# On Type Checking Delta-Oriented Product Lines*

Ferruccio Damiani and Michael Lienhardt

University of Torino, Italy
{ferruccio.damiani, michael.lienhardt}@unito.it

**Abstract.** A Software Product Line (SPL) is a set of similar programs generated from a common code base. Delta Oriented Programming (DOP) is a flexible approach to implement SPLs. Efficiently type checking an SPL (i.e., checking that all its programs are well-typed) is challenging. This paper proposes a novel type checking approach for DOP. Intrinsic complexity of SPL type checking is addressed by providing early detection of type errors and by reducing type checking to satisfiability of a propositional formula. The approach is tunable to exploit automatically checkable DOP guidelines for making an SPL more comprehensible and type checking more efficient. The approach and guidelines are formalized by means of a core calculus for DOP of product lines of Java programs.

## 1 Introduction

A *Software Product Line* (SPL) is a set of similar programs, called *variants*, with a common code base and well documented variability [6]. *Delta-Oriented Programming* (DOP) [18, 19, 5] is a flexible transformational approach to implement SPLs. A DOP product line is described by a *Feature Model* (FM), a *Configuration Knowledge* (CK), and an *Artifact Base* (AB). The FM provides an abstract description of variants in terms of *features*: each feature represents an abstract description of functionality and each variant is identified by a set of features, called a *product*. The AB provides language dependent code artifacts that are used to build the variants: it consists of a *base program* (that might be empty or incomplete) and of a set of *delta modules*, which are containers of modifications to a program (e.g., for Java programs, a delta module can add, remove or modify classes and interfaces). The CK connects the code artifacts in the AB with the features in the FM (thus defining a mapping from products to variants): it associates to each delta module an *activation condition* over the features and specifies an *application ordering* between delta modules [19]. DOP supports the automatic generation of variants based on a selection of features:

* The authors of this paper are listed in alphabetical order. This work has been partially supported by: project HyVar (www.hyvar-project.eu), which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 644298; by ICT COST Action IC1402 ARVI (www.cost-arvi.eu); and by Ateneo/CSP D16D15000360005 project RunVar.

once a user selects a product, the corresponding variant is derived by applying the delta modules with a satisfied activation condition to the base program according to the application ordering.

DOP is a generalization of *Feature-Oriented Programming* (FOP) [4, 22, 9]: the artifact base of a FOP product line consists of a set of *feature modules* which are delta modules that correspond one-to-one to features and do not contain remove operations. Hence FOP product line development always starts from base feature modules corresponding to mandatory features. Instead, DOP allows to use arbitrary code as a base program. For example, the base program can be empty and different variants can be used as base delta modules with pairwise disjoint activation conditions [20]. Therefore, DOP supports both proactive SPL development (i.e., planning all products/variants in advance) and extractive SPL development [15] (i.e., starting from existing programs). Moreover (see, e.g., [5]), the decoupling between features and delta modules allows to counter the optional feature problem [13], where additional glue code is needed in order to make optional features to cooperate properly. Due to the additional flexibility, in DOP it is more challenging than in FOP to efficiently type check a product line [5]. Type checking approaches for DOP have already been studied [8, 5], and implemented [1] for the ABS modeling language [12]. Although these approaches do not require to generate any variant, they involve an explicit iteration over the set of products, which becomes an issue when the number of products is large (a product line with $n$ features can have up to $2^n$ products).

In this paper we propose a novel type checking approach for DOP by building on ideas proposed for FOP [22, 9]. Our approach represents an achievement over previous type checking approaches for DOP [5, 8] since it provides earlier detection of some type errors and does not require to iterate over the set of products. Like the techniques in [22, 9], our approach requires to check the validity of a propositional formula (which is a co-NP-complete problem) and can take advantages of the many heuristics implemented in SAT solvers (a SAT solver can be used to check whether a propositional formula is valid by checking whether its negation is unsatisfiable)—[22, 9] report that the performance of using SAT solvers to verify the propositional formulas for four non-trivial product lines was encouraging and that, for the largest product line, applying the approach was even faster than generating and compiling a single product. Moreover, our approach is designed to be tunable to take advantage of automatically checkable DOP guidelines that make a product line more comprehensible and type checking more efficient. We formalize the approach and guidelines by means of IMPERATIVE FEATHERWEIGHT DELTA JAVA (IF$\Delta$J) [5], a core calculus for DOP product lines where variants are written in an imperative version of FEATHERWEIGHT JAVA (FJ) [11].

Section 2 introduces an example that will be used through the paper and recalls IF$\Delta$J. Section 3 introduces two DOP guidelines (*no-useless-operations* and *type-uniformity*). Section 4 gives a version of the approach tuned to exploit type-uniformity. Section 5 outlines a version that exploits no guidelines. Section 6 proposes other guidelines. Section 7 discusses related work. Section 8 concludes the paper by outlining planned future work. Proofs of the main results and a prototypical implementation are available in [2] (currently only the version of the approach in Section 4 is supported).

## 2 Model

In this section we introduce the running example of this paper and briefly recall the IF$\Delta$J [5] core calculus. A product line $L$ consist of a feature model, a configuration knowledge, and an artifact base. In IF$\Delta$J there is no concrete syntax for the feature model and the configuration knowledge. We use the following notations: $L$.`features` is the set of features; $L$.`products` specifies the products (i.e., a subset of the power set $2^{L.\texttt{features}}$); $L$.`activation` maps each delta module name `d` to its activation condition; and $L$.`order` (or $<_L$, for short) is the application ordering between the delta modules. Both the set of valid products and the activation condition of the delta modules are expressed as propositional logic formulas $\Phi$ where propositional variables are feature names $\varphi$ (see [3] for a discussion on other possible representations):

$$\Phi \;::=\; \textbf{true} \;\mid\; \varphi \;\mid\; \Phi \Rightarrow \Phi \;\mid\; \neg\Phi \;\mid\; \Phi \wedge \Phi \;\mid\; \Phi \vee \Phi.$$

As usual, we say that a propositional formula $\Phi$ is *valid* if it is true for all values of its propositional variables. To avoid over-specification, the order $<_L$ can be partial. We assume *unambiguity* of the product line, i.e., for each product, any total ordering of the activated delta modules that respects $<_L$ generates the same variant. We refer to [16, 5] for a discussion on an effective means to ensure unambiguity.

The running example of this paper is a version of the *Expression Product Line* (EPL) benchmark [17] (see also [5]) defined by the following grammar which describes a language of numerical expressions:

Exp ::= Lit | Add        Lit ::= <non−negative−integers>        Add ::= Exp "+" Exp

Each variant of the EPL contains a class `Exp` that represents an expression equipped with a subset of the following operations: `toInt`, which returns the value of the expression as an integer (an object of class `Int`); `toString`, which returns the expression as a `String`; and `eval`, which in some variants returns the value of the expression as a `Lit` (the subclass of `Exp` representing literals) and in the other variants returns it as an `Int`. The EPL has 6 products, described by two feature sets: one concerned with data—fLit, fAdd—and one concerned with operations —fToInt, fToString, fEval1, fEval2. Features fLit and fToInt are mandatory. The other features are optional with the two following constraints: exactly one between fEval1 and fEval2 must be selected; and fEval1 requires fToString. The EPL is illustrated in Figure 1. The partial order $L$.`order` is expressed as a total order on a partition of the set of delta modules. To make the example more readable, in the artifact base we use the JAVA syntax for operations on strings and sequential composition—encoding in IF$\Delta$J syntax is straightforward (see [5] for examples). Note that, in the method `Test.test` (in the base program), the

$$
\begin{array}{lcl}
EPL.\texttt{features} & = & \{\mathsf{fLit}, \mathsf{fAdd}, \mathsf{fToInt}, \mathsf{fToString}, \mathsf{fEval1}, \mathsf{fEval2}\} \\
EPL.\texttt{products} & = & \mathsf{fLit} \wedge \mathsf{fToInt} \wedge (\mathsf{fEval1} \Rightarrow \mathsf{fToString}) \wedge (\mathsf{fEval1} \vee \mathsf{fEval2}) \wedge \neg(\mathsf{fEval1} \wedge \mathsf{fEval2})
\end{array}
$$

---

$$
\begin{array}{lcl}
EPL.\texttt{order} & = & \{\texttt{dAdd}\} <_L \{\texttt{d\_notTostr}, \texttt{dAdd\_notTostr}\} <_L \{\texttt{dEval1}, \texttt{dEval2}\} \\
EPL.\texttt{activation} & = & \texttt{dAdd} \mapsto \mathsf{fAdd}, \\
 & & \texttt{d\_notTostr} \mapsto (\neg\mathsf{fToString}), \quad \texttt{dAdd\_notTostr} \mapsto (\mathsf{fAdd} \wedge \neg\mathsf{fToString}), \\
 & & \texttt{dEval1} \mapsto \mathsf{fEval1}, \quad \texttt{dEval2} \mapsto \mathsf{fEval1}
\end{array}
$$

---

```
// Base program
class Exp extends Object { // To be used only as a type (i.e., not to be instantiated)
  Int toInt() { return new Int(); }
  String toString() { return ""; }
}
class Lit extends Exp {
  Int val;
  Lit setLit(Int x) { this.val=x; return this; }
  Int toInt() { return this.val; }
  String toString() { return this.val.toString(); }
}
class Test extends Object {
  String test(Exp x) { return x.eval().toString(); }
}
// Delta Modules
delta dAdd {
  adds class Add extends Exp {
    Exp a;   Exp b;
    Int toInt() { return this.a.toInt().add(this.b.toInt()); }
    String toString() { return this.a.toString() + "+" + this.b.toString(); }
  }
}
delta d_notTostr {
  modifies class Exp { removes toString; }
  modifies class Lit { removes toString; }
}
delta dAdd_notTostr {   modifies class Add { removes toString; }   }
delta dEval1 {   modifies class Exp { adds Lit eval() {return (new Lit()).setLit(this.toInt());} }   }
delta dEval2 {   modifies class Exp { adds Int eval() {return this.toInt();} }   }
```

**Fig. 1.** Expression Product Line: FM (top), CK (middle), AB (bottom)

expression `x.eval()` has type `Lit` if feature fEval1 is selected (for this reason feature fEval1 requires feature fToString) and type `Int` otherwise.

