } \mathcal{X} = \mathcal{M}; S) : (\text{module type } \mathcal{X} = \mathcal{M}; S)}$$

Figure 2: Module typing rules – $\Gamma \vdash M : \mathcal{M}$

In our calculus, environments are simply signature on which we can "query" fields. To properly model our module calculus, we need to carefully consider how these environments are accessed. For this purpose, we consider the following operations on environments, which are defined in Figure 7

- $\Gamma(P)$: Lookup the path P in Γ to a module or module type.
- normalize $\Gamma(\mathcal{P})$: Normalizes the path \mathcal{P} in Γ to a canonical path.
 - $\operatorname{resolve}_{\Gamma}(\mathcal{P})$: Resolve the path \mathcal{P} in Γ to its shape (i.e., either an arrow or a signature).
 - shape_{Γ}(\mathcal{M}): Resolve the module \mathcal{M} in Γ to its shape (i.e., either an arrow or a signature).
- 3.3 Structures and qualified accesses
- 3.4 Subtyping and transparent ascriptions
- 3.5 Functors
- 3.6 Module operations

$$\frac{\operatorname{normalize}_{\Gamma}(\mathcal{P}) = \operatorname{normalize}_{\Gamma}(\mathcal{P}')}{\Gamma \vdash \mathcal{P} = \mathcal{P}'} \qquad \frac{\Gamma \vdash \{\!\!\!/ \mathcal{M} \}\!\!\! \Gamma <: \{\!\!\!/ \mathcal{M}' \}\!\!\! \Gamma \hookrightarrow \mathcal{M}'' }{\Gamma \vdash \mathcal{M} <: \mathcal{M}' \hookrightarrow \mathcal{M}''} \qquad \frac{\Gamma \vdash \Gamma(\mathcal{P}) <: \mathcal{M}^{\circ} \hookrightarrow \mathcal{M}}{\Gamma \vdash \mathcal{P} <: \mathcal{M}^{\circ} \hookrightarrow \mathcal{M}}$$

$$\frac{\Gamma \vdash \mathcal{M}^{\circ} <: \Gamma(\mathcal{P}) \hookrightarrow \mathcal{M}}{\Gamma \vdash \mathcal{M}^{\circ} <: \mathcal{P} \hookrightarrow \mathcal{M}} \qquad \frac{\Gamma \vdash \mathcal{P} = \mathcal{P}' \qquad \Gamma \vdash \operatorname{resolve}_{\Gamma}(\mathcal{P}) <: \mathcal{M} \hookrightarrow \mathcal{M}'' \qquad \Gamma \vdash \mathcal{M}'' <: \mathcal{M}' \hookrightarrow \mathcal{M}'}{\Gamma \vdash (= (\mathcal{P} <: \mathcal{M})) <: (= (\mathcal{P}' <: \mathcal{M}')) \hookrightarrow (= (\mathcal{P}' <: \mathcal{M}'))} \qquad \frac{\Gamma \vdash \mathcal{P} = \mathcal{P}' \qquad \Gamma \vdash \operatorname{resolve}_{\Gamma}(\mathcal{P}) <: \mathcal{M} \hookrightarrow \mathcal{M}'' \qquad \Gamma \vdash \mathcal{M}'' <: \operatorname{resolve}_{\Gamma}(\mathcal{P}) \hookrightarrow \mathcal{M}' \hookrightarrow \mathcal{M}''}{\Gamma \vdash (= (\mathcal{P} <: \mathcal{M})) <: (= \mathcal{P}') \hookrightarrow (= \mathcal{P}')} \qquad \frac{\Gamma \vdash \mathcal{P} = \mathcal{P}'}{\Gamma \vdash (= (\mathcal{P} <: \mathcal{M}'))} \qquad \frac{\Gamma \vdash \mathcal{P} = \mathcal{P}'}{\Gamma \vdash (= \mathcal{P}) <: (= (\mathcal{P}' <: \mathcal{M}'))} \qquad \frac{\Gamma \vdash \mathcal{P} = \mathcal{P}'}{\Gamma \vdash (= \mathcal{P}) <: (= (\mathcal{P}') \hookrightarrow (= \mathcal{P}'))} \qquad \frac{\Gamma \vdash \mathcal{M}'' <: \mathcal{M} \hookrightarrow \mathcal{M}' \hookrightarrow \mathcal{M}' \hookrightarrow \mathcal{M}''}{\Gamma \vdash (= \mathcal{P}) <: (\mathcal{P} <: \mathcal{M}^{\circ})} \qquad \frac{\Gamma \vdash \mathcal{M}'' <: \mathcal{M} \hookrightarrow \mathcal{M} \hookrightarrow \mathcal{M}' \hookrightarrow \mathcal{M}'$$