In the following, we first introduce the IFJ calculus, which is an imperative version of FJ [11], and then we introduce the constructs for variability on top of it. The abstract syntax of IFJ is presented in Figure 2 (top). Following [11], we use the overline notation for (possibly empty) sequences of elements: for instance $\overline{e}$ stands for a sequence of expressions. Variables `x` include the special variable `this` (implicitly bound in any method declaration $MD$), which may not be used as the name of a method's formal parameter. A program $P$ is a sequence of class declarations $\overline{CD}$. A class declaration **class** `C` **extends** `C′` $\{ \overline{AD} \}$ comprises the name `C` of the class, the name `C′` of the superclass (which must always be specified, even if it is the built-in class `Object`), and a list of field and method declarations $\overline{AD}$. All fields and methods are public, there is no field shadowing, there is no method overloading, and each class is assumed to have an implicit constructor that initializes all fields to **null**. The subtyping relation <: on classes, which is the reflexive and transitive closure of the immediate subclass relation

$$
\begin{array}{lll}
P & ::= \overline{CD} & \text{Program} \\
CD & ::= \textsf{\textbf{class}}\ \mathtt{C}\ \textsf{\textbf{extends}}\ \mathtt{C}\ \{\ \overline{AD}\ \} & \text{Class} \\
AD & ::= FD \ \mid\ MD & \text{Attribute (Field or Method)} \\
FD & ::= \mathtt{C}\ \mathtt{f} & \text{Field} \\
MD & ::= \mathtt{C}\ \mathtt{m}(\overline{\mathtt{C}\ \mathtt{x}})\ \{\textsf{\textbf{return}}\ e;\} & \text{Method} \\
e & ::= \mathtt{x} \ \mid\ e.\mathtt{f} \ \mid\ e.\mathtt{m}(\overline{e}) \ \mid\ \textsf{\textbf{new}}\ \mathtt{C}() \ \mid\ (\mathtt{C})e \ \mid\ e.\mathtt{f} = e \ \mid\ \textsf{\textbf{null}} & \text{Expression}
\end{array}
$$

$$
\begin{array}{lll}
L & ::= FM \quad CK \quad AB & \text{Product Line} \\
AB & ::= P \quad \overline{\Delta} & \text{Artifact Base} \\
\Delta & ::= \textsf{\textbf{delta}}\ \mathtt{d}\ \{\ \overline{CO}\ \} & \text{Delta Module} \\
CO & ::= \textsf{\textbf{adds}}\ CD \ \mid\ \textsf{\textbf{removes}}\ \mathtt{C} \ \mid\ \textsf{\textbf{modifies}}\ \mathtt{C}\ [\textsf{\textbf{extends}}\ \mathtt{C}']\ \{\ \overline{AO}\ \} & \text{Class Operation} \\
AO & ::= \textsf{\textbf{adds}}\ AD \ \mid\ \textsf{\textbf{removes}}\ \mathtt{a} \ \mid\ \textsf{\textbf{modifies}}\ MD & \text{Attribute Operation}
\end{array}
$$

**Fig. 2.** Syntax of IFJ (top) and of IF$\Delta$J (bottom)

(given by the **extends** clauses in class declarations), is assumed to be acyclic. Type system, operational semantics, and type soundness for IFJ are given in [5].

The abstract syntax of the language IF$\Delta$J is given in Figure 2 (bottom). An IF$\Delta$J program $L$ comprises: a feature model $FM$, a configuration knowledge $CK$, and an artifact base $AB$. Recall that we do not consider a concrete syntax for $FM$ and $CK$ and use the notations $L$.`features`, $L$.`products`, $L$.`activation`, and $L$.`order` ($<_L$ for short) introduced above. The artifact base comprises a possibly empty or incomplete IFJ program $P$, and a set of delta modules $\overline{\Delta}$.

A delta module declaration $\Delta$ comprises the name `d` of the delta module and class operations $\overline{CO}$ representing the transformations performed when the delta module is applied to an IFJ program. A class operation can add, remove, or modify a class. A class can be modified by (possibly) changing its super class and performing attribute operations $\overline{AO}$ on its body. An *attribute name* `a` is either a field name `f` or a method name `m`. An attribute operation can add or remove fields and methods, and modify the implementation of a method by replacing its body. The new body may call the special method `original`, which is implicitly bound to the previous implementation of the method and may not be used as the name of a method. The class operations in a delta module must act on distinct classes, and the attribute operations in a class operation must act on distinct attributes. The operational semantics of IF$\Delta$J variant generation is given in [5].

We conclude this section with some notations and definitions. First, in the rest of the document, we will use the term *module* to refer to the base program or a delta module: we denote with `p` the name of the base program, and extend $L$.`activation` by convention, stating that $L$.`activation`(`p`) = **true**. Second, the *projection* of a product line on a subset $S$ of its products is the product line obtained by restricting the $L$.`products` formula to describe only the products in $S$ and by ignoring delta modules that are never activated. Third, the following definitions introduce auxiliary structures and getters that are useful to type check an IF$\Delta$J product line.

**Definition 1 (FCST).** *A* Class Signature *(CS) is a class declaration deprived of the bodies of its methods, it comprises the name of the class and of its superclass, and a mapping from attribute names to types. A* Family Class Signature *(FCS) is a more liberal version of class signature that may extend multiple classes and associate more than one type to each attribute name. A* Family Class Signature table *(FCST) is a mapping that associates to each class name* `C` *an FCS for* `C`*. The subtyping relation <: described by an FCST can be cyclic. A* Class Signature Table *(CST) is a FCST that contains only class signatures and has an acyclic subtyping relation.*

To simplify the notation, except when stated otherwise, we always assume in the following a fixed product line $L = FM\ CK\ AB$. The FCST of $L$, denoted by $L$.`FCST`, contains for each class `C` declared in $AB$ all superclasses of `C` and all types of all attributes of `C`. Note that the FCST of $L$ is defined only in terms of $AB$ and it can be computed by a straightforward inspection of it. The FCST of a set of IFJ programs (or of a subset of $AB$) is defined similarly.

**Definition 2 (Getters on *AB*).** `add(C)` *is the set of modules that* add *the class* `C`*;* `remove(C)` *is the set of modules that* remove *the class* `C`*;* `modifyWEC(C)` *is the set of modules that modify the class* `C` without changing its **extends** clause*;* `modifyAEC(C)` *is the set of modules that modify the class* `C` also by changing its **extends** clause*;* `modify(C)` *is* `modifyWEC(C)`$\cup$`modifyAEC(C)`*;* `add(C.a)` *is the set of modules that* add *the attribute* `C.a`*;* `remove(C.a)` *is the set of modules that* remove *the attribute* `C.a`*;* `modify(C.a)` *is the set of modules that* modify *the attribute* `C.a`*;* `replace(C.m)` *is the set of modules that modify the method* `C.m` *without using calls to* `original` *(i.e.,* replace *its body); and* `wrap(C.m)` *is the set of modules that modify the method* `C.m` *by also using calls to* `original` *(i.e.,* wrap *its body).*

**Definition 3 (Getter on *FM* and *CK*).** *Let $\Phi$ be extended to include module names* `d` *as propositional variables. The formula* $L.\mathtt{FMandCK} \triangleq L.\mathtt{products} \wedge \bigwedge_{\mathtt{d}}(\mathtt{d} \Leftrightarrow L.\mathtt{activation}(\mathtt{d}))$ *specifies the products and binds each variable* `d` *to the activation condition of module* `d` *(i.e., it specifies which modules are activated for each product).*[1]

## 3 Two Delta-Oriented Programming Guidelines

The first guideline is to avoid *useless operations*, i.e., declarations in $P$ and **adds** or **modifies** in $\overline{\Delta}$ that introduce code that is never present in any of the variants.

**G1** Ensure that the product line does not contain useless operations.

For instance, in the product line obtained by projecting the EPL on the the products described by ¬fToString, the declarations of the methods with name toString in the base program and in the **adds** class operation in the delta module dAdd are useless. The notion of useless operation is formalized as follows (thus making Guideline G1 automatically checkable).

[1] The last occurrence of `d` in $L$.`FMandCK` is not used as a variable: it is used as argument of the map $L$.`activation` to obtain the activation condition of module `d`.

**Definition 4 (Useless operation and module).** *The declaration, addition or modification of an attribute* $\mathtt{C.a}$ *in a module* $\mathtt{d}$ *is* useless *iff the formula* $(L.\mathtt{FMandCK}\wedge\mathtt{d}) \Rightarrow \bigvee_{\mathtt{d}'} \mathtt{d}'$ $\quad$ (*with* $\mathtt{d}' \in \mathtt{remove(C.a)}\cup\mathtt{remove(C)}\cup\mathtt{replace(C.a)}$ *and* $(\mathtt{d} <_L \mathtt{d}')$) *is valid. An* **extends** *clause introduced in a class* $\mathtt{C}$ *by a module* $\mathtt{d}$ *is* useless *iff the formula* $(L.\mathtt{FMandCK}\wedge\mathtt{d}) \Rightarrow \bigvee_{\mathtt{d}'} \mathtt{d}'$ $\quad$ (*with* $\mathtt{d}' \in \mathtt{remove(C)}\cup\mathtt{modifyAEC(C)}$ *and* $(\mathtt{d} <_L \mathtt{d}')$) *is valid. A module* $\mathtt{d}$ *is* useless *iff* $L.\mathtt{products} \Rightarrow \neg L.\mathtt{activation(d)}$ *is valid.*

The second guideline is to have consistent declarations over the whole SPL (the FOP case-studies presented in [22] adhere to this guideline). For IF$\Delta$J (since IFJ has no method overloading and field shadowing), this means that two declarations of the same attribute (of the same class) in two different modules must have the same type.[2] We call this property *type-uniformity*. It can be straightforwardly formalized by exploiting the family class signature table of the product line.

**Definition 5 (Type-uniformity).** *A FCST* $\mathtt{FCST}$ *is* type-uniform *iff:*

- $\forall\mathtt{C} \in \mathrm{dom}(\mathtt{FCST})$, $\forall\mathtt{a} \in \mathrm{dom}(\mathtt{FCST(C)})$ *the set* $\mathtt{FCST(C.a)}$ *is a singleton; and*
- $\forall\mathtt{C}_1,\mathtt{C}_2,\mathtt{C}_3 \in \mathrm{dom}(\mathtt{FCST})$ *such that* $\mathtt{C}_1 <: \mathtt{C}_2$ *and* $\mathtt{C}_1 <: \mathtt{C}_3$*, we have:* $\forall\mathtt{a} \in \mathrm{dom}(\mathtt{FCST}(\mathtt{C}_2)) \cap \mathrm{dom}(\mathtt{FCST}(\mathtt{C}_3))$, $\mathtt{FCST}(\mathtt{C}_2.\mathtt{a}) = \mathtt{FCST}(\mathtt{C}_3.\mathtt{a})$

*An IF$\Delta$J product line (or a subset of its artifact base, or a set of IFJ programs) is* type-uniform *iff its FCST is type-uniform.*

Our second guideline is thus stated as follows (and it can automatically be checked by a straightforward inspection of the FCST).

**G2** Ensure that the product line is type-uniform.

The EPL is not type-uniform, because of the method `eval` of class `Exp`, that is added with two different types by delta modules `dEval1` and `dEval2`, respectively. Instead, both its two projections respectively described by the mutually exclusive features fEval1 and fEval2 are type-uniform.

We say that an IF$\Delta$J product line is *variant-type-uniform* to mean that: (i) its variants can be generated; and (ii) the FCST of the set of its variants is type-uniform. The following proposition illustrate how type-uniformity relates to variant-type-uniformity.

**Proposition 1.** *Let L be an IF$\Delta$J product line such that its variants can be generated. If L is type-uniform, then it is variant-type-uniform. If L satisfies Guideline G1 and is variant-type-uniform, then it is type-uniform.*

## 4 Type Checking for Type-Uniform IF$\Delta$J

This section presents a version of the type checking approach tuned to exploit Guideline G2 and states its correctness and completeness. Type-uniformity makes type checking more efficient. The approach is modularized in three independent parts: *partial typing*, *applicability*, and *dependency*. All the parts rely on the FCST of the product line (see Definition 1).

[2] Note that, since the type system of IFJ is nominal, a class may have different sets of attributes in different variants.

**Product Line Partial Typing** Partial typing checks that all fields, methods and classes in $AB$ type-check with respect to the product line FCST (i.e., with respect to declarations made in $AB$). Partial typing does not use any knowledge about valid feature combinations (it does not use $FM$ and $CK$), so it does not guarantee that variants are well-typed, as delta modules may be activated or not. However, it guarantees that variants that have their inner dependencies satisfied (i.e., all used classes, methods and fields are declared) are well-typed.

The IF$\Delta$J partial-type-system is a straightforward extension of the (standard) IFJ type system [5] that: (i) includes rules for the new syntactic constructs of IF$\Delta$J; (ii) checks well-typedness with respect to the product line FCST (instead of the program CST); and (iii) allows to introduce a same class or attribute in different modules of $AB$ (e.g., a class of name `C` may be added by different delta modules).

The projection of the EPL described by feature `fEval1` is type-uniform. Its artifact base (which is obtained from the EPL artifact base in Fig. 1 by dropping the delta module `dEval2`) is accepted by partial typing, even if the method `Exp.eval` might not be available in some variant (in principle the delta module `dEval1` might not be selected). This is because the way the method `Exp.eval` is used in the method `Test.test` in the base program is correct with respect to its definition in the delta module `dEval1` (it takes no parameters and returns a `Lit` object).

**Product Line Applicability** Applicability ensures that variants can actually be generated (variant generation fails if, e.g., a delta module that adds a class `C` is applied to an intermediate variant that already contains a class named `C`). It is formalized by a constraint ensuring that, during variant generation, each delta operation is applied to an intermediate variant on which that operation is defined. For instance, for adding a class `C`, this class must not be present in the intermediate variant (either it never was added, or it was removed at some point). The applicability constraint comprises three validation parts: element addition (either a class or an attribute), element removal, and element modification.

In the following we use $\rho$ to denote either a class name `C` or a fully qualified attribute name `C.a`. The constraint for checking that an element $\rho$ can be added is as follows:

$$\mathtt{appADD}(\rho) \triangleq \bigwedge_{\mathtt{d}\neq\mathtt{d}'} \mathtt{d} \wedge \mathtt{d}' \Rightarrow \bigvee_{\mathtt{d}''} \mathtt{d}'' \quad \text{with} \begin{cases} \mathtt{d}, \mathtt{d}' \in \mathtt{add}(\rho),\ \mathtt{d}'' \in \mathtt{remove}(\rho) \\ \text{and } \mathtt{d} <_L \mathtt{d}'' <_L \mathtt{d}' \end{cases}$$

It ensures that all **adds** operations are performed on a partial variant that does not contain the added element: basically, it requires that if two delta modules $\mathtt{d}$ and $\mathtt{d}'$ add the same element, then there must be another delta module $\mathtt{d}''$ in between that removes it.

The constraint for removal of an element $\rho$ is slightly more complex:

$$\mathtt{appRM}(\rho) \triangleq \bigwedge_{\mathtt{d}} \mathtt{d} \Rightarrow (\bigvee_{\mathtt{d}_1} \mathtt{d}_1 \wedge \bigwedge_{\mathtt{d}'} (\mathtt{d}' \Rightarrow \bigvee_{\mathtt{d}_2} \mathtt{d}_2)) \quad \text{with} \begin{cases} \mathtt{d}, \mathtt{d}' \in \mathtt{remove}(\rho),\ \mathtt{d}_1, \mathtt{d}_2 \in \mathtt{add}(\rho) \\ \mathtt{d}_1 <_L \mathtt{d} <_L \mathtt{d}_2 <_L \mathtt{d}' \end{cases}$$

In comprises two parts: the first part ($\mathtt{d} \Rightarrow \bigvee_{\mathtt{d}_1}$) ensures that the element $\rho$ is added to the partial variant (by some $\mathtt{d}_1$) before it is removed (by $\mathtt{d}$); the second part ensures that if two delta modules $\mathtt{d}$ and $\mathtt{d}'$ remove $\rho$, then there is another delta module $\mathtt{d}_2$ in between that adds it.

The constraint for modification of an element $\rho$ simply ensures that $\rho$ is present for the modification:

$$\mathtt{appMOD}(\rho) \triangleq \bigwedge_{\mathtt{d}} \mathtt{d} \Rightarrow (\bigvee_{\mathtt{d}'} \mathtt{d}' \wedge \bigwedge_{\mathtt{d}''} \neg \mathtt{d}'') \quad \text{with } \begin{cases} \mathtt{d} \in \mathtt{modify}(\rho),\ \mathtt{d}'' \in \mathtt{remove}(\rho) \\ \mathtt{d}' \in \mathtt{add}(\rho),\ \mathtt{d}' <_L \mathtt{d}'' <_L \mathtt{d} \end{cases}$$

Basically, it checks that there is a delta module $\mathtt{d}'$ that adds the element before it is modified by $\mathtt{d}$, and that there is no delta module $\mathtt{d}''$ in between that removes it.

The formula $\mathtt{app}(L) \triangleq \bigwedge_{\rho \in \mathtt{add}(L)} \mathtt{appADD}(\rho) \wedge \mathtt{appRM}(\rho) \wedge \mathtt{appMOD}(\rho)$ combines the constraints described above, and the formula $\mathtt{ac}(L) \triangleq L.\mathtt{FMandCK} \Rightarrow \mathtt{app}(L)$ associates to each product of $L$ its applicability constraints. Applicability-consistency (i.e., the fact that variants of $L$ can be generated) is therefore formalized as follows.

**Definition 6 (Applicability-consistency).** *A product line $L$ is* applicability-consistent *iff the formula $\mathtt{ac}(L)$ is valid.*

**Product Line Dependency** Dependency ensures that no generated variant has a missing dependency, which can be straightforwardly expressed by means of constraints on attributes and classes. For instance, the dependencies induced by "class $\mathtt{C}$ extends class $\mathtt{C}'$" could be encoded with the constraint $\mathtt{decl}(\mathtt{C}) \Rightarrow (\mathtt{decl}(\mathtt{C}') \wedge \neg\mathtt{sub}(\mathtt{C}', \mathtt{C}))$, as the declaration of $\mathtt{C}$ requires that the declaration of $\mathtt{C}'$ is present and that $\mathtt{C}'$ is not a subtype of $\mathtt{C}$ (to ensure that the inheritance graph has no loops). In DOP, since each declaration is made in a module that can be activated or not, dependency constraints must be lifted at the module level. For instance, if the fact that $\mathtt{C}$ extends $\mathtt{C}'$ is declared in the module $\mathtt{d}$, then the previous constraint becomes: $\mathtt{d} \Rightarrow \neg\mathtt{rm}(\mathtt{d}, \mathtt{C}) \Rightarrow \neg\mathtt{modifyEC}(\mathtt{d}, \mathtt{C}) \Rightarrow (\mathtt{decl}(\mathtt{C}') \wedge \neg\mathtt{sub}(\mathtt{C}', \mathtt{C}))$, basically stating that if the module $\mathtt{d}$ is activated and no other module that removes $\mathtt{C}$ or changes its **extends** clause is activated afterward, then the class $\mathtt{C}'$ must be present in the generated variant and must not be a subtype of $\mathtt{C}$.

The product line dependency constraint is generated by exploiting the rules in Figures 3 and 4, which infer a dependency constraint for each expression and declaration, respectively. It is based on the following atomic constraints: $\mathtt{rm}(\mathtt{d}, \mathtt{C})$ (resp. $\mathtt{rm}(\mathtt{d}, \mathtt{C.a})$) ensures that the class $\mathtt{C}$ (resp. attribute $\mathtt{C.a}$) added by the delta module $\mathtt{d}$ will be removed afterward; $\mathtt{modifyEC}(\mathtt{d}, \mathtt{C})$ ensures that the class $\mathtt{C}$ added or modified by the delta module $\mathtt{d}$ will have its **extends** clause modified by another delta module afterward; $\mathtt{replace}(\mathtt{d}, \mathtt{C.m})$ ensures that the method $\mathtt{C.m}$ added or modified by the delta module $\mathtt{d}$ will be replaced by another delta module afterward; $\mathtt{sub}(T, \mathtt{C}')$ ensures that $T$ (either a class or **null**) is a subtype of $\mathtt{C}'$; $\mathtt{decl}(\mathtt{C})$ (resp. $\mathtt{decl}(\mathtt{C.a})$) ensures that the class $\mathtt{C}$ (resp. the attribute $\mathtt{a}$) is present in the generated variant (resp. is an attribute of the class $\mathtt{C}$, possibly through inheritance).