$$\overline{\Gamma \vdash (\mathsf{module} \ X : \mathcal{M}_1) <: (\mathsf{module} \ X : \mathcal{M}_2) \leadsto (\mathsf{module} \ X : \mathcal{M})} \\ \overline{\Gamma \vdash (\mathsf{module} \ X : \mathcal{M}_1) <: \emptyset \leadsto (\mathsf{module} \ X : \mathcal{M})} \\ \overline{\Gamma \vdash (\mathsf{module} \ X : \mathcal{M}_1) <: \emptyset \leadsto (\mathsf{module} \ X : \mathcal{M})} \\ \overline{\Gamma \vdash (\mathsf{module} \ \mathsf{type} \ \mathcal{X} = \mathcal{M}_1) <: (\mathsf{module} \ \mathsf{type} \ \mathcal{X} = \mathcal{M}_2) \leadsto (\mathsf{module} \ \mathsf{type} \ \mathcal{X} = \mathcal{M})} \\ \overline{\Gamma \vdash (\mathsf{module} \ \mathsf{type} \ \mathcal{X} = \mathcal{M}) <: \emptyset \leadsto (\mathsf{module} \ \mathsf{type} \ \mathcal{X} = \mathcal{M})} \\ \overline{\Gamma \vdash (\mathsf{module} \ \mathsf{type} \ \mathcal{X} = \mathcal{M})} \\ \overline{\Gamma \vdash (\mathsf{type} \ t : \tau_1) <: (\mathsf{val} \ x : \tau_2) \leadsto (\mathsf{val} \ x : \tau_2)}} \\ \overline{\Gamma \vdash (\mathsf{type} \ t) <: (\mathsf{type} \ t) \leadsto (\mathsf{type} \ t)}} \\ \overline{\Gamma \vdash (\mathsf{type} \ t = \tau_1) <: (\mathsf{type} \ t = \tau_1)} \\ \overline{\Gamma \vdash (\mathsf{type} \ t) <: (\mathsf{type} \ t = \tau)} \\ \overline{\Gamma \vdash (\mathsf{type} \ t = \tau) <: (\mathsf{type} \ t) \leadsto (\mathsf{type} \ t = \tau)}} \\ \overline{\Gamma \vdash (\mathsf{type} \ t = \tau) <: (\mathsf{type} \ t) \leadsto (\mathsf{type} \ t = \tau)}} \\ \overline{\Gamma \vdash (\mathsf{type} \ t = \tau) <: (\mathsf{type} \ t) \leadsto (\mathsf{type} \ t = \tau)}}$$

Figure 3: Module subtyping rules – $\Gamma \vdash \mathcal{M} <: \mathcal{M}' \leadsto \mathcal{M}_r$

Figure 4: Module strengthening operation – \mathcal{M}/\mathcal{P}

$$(\mathcal{X}_{[C]})_{\Gamma} \to (\Gamma(\mathcal{X})_{[C]})_{\Gamma}$$

$$(\mathcal{P}.\mathcal{X}_{[C]})_{\Gamma} \to (\Gamma(\mathcal{P}.\mathcal{X})_{[C]})_{\Gamma}$$

$$((=\mathcal{P})_{[C]})_{\Gamma} \to (=\mathcal{P})$$

$$((X:\mathcal{M}) \to \mathcal{M}')_{[C]})_{\Gamma} \to \text{fail}$$

$$(\text{sig}_{X} \ \mathcal{S} \ \text{end}_{[C]})_{\Gamma} \to \text{sig}_{X} \ (\mathcal{S}_{[C]})_{\Gamma} \ \text{end}$$

$$(\text{type } t = \tau; \mathcal{S}_{[t=\tau']})_{\Gamma} \to \text{type } t = \tau'; \mathcal{S}$$

$$(\text{module } X:\mathcal{M}; \mathcal{S}_{[X:\mathcal{M}']})_{\Gamma} \to \text{module } X:\mathcal{M}'; \mathcal{S}$$

$$(\text{module } X:\mathcal{M}; \mathcal{S}_{[X:\mathcal{P}.t=\tau]})_{\Gamma} \to \text{module } X: (\mathcal{M}_{[P.t=\tau]})_{\Gamma}; \mathcal{S}$$

$$(\text{module } X:\mathcal{M}; \mathcal{S}_{[X.P:t=\tau]})_{\Gamma} \to \text{module } X: (\mathcal{M}_{[P:\mathcal{M}']})_{\Gamma}; \mathcal{S}$$

$$((\mathcal{D};\mathcal{S})_{[C]})_{\Gamma} \to \mathcal{D}; (\mathcal{S}_{[C]})_{\Gamma}$$

$$((\mathcal{M}/P)_{[C]})_{\Gamma} \to ((\mathcal{M}/P)_{\Gamma_{[C]}})_{\Gamma}$$

$$(\mathcal{M}_{[C']_{[C]}})_{\Gamma} \to ((\mathcal{M}/P)_{\Gamma_{[C]}})_{\Gamma}$$

$$(\mathcal{M}_{[C']_{[C]}})_{\Gamma} \to ((\mathcal{M}/P)_{\Gamma_{[C]}})_{\Gamma}$$