Dependency generation rules for expressions perform a type analysis to know what is the type of each expression, which is used to compute the appropriate

E:VAR
$$\frac{\Gamma(\mathtt{x}) = \mathtt{C}}{\Gamma \vdash \mathtt{x} : \mathtt{C} \mid \mathbf{true}}$$

E:FIELD
$$\frac{\Gamma \vdash e : \mathtt{C} \mid \Phi \qquad \mathtt{FCST}(\mathtt{C.f}) = \mathtt{C}'}{\Gamma \vdash e.\mathtt{f} : \mathtt{C}' \mid \Phi \wedge \mathtt{decl}(\mathtt{C.f})}$$

E:NULL
$$\Gamma \vdash \textsf{null} : \bot \mid \mathbf{true}$$

E:METH
$$\frac{\begin{array}{c}\Gamma \vdash e : \mathtt{C} \mid \Phi \qquad \mathtt{FCST}(\mathtt{C.m}) = \mathtt{C}'(\mathtt{C}_1, \ldots, \mathtt{C}_n) \\ \Gamma \vdash e_i : T_i \mid \Phi_i \qquad \Phi'_i = \mathtt{sub}(T_i, \mathtt{C}_i)\end{array}}{\Gamma \vdash e.\mathtt{m}(e_1, \ldots, e_n) : \mathtt{C}' \mid \bigwedge_i (\Phi_i \wedge \Phi'_i) \wedge \Phi \wedge \mathtt{decl}(\mathtt{C.m})}$$

D:NEW
$$\Gamma \vdash \textsf{new}\ \mathtt{C}() : \mathtt{C} \mid \mathtt{decl}(\mathtt{C})$$

E:CAST
$$\frac{\Gamma \vdash e : T \mid \Phi}{\Gamma \vdash (\mathtt{C})e : \mathtt{C} \mid \Phi \wedge (\mathtt{sub}(T, \mathtt{C}) \vee \mathtt{sub}(\mathtt{C}, T))}$$

E:ASSIGN
$$\frac{\Gamma \vdash e.\mathtt{f} : \mathtt{C} \mid \Phi_1 \qquad \Gamma \vdash e' : T \mid \Phi_2}{\Gamma \vdash e.\mathtt{f} = e' : \mathtt{C} \mid \Phi_1 \wedge \Phi_2 \wedge \mathtt{sub}(T, \mathtt{C})}$$

**Fig. 3.** Dependency Generation for Expressions

D:FIELD
$$\mathtt{d}, \mathtt{C} \vdash \mathtt{C}'\ \mathtt{f} : \neg\mathtt{rm}(\mathtt{d}, \mathtt{C.f}) \Rightarrow \mathtt{decl}(\mathtt{C}')$$

D:METH
$$\frac{\mathtt{this} : \mathtt{C}; \mathtt{x_i} : \mathtt{C_i} \vdash e : \mathtt{C}' \mid \Phi}{\begin{array}{c}\mathtt{d}, \mathtt{C} \vdash \mathtt{C}_0\ \mathtt{m}(\mathtt{C}_1\ \mathtt{x}_1, \ldots, \mathtt{C}_n\ \mathtt{x}_n)\ \{\textsf{return}\ e\} \\ : \neg(\mathtt{rm}(\mathtt{d}, \mathtt{C.m}) \vee \mathtt{replace}(\mathtt{d}, \mathtt{C.m})) \Rightarrow (\bigwedge_i \mathtt{decl}(\mathtt{C}_i) \wedge \Phi \wedge \mathtt{sub}(\mathtt{C}', \mathtt{C}_0))\end{array}}$$

D:CLASS
$$\frac{\mathtt{d}, \mathtt{C} \vdash AD_i : \Phi_i}{\begin{array}{c}\mathtt{d} \vdash \textsf{class}\ \mathtt{C}\ \textsf{extends}\ \mathtt{C}'\ \{AD_1 \ldots FD_n\} \\ : \neg\mathtt{rm}(\mathtt{d}, \mathtt{C}) \Rightarrow \bigwedge_i \Phi_i \wedge (\neg\mathtt{modifyEC}(\mathtt{d}, \mathtt{C}) \Rightarrow \mathtt{decl}(\mathtt{C}') \wedge \neg\mathtt{sub}(\mathtt{C}', \mathtt{C}))\end{array}}$$

D:MODMD
$$\frac{\mathtt{d}, \mathtt{C} \vdash MD : \Phi}{\mathtt{d}, \mathtt{C} \vdash \textsf{modifies}\ MD : \Phi}$$

D:ADDATT
$$\frac{\mathtt{d}, \mathtt{C} \vdash AD : \Phi}{\mathtt{d}, \mathtt{C} \vdash \textsf{adds}\ AD : \Phi}$$

D:RMATT
$$\mathtt{d}, \mathtt{C} \vdash \textsf{removes}\ \mathtt{a} : \mathbf{true}$$

D:RMCLASS
$$\mathtt{d} \vdash \textsf{removes}\ \mathtt{C} : \mathbf{true}$$

D:ADDCLASS
$$\frac{\mathtt{d} \vdash CD : \Phi}{\mathtt{d} \vdash \textsf{adds}\ CD : \Phi}$$

D:MODCLASS1
$$\frac{\mathtt{d}, \mathtt{C} \vdash AO_i : \Phi_i}{\begin{array}{c}\mathtt{d} \vdash \textsf{modifies}\ \mathtt{C}\ \{AO_1 \ldots AO_n\} \\ : \neg\mathtt{rm}(\mathtt{d}, \mathtt{C}) \Rightarrow \bigwedge_i \Phi_i\end{array}}$$

D:MODCLASS2
$$\frac{\mathtt{d}, \mathtt{C} \vdash AO_i : \Phi_i}{\begin{array}{c}\mathtt{d} \vdash \textsf{modifies}\ \mathtt{C}\ \textsf{extends}\ \mathtt{C}'\{AO_1 \ldots AO_n\} \\ : \neg\mathtt{rm}(\mathtt{d}, \mathtt{C}) \Rightarrow \bigwedge_i \Phi_i \wedge (\neg\mathtt{modifyEC}(\mathtt{d}, \mathtt{C}) \Rightarrow \mathtt{decl}(\mathtt{C}') \wedge \neg\mathtt{sub}(\mathtt{C}', \mathtt{C}))\end{array}}$$

D:DELTA
$$\frac{\mathtt{d} \vdash CO_i : \Phi_i}{\vdash \textsf{delta}\ \mathtt{d}\ \{CO_1 \ldots CO_n\} : \mathtt{d} \Rightarrow \bigwedge_i \Phi_i}$$

D:P
$$\frac{\mathbf{true} \vdash CD_i : \Phi_i \qquad \vdash \Delta_j : \Phi'_j}{\vdash \Phi\ \Delta_1 \ldots \Delta_n\ CD_1 \ldots CD_m : \bigwedge_i \Phi_i \wedge \bigwedge_j \Phi'_j}$$

**Fig. 4.** Dependency Generation for Declarations

dependency. They have judgments of the form $\Gamma \vdash e : T \mid \Phi$, where: $\Gamma$ is an environment giving the type of each variable; $e$ is the parsed expression; $T$ is its type; and $\Phi$ is the generated dependency constraint. The rules for expressions are quite direct: accessing a variable (rule (E:VAR)) does not raise any dependency, while accessing a field requires for this field to be accessible (rule (E:FIELD)); method calls (rule (E:METH)) require that the method is accessible and that the parameters have a type consistent with the method's declaration; object creation requires for the class of the object to be defined (rule (E:NEW)); and **null** does not raise any dependency (rule (E:NULL)), while casting and assignment generate constraints ensuring that the right inheritance relation holds (rules (E:CAST) and (E:ASSIGN)).

Dependency generation rules for declarations have judgments of the form $\Omega \vdash A : \Phi$ where $\Omega$ can either be empty, $\mathtt{d}$ (meaning that we are parsing the content of the module $\mathtt{d}$), or $\mathtt{d}, \mathtt{C}$ (meaning that we are parsing the content of the class $\mathtt{C}$ inside $\mathtt{d}$); $A$ is the parsed declaration (e.g., an attribute, a class operation); and $\Phi$ is the generated constraint. Rules (D:Field) for field and (D:Meth) for method declarations are quite direct: if the attribute is not removed afterward, the dependencies it generates must be validated. The rule (D:Class) for class declaration is similar (if the class is not removed, its inner dependencies must be validated), with an additional clause for the **extends** clauses (as previously discussed). Rules (D:ModMD) for modifying methods and (D:AddAtt) and (D:AddClass) for adding attributes and classes simply forward the constraints generated from the inner declaration, while removing an attribute or a class (rules (D:RmAtt) and (D:RmClass)) does not generate any dependency. The rules (D:AddClass1) and (D:AddClass2) for modifying a class are simple variations on the rule for class declaration. Finally, the dependencies of a delta module body are activated only if the delta module is activated (rule (D:Delta)), and the dependencies of a whole program is the conjunction of the dependencies of all its parts (rule (D:P)). The resulting constraint thus has the form $\bigwedge_i \mathtt{d}_i \Rightarrow \Phi_i$, giving for all module $\mathtt{d}_i$ its dependencies $\Phi_i$. Let then $\mathtt{dep}(L)$ be the constraint generated for the product line $L$. The formula $\mathtt{dc}(L) \triangleq L.\mathtt{FMandCK} \Rightarrow \mathtt{dep}(L)$ associates to each product of $L$ its dependency constraints. Dependency-consistency (i.e., variants of $L$ have all their dependencies fulfilled) is therefore formalized as follows.

**Definition 7 (Dependency-consistency).** *A product line $L$ is* dependency-consistent *iff the formula $\mathtt{dc}(L)$ is valid.*

**Correctness and Completeness of the Approach** The following theorem states that, if the product line follows Guideline G2, then the presented IF$\Delta$J product line type checking approach is correct with respect to generating variants and checking them using the IFJ type system. The approach is complete (i.e., if the check performed by the approach fails then at least one variant is not a well-typed IFJ program) if also Guideline G1 is followed.

**Theorem 1.** *Let $L$ be a type-uniform product line. Consider the properties:*
*i. $L$ is well partially-typed, applicability- and dependency-consistent.*
*ii. Variants of $L$ can be generated and are well-typed IFJ programs.*
*Then: (i) implies (ii); and if $L$ has no useless operations then (ii) implies (i).*

## 5 Type Checking for IF$\Delta$J without Guidelines

In this section we outline how the type checking approach presented in Section 4 can be tuned to non type-uniform product lines (i.e., not to exploit any guidelines). This modification is quite straightforward, although it involves many technical details. Partial typing must be adapted since the product line FCST maps attribute names to sets of types with possibly more than one element, and expressions can have more than one type. E.g., a method call expression $e.\mathtt{m}(\overline{e'})$ can use any declaration of the method $\mathtt{C.m}$ (considering that $e$ is typed $\mathtt{C}$) whose type accepts a combination of types of the call's arguments. So partial typing may carry a combinatorial explosion.

Applicability does not need any modification to analyze non-uniform programs. This is due to the fact that the applicability criteria focuses on the interplay between delta operations and do not consider attribute types.

Dependency is the part that changes more: it now has to be type-aware, and thus subsumes partial typing. We illustrate it on the rule that generates the dependency for field usage (second rule in Fig. 3). This rule must be extended in two ways to manage non-uniform programs: (i) $e$ can have more than one type; (ii) the field type lookup `FCST(C.f)` can return different possible types for `C.f`, depending on which modules are activated. Consequently, the dependency generation judgment for expressions now has the form $\Gamma \vdash e : [\Phi_i \mapsto T_i]_{i \in I}$ where $T_i$ are the possible types of $e$, and $\Phi_i$ is the condition (i.e. which module must or must not be activated) for $e$ to have the type $T_i$ in the final product. Hence, the rule becomes as displayed on the right,

$$\frac{\Gamma \vdash e : [\Phi_i \mapsto \mathtt{C}_i]_i \cup [\Phi_{i'} \mapsto \bot]_{i'} \qquad \mathtt{FCST}(\mathtt{C}_i.\mathtt{f}) = [\Phi_{i,j} \mapsto \mathtt{C}_{i,j}]_j}{\Gamma \vdash e.\mathtt{f} : [\Phi_i \wedge \Phi_{i,j} \mapsto \mathtt{C}_{i,j}]_{i,j}}$$

where $\Phi_{i,j}$ is the formula that enforces that the field `f` accessible from the class $\mathtt{C}_i$ has the type $\mathtt{C}_{i,j}$ in the final product.

Correctness and completeness are stated as in Theorem 1 by dropping the assumption that the product line is type-uniform.

## 6 Three other Guidelines

Our type-checking approach is modularized in three parts: i) partial typing performs a preliminary type analysis that can be exploited by an IDE for prompt notification of type-errors and auto-completing code; ii) applicability ensures that variants can be generated; and iii) dependency completes the analysis done by the partial typing. The approach is tunable to exploit DOP guidelines that enforce structural regularities in product line implementation. In Section 4 we have presented a version tuned to exploit type-uniformity. In this section we briefly discuss three other automatically checkable guidelines (other useful guidelines could be devised).

First, whenever it is possible to enforce the following guideline (satisfied by the EPL), the dependency analysis can be simplified, as it is no longer needed to check the absence of inheritance loop in the generated variant (cf. dependency generation for class declaration and modification in Figure 4).

**G3** Ensure that the product line FCST subtyping relation is acyclic.

If a product line cannot be made variant-type-uniform, then guideline G2 cannot be enforced (see Proposition 1), and understanding the structure of the SPL may become an issue. The following guideline (satisfied by the EPL) aims at helping the understanding of an SPL implementation by decoupling the sources of non type-uniformity.

**G4** Ensure that, for all distinct modules $\mathtt{d}_1$ and $\mathtt{d}_2$, if the set comprising $\mathtt{d}_1$ and $\mathtt{d}_2$ is not type-uniform then their activation conditions are mutually exclusive.

Consider for instance a module `d` that declares an attribute `C.a` with a type $t$. Then, if the SPL follows G3, we are sure that each variant using `d` in its construction will have `C.a` typed $t$ when it contains this attribute.

We introduce our final guideline with the following consideration: implementing or modifying a product line involves editions of the feature model, the configuration knowledge and the artifact base that may affect only a subset of the products. For example, adding, removing or modifying a delta module `d` and its activation condition will affect only the products that activate `d`. Therefore, only the projection of the product line on the affected products needs to be re-analyzed. If such a projection is type-uniform, then the more efficient type checking technique of Section 4 can be used (even if the whole product line is not type-uniform). The following guideline naturally arises.

**G5** i) Ensure that the set of products is partitioned in such a way that: each part $S$ is type-uniform (i.e., the projection of the SPL on $S$ is type uniform), and the union of any two distinct parts is not type-uniform.
ii) If the number of parts of such a partition is "too big", then merge some of them to obtain a "small enough" partition where only one part is not type-uniform.

The goal of this guideline is to allow to use as much as possible the version of the approach presented in Section 4. For the EPL the partition that satisfies Guideline G5.i is unique: the two products with feature fEval1 and the four products with feature fEval2. However, in general, such a partition may be not unique and tool support for identifying a partition that satisfies G5.i and further conditions (e.g., having a minimal number of parts) or G5.ii and other conditions (e.g., the number of products in the non type-uniform part is as small as possible) would be valuable.

## 7 Related Work

Product line analysis approaches can be classified into three main categories [23]: *Product-based* analyses operate only on generated variants (or models of variants); *Family-based* analyses operate only on the AB by exploiting the FM and the CK to obtain results about all variants; *Feature-based* analyses operate on the building blocks of the different variants (feature modules in FOP and delta modules in DOP) in isolation (without using the FM and the CK) to derive results on all variants. We refer to [23] for a survey on product line type checking. Here we discuss previous type checking approaches for DOP [5, 8] and the two approaches for FOP that are closets to our proposal [9, 22].

The type checking approach for DOP in [5] comprises: a feature-based analysis that uses a constraint-based type system for IFJ to infer a type abstraction for each delta module; and a product-based step that uses these type abstractions to generate, for each product of the SPL, a type abstraction (of the associated variant) that is checked to establish whether the associated variant type checks. The approach of [5] is enhanced in [8] by introducing a family-based step that builds a product family generation tree which is then traversed in order to perform optimized generation and check of type abstractions of all variants. The approach proposed in this paper, which is feature-family-based, represents an achievement over [5, 8] since it does not require to iterate over the set of products (cf. Section 1) and supports earlier detection of errors via partial typing.

The paper [22] informally illustrates the implementation of a family-based approach for the AHEAD system [4]. The approach comprises: i) a family-feature-based step that computes for each class a stub (all stubs can be understood as a type-uniform FCST for the product line) and compiles each feature module in the context of all stubs (thus performing checks corresponding to our type-uniformity and partial-typing); and ii) a family-based step that infers a set of constraints that are combined with the FM to generate a formula (corresponding to our type-uniform applicability and dependency) whose satisfiability should imply that all variants successfully compile.

The paper [9] formalizes a feature-family-based approach for the Lightweight Feature Java (LFJ) calculus, which models FOP for the Lightweight Java (LJ) [21] calculus. The approach comprises: i) a feature-based step that uses a constraint-based type system for LFJ to analyze each feature module in isolation and infer a set of constraints for each feature module; and ii) a family-based step where the FM and the previously inferred constraints are used to generate a formula whose satisfiability implies that all variants type check. The applicability and dependency analyses presented in Section 5 provide an extension to DOP of these two steps. Moreover, our approach provides partial typing for early error detection and is tunable to exploit different programming guidelines.

## 8 Conclusions and Future Work

We have proposed a modular and tunable approach for type checking DOP product lines. A prototypical implementation is available [2] (currently only the version of the approach exploiting type-uniformity is supported).

In future work we plan to: implement our approach for both DeltaJ 1.5 [14] (a prototypical implementation of DOP that supports full Java 1.5) and ABS [12] (this would allow experimental comparison with the approaches of [5, 8], which have been implemented for ABS [1]); to develop case studies to evaluate the effectiveness of the approach and of the proposed guidelines; to investigate further DOP guidelines; and to develop tool support to allow the programmer to choose the guidelines to be automatically enforced. We also plan to investigate whether the proposed DOP guidelines (or other guidelines) could be useful for other kind of product line analyses. In particular, we would like to consider formal verification (proof systems for the verification of DOP product lines of Java programs have been recently proposed [10, 7]).

**Acknowledgements** We are grateful to Don Batory for clarifications about previous work on type checking FOP, and to Ina Schaefer and Thomas Thüm for discussions about how to classify SPL type checking approaches. We also thank the iFM 2016 anonymous reviewers for insightful comments and suggestions.

## References


1. https://github.com/abstools/abstools/tree/master/frontend/src/abs/frontend/delta.
2. https://github.com/gzoumix/IFDJTS.
3. D. Batory. Feature Models, Grammars, and Propositional Formulas. In *Proc. of SPLC 2005*, volume 3714 of *LNCS*, pages 7–20. Springer, 2005.
4. D. Batory, J. N. Sarvela, and A. Rauschmayer. Scaling step-wise refinement. In *Proc. of ICSE 2003*, pages 187–197. IEEE, 2003.
5. L. Bettini, F. Damiani, and I. Schaefer. Compositional type checking of delta-oriented software product lines. *Acta Informatica*, 50(2):77–122, 2013.
6. P. Clements and L. Northrop. *Software Product Lines: Practices and Patterns*. Addison Wesley Longman, 2001.
7. F. Damiani, O. Owe, J. Dovland, I. Schaefer, E. B. Johnsen, and I. C. Yu. A transformational proof system for delta-oriented programming. In *Proc. of SPLC 2012 - Volume 2*, pages 53–60. ACM, 2012.
8. F. Damiani and I. Schaefer. Family-based analysis of type safety for delta-oriented software product lines. In *Proc. of ISoLA*, volume 7609 of *LNCS*. Springer, 2012.
9. B. Delaware, W. R. Cook, and D. Batory. Fitting the pieces together: A machine-checked model of safe composition. In *Proc. of ESEC/FSE 2009*. ACM, 2009.
10. R. Hähnle and I. Schaefer. A Liskov Principle for Delta-Oriented Programming. In *Proc. of ISoLA*, volume 7609 of *LNCS*, pages 32–46. Springer, 2012.
11. A. Igarashi, B. Pierce, and P. Wadler. Featherweight Java: A minimal core calculus for Java and GJ. *ACM TOPLAS*, 23(3):396–450, 2001.
12. E. B. Johnsen, R. Hähnle, J. Schäfer, R. Schlatte, and M. Steffen. ABS: A core language for abstract behavioral specification. In *Formal Methods for Components and Objects*, volume 6957 of *LNCS*, pages 142–164. Springer, 2012.
13. C. Kästner, S. Apel, S. S. ur Rahman, M. Rosenmüller, D. Batory, and G. Saake. On the impact of the optional feature problem: analysis and case studies. In *Proc. of SPLC 2009*, pages 181–190, 2009.
14. J. Koscielny, S. Holthusen, I. Schaefer, S. Schulze, L. Bettini, and F. Damiani. Deltaj 1.5: Delta-oriented programming for java 1.5. In *Proc. of PPPJ 2014*, pages 63–74. ACM, 2014.
15. C. Krueger. Eliminating the Adoption Barrier. *IEEE Software*, 19(4):29–31, 2002.
16. M. Lienhardt and D. Clarke. Conflict detection in delta-oriented programming. In *ISoLA*, pages 178–192, 2012.
17. R. E. Lopez-Herrejon, D. S. Batory, and W. R. Cook. Evaluating Support for Features in Advanced Modularization Technologies. In *Proc. of ECOOP 2005*, volume 3586 of *LNCS*, pages 169–194. Springer, 2005.
18. I. Schaefer. Proc. of VaMoS '10. In *Intl. Workshop on Variability Modelling of Software-intensive Systems*, 2010.
19. I. Schaefer, L. Bettini, V. Bono, F. Damiani, and N. Tanzarella. Delta-oriented Programming of Software Product Lines. In *Proc. of SPLC 2010*, volume 6287 of *LNCS*, pages 77–91. Springer, 2010.
20. I. Schaefer and F. Damiani. Pure Delta-oriented Programming. In *Proc. of FOSD 2010*, pages 49–56. ACM, 2010.
21. R. Strniša, P. Sewell, and M. Parkinson. The Java module system: core design and semantic definition. In *Proc. of OOPSLA 2007*, pages 499–514. ACM, 2007.
22. S. Thaker, D. Batory, D. Kitchin, and W. Cook. Safe composition of product lines. In *Proc. of GPCE '07*, pages 95–104. ACM, 2007.
23. T. Thüm, S. Apel, C. Kästner, I. Schaefer, and G. Saake. A classification and survey of analysis strategies for software product lines. *ACM Comput. Surv.*, 2014.