Figure 5: Enrichment operation – $\mathcal{M}_{[\mathcal{C}]}$

$$\begin{split} (\text{let } X: (=\mathcal{P}) \text{ in } \mathcal{M}|_{\Gamma} &\to \mathcal{M}[X \mapsto \mathcal{P}] \\ (\text{let } X: \mathcal{M} \text{ in } \mathcal{M}'|_{\Gamma} &\to \mathcal{M}' \end{split} \qquad \text{when } X \notin FV(\mathcal{M}') \end{split}$$

TODO: Provide something more elaborate

Figure 6: Subst operation – let $X : \mathcal{M}$ in \mathcal{M}'

$$\frac{(\text{module }X:\mathcal{M})\in\Gamma}{\Gamma(X)=\mathcal{M}} \qquad \frac{\text{shape}_{\Gamma}(\Gamma(\mathcal{P}))=\text{sig}_{Y}\ \mathcal{S}_{1};\text{module }X:\mathcal{M};\mathcal{S}_{2}\text{ end}}{\Gamma(\mathcal{P}X)=\mathcal{M}[Y\mapsto\mathcal{P}]}$$

$$\frac{(\text{module type }\mathcal{X}=\mathcal{M})\in\Gamma}{\Gamma(\mathcal{X})=\mathcal{M}} \qquad \frac{\text{shape}_{\Gamma}(\Gamma(\mathcal{P}))=\text{sig}_{Y}\ \mathcal{S}_{1};\text{module type }\mathcal{X}=\mathcal{M};\mathcal{S}_{2}\text{ end}}{\Gamma(\mathcal{P}X)=\mathcal{M}[Y\mapsto\mathcal{P}]}$$

$$\frac{\Gamma(\mathcal{P})=\mathcal{M}}{\Gamma(\mathcal{P})=\mathcal{M}} \qquad \frac{\Gamma\vdash\mathcal{M}<:\mathcal{M}'\leadsto\mathcal{M}_{r}}{\Gamma((\mathcal{P}<:\mathcal{M}'))=\mathcal{M}_{r}} \qquad \frac{\text{shape}_{\Gamma}(\Gamma(\mathcal{P}))=(X:\mathcal{M}_{a})\to\mathcal{M}_{r}}{\Gamma(\mathcal{P}_{f}(\mathcal{P}_{a}))=\text{let }Y:\mathcal{M}\text{ in }\mathcal{M}_{r}[X\mapsto Y]}$$

$$(a) \text{ Lookup rules } -\Gamma(\mathcal{P})=\mathcal{M}$$

$$\text{normalize}_{\Gamma}(\mathcal{P})=\begin{cases} \text{normalize}_{\Gamma}(\mathcal{P}') & \text{when } \Gamma(\mathcal{P})=(=\mathcal{P}')\\ \mathcal{P} & \text{otherwise} \end{cases} \qquad \frac{\text{shape}_{\Gamma}(\mathcal{M})=\text{shape}_{\Gamma}(\mathcal{M})_{\Gamma}}{\text{shape}_{\Gamma}((=\mathcal{P}))=\text{shape}_{\Gamma}(\text{resolve}_{\Gamma}(\mathcal{P}))}$$

$$\text{shape}_{\Gamma}(\mathbb{Sig}_{X}\ \mathcal{S}\text{ end})=\text{sig}_{X}\ \mathcal{S}\text{ end}$$

$$\text{resolve}_{\Gamma}(\mathcal{P})=\Gamma(\text{normalize}_{\Gamma}(\mathcal{P}))/\text{normalize}_{\Gamma}(\mathcal{P})} \qquad \frac{\text{shape}_{\Gamma}(\mathbb{Sig}_{X}\ \mathcal{S}\text{ end})=\text{sig}_{X}\ \mathcal{S}\text{ end}}{\text{shape}_{\Gamma}((X:\mathcal{M})\to\mathcal{M}')=(X:\mathcal{M})\to\mathcal{M}'}$$

$$(b) \text{ Path normalization and resolution} \qquad (c) \text{ Shape resolution } -\text{shape}_{\Gamma}(\mathcal{M})=\mathcal{M}'$$

Figure 7: Environment access, normalization and resolution