## A Appendix

This appendix presents proofs and construction completing the main document: in Section A.1 we prove Proposition 1; in Section A.2 we prove Theorem 1; and in Section A.3 we present the complete set of rules for partial typing and dependency analysis for the version of the approach outlined in Section 5.

### A.1 Proof of Proposition 1

*Proof (Proposition 1).* Let note in this proof $\texttt{FCST}'$ the union of all the product line's variant's CST.

$(i) \Rightarrow (ii)$. By construction, for all variant, its CST is included in the FCST of the product line. Hence $\texttt{FCST}'$ is still contained in the FCST of the product line. By definition of type uniform, $\texttt{FCST}'$ is thus type uniform, which proves this case.

$(ii) \Rightarrow (i)$. If the product line does not containt useless operations, it means that all elements of its FCST is part of one of the product line's variant. Hence, it is part of this variant's CST. We thus have that $\texttt{FCST}'$ contains the FCST of the product line. By construction of the variant's CST, we thus have that $\texttt{FCST}'$ is equals to the FCST of the product line. Hence, by hypothesis, we have that the product line is type-uniform.

### A.2 Proof of Theorem 1

We base our proof on different preliminary definitions: Section A.2.1 briefly recalls the Type System for the IFJ language; Section A.2.2 we formally describes our Partial Typing algorithm; Section A.2.3 gives the formal definition of the different predicates used in our constraint generation algorithm. The actual proof, with preliminary lemmas, is presented in Section A.2.4.

**A.2.1 A Recollection of the Typing Rules for IFJ.** The Type System for IFJ is presented as usual with a set of rule described in Figure 5. As usual, these rules are based on the class signature table `CST` (see Definition 1) of the program.[3] This table is also completed with the builtin `Object` class. Moreover, the rules are also based on the *subtyping order* $<:$ which is constructed as the transitive closure of the **extends** relation declared in the input program, extended as follows: all classes are a supertype of the special type $\perp$ that types the **null** expression. Finally, we note $\texttt{C} \in \mathrm{dom}(\texttt{CST})$ if the class `C` is declared in `CST`, and `CST(C.a)` the lookup operation that gets the header of the attribute `a` accessible from the class `C` declared in `CST`. Note that this lookup operation follows the

[3] It is implicitely supposed that a table `CST` can be computed from a program only if a class is declared only once in it.

$$\textsc{T-Var}\ \frac{\Gamma(\mathtt{x}) = \mathtt{C}}{\mathtt{CST}, \Gamma \vdash \mathtt{x} : \mathtt{C}} \qquad \textsc{T-Null}\ \mathtt{CST}, \Gamma \vdash \textsf{null} : \bot \qquad \textsc{T-Access}\ \frac{\mathtt{CST}, \Gamma \vdash e : \mathtt{C} \qquad \mathtt{CST}(\mathtt{C}.\mathtt{f}) = \mathtt{C}'}{\mathtt{CST}, \Gamma \vdash e.\mathtt{f} : \mathtt{C}'}$$

$$\textsc{T-Call}\ \frac{\begin{array}{c}\mathtt{CST}, \Gamma \vdash e : \mathtt{C} \qquad \mathtt{CST}, \Gamma \vdash e_i : T_i \\ \mathtt{CST}(\mathtt{C}.\mathtt{m}) = \mathtt{C}_0(\mathtt{C}'_1, \ldots, \mathtt{C}'_n) \qquad T_i <: \mathtt{C}'_i\end{array}}{\mathtt{CST}, \Gamma \vdash e.\mathtt{m}(e_1, \ldots, e_n) : \mathtt{C}_0} \qquad \textsc{T-Assign}\ \frac{\begin{array}{c}\mathtt{CST}, \Gamma \vdash e.\mathtt{f} : \mathtt{C} \\ \mathtt{CST}, \Gamma \vdash e' : T \qquad T <: \mathtt{C}\end{array}}{\mathtt{CST}, \Gamma \vdash e.\mathtt{f} = e' : \mathtt{C}} \qquad \textsc{T-New}\ \frac{\mathtt{C} \in \mathrm{dom}(\mathtt{CST})}{\mathtt{CST}, \Gamma \vdash \textsf{new}\ \mathtt{C}() : \mathtt{C}}$$

$$\textsc{T-Cast}\ \frac{\mathtt{CST}, \Gamma \vdash e : T \qquad T <: \mathtt{C} \vee \mathtt{C} <: T}{\mathtt{CST}, \Gamma \vdash (\mathtt{C})e : \mathtt{C}} \qquad \textsc{T-Field}\ \frac{\mathtt{C}' \in \mathrm{dom}(\mathtt{CST}) \qquad \mathtt{CST}(\mathtt{C}.\mathtt{f}) = \mathtt{C}'}{\mathtt{CST}, \mathtt{C} \vdash \mathtt{C}'\ \mathtt{f}}$$

$$\textsc{T-Method}\ \frac{\begin{array}{c}\mathtt{CST}, \mathtt{this} : \mathtt{C}; \mathtt{x_i} : \mathtt{C_i} \vdash e : \mathtt{C}' \qquad \mathtt{C}' <: \mathtt{C}_0 \\ \mathtt{CST}(\mathtt{C}.\mathtt{m}) = \mathtt{C}_0(\mathtt{C}_1, \ldots, \mathtt{C}_n) \qquad \mathtt{C}_i \in \mathrm{dom}(\mathtt{CST})\end{array}}{\mathtt{CST}, \mathtt{C} \vdash \mathtt{C}_0\ \mathtt{m}(\mathtt{C}_1\ \mathtt{x}_1, \ldots, \mathtt{C}_n\ \mathtt{x}_n)\ \{\textsf{return}\ e\}} \qquad \textsc{T-Class}\ \frac{\mathtt{CST}, \mathtt{C} \vdash AD_i \qquad \mathtt{C}' \in \mathrm{dom}(\mathtt{CST}) \qquad \mathtt{C}' \not<: \mathtt{C}}{\mathtt{CST} \vdash \textsf{class}\ \mathtt{C}\ \textsf{extends}\ \mathtt{C}'\ \{AD_1 \ldots AD_n\}}$$

$$\textsc{T-Program}\ \frac{P = CD_1\ \ldots\ CD_n \qquad \mathtt{CST}[P] \vdash CD_i}{\vdash P}$$

**Fig. 5.** IFJ Typing Rules

inheritance hierarchy to find the field or method declaration in the super class if it is not locally declared.

Our set of rules starts with the rules to type expressions which use statements of the form $\mathtt{CST}, \Gamma \vdash e : T$ where: CST is the class signature table of the considered program; $\Gamma$ is the typing environment, mapping all declared variables to their declared type; $e$ is the typed expression; and $T$ is the computed type of $e$, which can either be a class name C, or the type of null $\bot$. These rules are quite standard. Rule (T-Var) simply states that the type of a variable is the one used in its declaration (and stored in $\Gamma$). Rule (T-Null) states that the type of null is the special type $\bot$. Rule (T-Access) first computes the type C of $e$ (it must be a class), and lookup in the table CST to get the type of its field f. Rule (T-Call) first compute the types of the calee and all the method call's arguments, lookup the signature of the method in CST, and matches it against the arguments' types. Rule (T-Assign) computes the type of the specified field $e$.f, the type of the expression $e'$ and matches them. The type of the whole expression then is the one of the assigned field. Rule (T-New) simply types the **new** operation, checking that the specified class is declared in the program. Finally, rule (T-Cast) checks that a cast operations corresponds to an up- or down-cast.

The two typing rules for attribute declarations have statements of the form $\mathtt{CST}, \mathtt{C} \vdash AD$: here, C is the name of the class in which the attribute is declared. Rule (T-Field) checks that the type of the field is declared and corresponds to what is stated in the table CST. Rule (T-Method) first controls that the types used in the method's header are declared in CST, types its body with the right typing environment $\Gamma$, and checks that the resulting type matches the declared returned type of the method.

Rule (T-Class) typing a class declaration simply checks that the body of the class is well typed, that its super class exists and that there are no loops in the inheritance relation, and rule (T-Program) type checks a full program $P$ by typing all its declared class using its signature class table $\mathtt{CST}[P]$.

Finally, we conclude this section by defining when a program is well typed:

**Definition 8.** *A program $P$ is* well typed *iff the statement $\vdash P$ holds.*

**A.2.2 Rules for Partial Typing in IF$\Delta$J.** We present the rules for partial typing in Figure 6. Partial typing statements for expressions have the form $\mathtt{FCST}, \Gamma \vdash e : T$ where: FCST is the family class signature table containing all the declarations made in the program; $\Gamma$ is the typing environment giving the type of all the declared variables; $e$ is the typed expression; and $T$ is its computed type. These rules are very standard, with only few things to note: i) when typing field accesses or method calls, we enforce the base expression to be typed with a class name, thus forbidding expressions of the form **null**.f; note that this also ensures, given that FCST is uniform, that the field's type (or method's signature) is unique. ii) to enforce the fact that the presented rules correspond to a deterministic algorithm, we did not write a dedicated rule for subtyping, instead we included it in the rules for assignment and method call.

Partial typing statements for declarations can take three forms: one to type attributes and delta operations on attributes; one to type classes and delta operations on classes; and one to type delta modules and programs. Partial typing statements for (delta operations on) attributes have the form $\mathtt{CST}, \mathtt{C} \vdash A$ where: FCST is the family class signature table containing all the declarations made in the program; C is the class in which this attribute is being declared or modified; $A$ (either $FD$, $MD$ or $AO$) is the typed attribute. Partial typing statements for (delta operations on) classes have the same form as for attributes, except for the class name and partial typing statements for delta modules and programs have the same form as for classes, except that CST is replaced by the more precise delta module signature table DMST, to ensure while checking the declaration of a delta module, that all the used delta modules are declared. Most of the rules are quite direct and basically ensures that everything is declared as stated in the table (FCST or DMST), and that the code of methods is well typed. The only non-classic rules is for the method modification, in which we extend the FCST by giving a signature to the special method original: the same as the one of the method we are currently checking.

Finally, we conclude this section by defining when a IF$\Delta$J software product line is well partially typed:

**Definition 9.** *A product line $L$ is* well partially typed *iff it is type-uniform and the statement $\vdash L$ holds.*

**A.2.3 Definition of our Type System base Predicates.** We present in this section a full definition of the base constraints introduced in Section 4. To simplify the presentation of these constraints, we assume in the following, as

$$\frac{\Gamma(\mathtt{x}) = \mathtt{C}}{\mathtt{FCST}, \Gamma \vdash \mathtt{x} : \mathtt{C}}\ \textsc{PT-Var} \qquad \mathtt{FCST}, \Gamma \vdash \textbf{null} : \bot\ \textsc{PT-Null} \qquad \frac{\mathtt{FCST}, \Gamma \vdash e : \mathtt{C} \quad \mathtt{FCST}(\mathtt{C}.\mathtt{f}) = \mathtt{C}'}{\mathtt{FCST}, \Gamma \vdash e.\mathtt{f} : \mathtt{C}'}\ \textsc{PT-Access}$$

$$\frac{\mathtt{FCST}, \Gamma \vdash e : \mathtt{C} \quad \mathtt{FCST}, \Gamma \vdash e_i : T_i \quad \mathtt{FCST}(\mathtt{C}.\mathtt{m}) = \mathtt{C}_0(\mathtt{C}'_1, \dots, \mathtt{C}'_n) \quad T_i <: \mathtt{C}'_i}{\mathtt{FCST}, \Gamma \vdash e.\mathtt{m}(e_1, \dots, e_n) : \mathtt{C}_0}\ \textsc{PT-Call}$$

$$\frac{\mathtt{FCST}, \Gamma \vdash e.\mathtt{f} : \mathtt{C} \quad \mathtt{FCST}, \Gamma \vdash e' : T \quad T <: \mathtt{C}}{\mathtt{FCST}, \Gamma \vdash e.\mathtt{f} = e' : \mathtt{C}}\ \textsc{PT-Assign} \qquad \frac{\mathtt{C} \in \mathrm{dom}(\mathtt{FCST})}{\mathtt{FCST}, \Gamma \vdash \textbf{new}\ \mathtt{C}() : \mathtt{C}}\ \textsc{PT-New}$$

$$\frac{\mathtt{FCST}, \Gamma \vdash e : T \quad T <: \mathtt{C} \lor \mathtt{C} <: T}{\mathtt{FCST}, \Gamma \vdash (\mathtt{C})e : \mathtt{C}}\ \textsc{PT-Cast} \qquad \frac{\mathtt{C}' \in \mathrm{dom}(\mathtt{FCST}) \quad \mathtt{FCST}(\mathtt{C}.\mathtt{f}) = \mathtt{C}'}{\mathtt{FCST}, \mathtt{C} \vdash \mathtt{C}'\ \mathtt{f}}\ \textsc{PT-Field}$$

$$\frac{\mathtt{FCST}, \mathtt{this} : \mathtt{C}; \mathtt{x_i} : \mathtt{C_i} \vdash e : \mathtt{C}' \quad \mathtt{C}' <: \mathtt{C}_0 \quad \mathtt{FCST}(\mathtt{C})(\mathtt{m}) = \mathtt{C}_0(\mathtt{C}_1, \dots, \mathtt{C}_n) \quad \mathtt{C}_i \in \mathrm{dom}(\mathtt{FCST})}{\mathtt{FCST}, \mathtt{C} \vdash \mathtt{C}_0\ \mathtt{m}(\mathtt{C}_1\ \mathtt{x}_1, \dots, \mathtt{C}_n\ \mathtt{x}_n)\ \{\textbf{return}\ e\}}\ \textsc{PT-Method}$$

$$\frac{\mathtt{FCST}, \mathtt{C} \vdash AD_i \quad \mathtt{C}' \in \mathrm{dom}(\mathtt{FCST})}{\mathtt{FCST} \vdash \textbf{class}\ \mathtt{C}\ \textbf{extends}\ \mathtt{C}'\ \{AD_1 \dots AD_n\}}\ \textsc{PT-Class}$$

$$\frac{\mathtt{FCST}, \mathtt{C} \vdash AD}{\mathtt{FCST}, \mathtt{C} \vdash \textbf{adds}\ AD}\ \textsc{PT-A-Add} \qquad \frac{\mathtt{FCST}[\mathtt{C.original} \mapsto \mathtt{FCST}(\mathtt{C.name}(MD))], \mathtt{C} \vdash MD}{\mathtt{FCST}, \mathtt{C} \vdash \textbf{modifies}\ MD}\ \textsc{PT-A-Mod}$$

$$\frac{\mathtt{FCST} \vdash AO_i}{\mathtt{FCST} \vdash \textbf{modifies}\ \mathtt{C}\ \{AO_1 \dots AO_n\}}\ \textsc{PT-C-Mod} \qquad \frac{\mathtt{FCST} \vdash AO_i \quad \mathtt{C}' \in \mathrm{dom}(\mathtt{FCST})}{\mathtt{FCST} \vdash \textbf{modifies}\ \mathtt{C}\ \textbf{extends}\ \mathtt{C}'\{AO_1; \dots AO_n;\}}\ \textsc{PT-C-Mod+}$$

$$\frac{\mathtt{FCST} \vdash CD}{\mathtt{FCST} \vdash \textbf{adds}\ CD}\ \textsc{PT-C-Add} \qquad \frac{\mathtt{FCST} \vdash CO_i}{\mathtt{DMST} \vdash \textbf{delta}\ \mathtt{d}\ \{CO_1; \dots; CO_n\}}\ \textsc{PT-Delta} \qquad \mathtt{FCST} \vdash \textbf{removes}\ \mathtt{a}\ \textsc{PT-A-Del} \qquad \mathtt{FCST} \vdash \textbf{removes}\ \mathtt{C}\ \textsc{PT-C-Del}$$

$$\frac{L = \Pi\ \Delta_1 \dots \Delta_n\ CD_1 \dots CD_m \quad \mathtt{FCST}[L] \vdash CD_i \quad \mathtt{FCST}[L] \vdash \Delta_i}{\vdash L}\ \textsc{PT-Program}$$

**Fig. 6.** IF$\Delta$J Partial Typing Rules for Type-Uniform SPLs

stated in the end of Section 2, a fixed product line $L = P\ \overline{\Delta}\ \Pi$. Moreover, we use the following notations:

- $\mathtt{module}$ is the set of module names in the program $L$;
- $\mathtt{after}(\mathtt{d})$ is the set of modules that are applied after $\mathtt{d}$:

$$\mathtt{after}(\mathtt{d}) \triangleq \{\mathtt{d}' \mid \mathtt{d}' \in \mathtt{module} \land \mathtt{d} <_L \mathtt{d}'\}$$

- $\mathtt{ext}(\mathtt{C}, \mathtt{C}')$ is the set of module stating that $\mathtt{C}$ extends $\mathtt{C}'$;
- and $\mathtt{sub}(\mathtt{C}_0, \mathtt{C}_n)$ is the set of sequences $\mathtt{C}_0 \xrightarrow{\mathtt{d}_1} \mathtt{C}_1 \xrightarrow{\mathtt{d}_2} \dots \xrightarrow{\mathtt{d}_n} \mathtt{C}_n$ such that for all $1 \leq i \leq n$, $\mathtt{d}_i \in \mathtt{ext}(\mathtt{C}_{i-1}, \mathtt{C}_i)$.

**$\mathtt{rm}(\mathtt{d}, \mathtt{C})$ and $\mathtt{rm}(\mathtt{d}, \mathtt{C.a})$.** These formula ensures that the class $\mathtt{C}$ (resp. attribute $\mathtt{C.a}$) added by the module $\mathtt{d}$ will not be present in the generated variant. We define then as follows:

$$\texttt{rm}(\texttt{d},\texttt{Object}) \triangleq \texttt{rm}(\texttt{d},\bot) \triangleq \textbf{false}$$

$$\texttt{rm}(\texttt{d},\texttt{C}) \triangleq \bigvee_{\texttt{d}'} \texttt{d}' \qquad \text{with } \texttt{d}' \in \texttt{remove}(\texttt{C}) \cap \texttt{after}(\texttt{d})$$

$$\texttt{rm}(\texttt{d},\texttt{C.a}) \triangleq \bigvee_{\texttt{d}'} \texttt{d}' \qquad \text{with } \texttt{d}' \in (\texttt{remove}(\texttt{C.a}) \cup \texttt{remove}(\texttt{C})) \cap \texttt{after}(\texttt{d})$$

In addition to the classic base cases of the `Object` class and **null** type, this constraint simply ensures that at least one module $\texttt{d}'$ that removes the class `C` (or the attribute `C.a`) is applied after `d`.

**$\texttt{modifyEC}(\texttt{d},\texttt{C})$ and $\texttt{replace}(\texttt{d},\texttt{C.m})$.** These formula ensures that the class `C` (resp. the method `C.m`) added or modified by the module `d` will have its **extends** statement replaced (resp. its implementation entirely replaced) by another module afterward:

$$\texttt{modifyEC}(\texttt{d},\texttt{C}) \triangleq \bigvee_{\texttt{d}'} \texttt{d}' \quad \text{with } \texttt{d}' \in \texttt{modifyAEC}(\texttt{C}) \cap \texttt{after}(\texttt{d})$$

$$\texttt{replace}(\texttt{d},\texttt{C.m}) \triangleq \bigvee_{\texttt{d}'} \texttt{d}' \quad \text{with } \texttt{d}' \in \texttt{replace}(\texttt{C.m}) \cap \texttt{after}(\texttt{d})$$

**$\texttt{O}(\texttt{d},\texttt{C},\texttt{C}')$.** This formula wasn't introduced in Section 4, as it is only used to defined the two formula $\texttt{sub}(T,\texttt{C})$ and $\texttt{decl}(\texttt{C.a})$. This formula ensures that the inheritance statement `C` **extends** $\texttt{C}'$ declared in the module `d` is not replaced by some other module $\texttt{d}'$:

$$\texttt{O}(\texttt{d},\texttt{C},\texttt{C}') \triangleq \bigwedge_{\texttt{d}'} \neg\texttt{d}' \qquad \text{with } \texttt{d}' \in (\texttt{modifyAEC}(\texttt{C}) \setminus \texttt{ext}(\texttt{C},\texttt{C}')) \cap \texttt{after}(\texttt{d})$$

This constraint simply ensures that no module $\texttt{d}'$, that replace the **extends** statement on `C` to a class name different from $\texttt{C}'$, is activated after `d`.

**$\texttt{sub}(T,\texttt{C})$.** This formula ensures that $T$ is a subtype of `C`:

$$\texttt{sub}(\texttt{C},\texttt{Object}) \triangleq \texttt{sub}(\bot,\texttt{C}) \triangleq \textbf{true}$$

$$\texttt{sub}(\texttt{C},\texttt{C}') \triangleq \bigvee_{\texttt{C}_0 \xrightarrow{\texttt{d}_1} \texttt{C}_1 \xrightarrow{\texttt{d}_2} \dots \xrightarrow{\texttt{d}_n} \texttt{C}_n} \bigwedge_{1 \leq i < n} (\texttt{d}_i \wedge \texttt{O}(\texttt{d}_i,\texttt{C}_{i-1},\texttt{C}_i))$$

$$\text{with } \texttt{C}_0 \xrightarrow{\texttt{d}_1} \texttt{C}_1 \xrightarrow{\texttt{d}_2} \dots \xrightarrow{\texttt{d}_n} \texttt{C}_n \in \texttt{sub}(\texttt{C},\texttt{C}')$$

Basically, in addition to the two classic special cases for the `Object` class and the **null** type, this formula looks at all the possible inheritance path between `C` and $\texttt{C}'$ and enforces, using the previous formula, that one of them must be present in the generated variant.

**$\texttt{decl}(\texttt{C})$ and $\texttt{decl}(\texttt{C.a})$.** These formula ensure that the class `C` (resp. the attribute `C.a`) is present in the generated variant:

$$\texttt{decl}(\texttt{Object}) \triangleq \texttt{decl}(\bot) \triangleq \textbf{true}$$

$$\texttt{decl}(\texttt{C}) \triangleq \bigvee_{\texttt{d}} (\texttt{d} \wedge \neg\texttt{rm}(\texttt{d},\texttt{C})) \qquad \text{with } \texttt{d} \in \texttt{add}(\texttt{C})$$

$$\texttt{decl}(\texttt{C.a}) \triangleq \bigvee_{(\texttt{d},\texttt{C}')} (\texttt{d} \wedge \neg\texttt{rm}(\texttt{d},\texttt{C}'.\texttt{a}) \wedge \texttt{sub}(\texttt{C},\texttt{C}')) \quad \text{with } \texttt{d} \in \texttt{add}(\texttt{C}'.\texttt{a})$$

Ensuring that a class is present is quite direct, as we only need an activated delta module `d` that adds it, with no delta module `d′` afterward that removes it. It is a little bit more complex for attributes, as they can be accessed through inheritance. Hence in this case, we look for a delta module `d` that declares the attribute `a` in an annex class `C′`, with not other delta module `d′` that removes that attribute afterward, and with an active inheritance path between `C` and `C′`.

### A.2.4 Proof of Theorem 1.

*Preliminary Lemma*

**Lemma 1.** *For all products of the product line $L$, there exists one and only one solution of the formula $\Pi$.*`FMandCK` *that entirely describe the construction of this product. More precisely, for a product, its corresponding solution $\sigma$ is such that: i) all the variables corresponding to features selected for that product are set to* **true** *in $\sigma$; ii) all the variables corresponding to modules activated for the construction of this product's variant are set to* **true** *in $\sigma$; and iii) all the other variables in* $\mathrm{dom}(\sigma)$ *are set to* **false** *in $\sigma$.*

*Proof.* This is a direct consequence of how the formula is constructed.

*Applicability Consistency*

**Lemma 2.** *The product line $L$ is applicable-consistent if and only if all its variants can be generated without error.*

*Proof.* In case $L$ has no product, this lemma is trivial. Let now consider that the product line has at least one product: we prove the equivalence by proving each implication independently.

$\Rightarrow$. Suppose chosen a specific product of $L$: by Lemma 1, there exists one solution $\sigma$ of $\Pi$.`FMandCK` that entirely describes the construction of the product. Because `ac`($L$) is a tautology, $\sigma$ is thus also a solution of `ac`($L$). Let prove this implication by contradiction: suppose that the variant cannot be generated because an operation *op* of a module `d` cannot be applied. We have six cases: adding, modifying and removing classes or attributes. Let consider the case of adding a class `C` (the other cases are similar). By definition of the operation **adds**, the variant cannot be generated because `d` is applied on a program that already contains a class `C`, that was added by another module `d′` and wasn't removed afterward. However, this is in contradiction with the formula `appADD(C)`. Hence the hypothesis is wrong, and the variant can be generated.

$\Leftarrow$. Let prove this implication by contradiction: suppose that all the products can be generated but `ac`($L$) is not a tautology. As `ac`($L$), there exist a solution $\sigma$ of $\Pi$.`FMandCK` which puts to false the right part of the implication in `ac`($L$). By Lemma 1, $\sigma$ corresponds to a product of $L$ which, by hypothesis, can be generated. The right part of the implication in `ac`($L$) can be false for three possible reason: `appADD`($\rho$) or `appRM`($\rho$) or `appMOD`($\rho$) could be false for some reference $\rho$. Let consider the case where $\texttt{appADD}(\texttt{C})$ for some class name $\texttt{C}$ (the other cases are similar). This constraint is false if and only if there the variant of the considered product is generated with two modules $\texttt{d}$ and $\texttt{d}'$ that add the class $\texttt{C}$ without a third module inbetween to remove it: this is impossible as the variant can be generated. Hence the hypothesis is wrong, and $\texttt{ac}(L)$ is a tautology.

*Dependency Consistency*

**Lemma 3.** *Suppose that $L$ is dependency-consistent. Consider a generated variant of $L$. Then for all expressions $e$ in a method $\texttt{C.m}$ of the variant such that $\texttt{FCST}, \Gamma \vdash e : T$ is a valid partial typing derivation in $L$, there exists a valid derivation of $\texttt{CST}, \Gamma \vdash e \vdash T$, where $\texttt{CST}$ is the CST of the variant.*

*Proof.* By Lemma 1, the considered variant corresponds to a solution $\sigma$ of $\Pi.\texttt{FMandCK}$. Hence, by construction and because $L$ is dependency-consistent, $\sigma$ is also a solution of $\texttt{dc}(L)$. Moreover, by construction of the variant, there exists a module $\texttt{d}$ that defines $e$ in $\texttt{C.m}$ with

$$\begin{array}{ll} \sigma(\texttt{d}) = \textbf{true} & \\ \sigma(\texttt{d}') = \textbf{false} & \forall \texttt{d}' \in (\texttt{remove(C)} \cup \texttt{remove(C.m)} \cup \texttt{replace(C.m)}) \cap \texttt{after(d)} \end{array} \tag{1}$$

We prove this lemma by induction on the stucture of $e$. For simplicity, when possible we will use the same notation as in the partial typing rules presented in Figure 6, and we will only present the most relevant cases in this demonstration:

- **Variable $e = \texttt{x}$.** This case is straightforward using the typing rule (T-VAR);
- **Field Access $e = e.\texttt{f}$.** By hypothesis, we have that $\texttt{CST}, \Gamma \vdash e : \texttt{C}$. Moreover, by construction of $\texttt{dc}(L)$, the following formula is validated by $\sigma$:

  $$\texttt{d} \Rightarrow \neg\texttt{rm(d,C)} \Rightarrow \neg(\texttt{rm(d,C.m)} \vee \texttt{replace(d,C.m)}) \Rightarrow \texttt{decl(C.f)}$$

  By equation 1, we thus have that $\sigma$ validates $\texttt{decl(C.f)}$. By construction of the constraint $\texttt{decl(C.f)}$, we have that $\texttt{C.f}$ is declared in $\texttt{CST}$. By construction of $\texttt{CST}$, we also have that $\texttt{CST(C.f)} = \texttt{FCST(C.f)} = \texttt{C}'$. We can thus apply the typing rule (T-ACCESS) to have the result;
- **Method Call $e.\texttt{m}'(e_1, \ldots, e_n)$.** By hypothesis, we have that

  $$\begin{array}{ll} \texttt{FCST}, \Gamma \vdash e : \texttt{C} & \texttt{FCST}, \Gamma \vdash e_i : T_i \\ \texttt{FCST(C.m}') = \texttt{C}_0(\texttt{C}'_1, \ldots, \texttt{C}'_n) & T_i <: \texttt{C}'_i \end{array}$$

  Moreover, by construction of $\texttt{dc}(L)$, the following formula is validated by $\sigma$:

  $$\texttt{d} \Rightarrow \neg\texttt{rm(d,C)} \Rightarrow \neg(\texttt{rm(d,C.m)} \vee \texttt{replace(d,C.m)}) \Rightarrow \bigwedge_i (\texttt{sub}(T_i, \texttt{C}_i)) \wedge \texttt{decl(C.m}')$$

  By equation 1, we thus have that $\sigma$ validates $\bigwedge_i (\texttt{sub}(T_i, \texttt{C}_i)) \wedge \texttt{decl(C.m}')$. Using the same reasoning as in the previous case, using how the constraint $\texttt{decl}$ and $\texttt{sub}$ are constructed, we can remark that: i) $\texttt{C.m}'$ is declared in $\texttt{CST}$, ii) $\texttt{CST(C.m}') = \texttt{FCST(C.m}') = \texttt{C}_0(\texttt{C}'_1, \ldots, \texttt{C}'_n)$, and iii) $T_i <: \texttt{C}'_i$ is declared in $\texttt{CST}$ for all $i$. We can hence use the typing rule (T-CALL) to get the result in this case.

**Lemma 4.** *Consider a variant of $L$ that can be generated and that contains an expression in a method* $\texttt{C.m}$ *coming from a module* $\texttt{d}$ *such that* $\texttt{CST}, \Gamma \vdash e : T$ *holds. Consider also $\sigma$ to be the solution of* $\texttt{ac}(L)$ *corresponding to the variant by Lemma 1. Then there exists a partial typing derivation of* $\texttt{FCST}, \Gamma \vdash e : T$*, and a dependency generation derivation of* $\Gamma \vdash e : T \mid \Phi$ *with: i)* $\sigma(\texttt{d}) = \textbf{true}$*, ii)* $\sigma$ *validating* $\neg\texttt{rm}(\texttt{d}, \texttt{C})$ *and* $\neg(\texttt{rm}(\texttt{d}, \texttt{C.m}) \vee \texttt{replace}(\texttt{d}, \texttt{C.m}))$*, and iii)* $\sigma$ *validating* $\Phi$*.*

*Proof.* Because $\sigma$ corresponds to the variant, we have that:

$$\begin{array}{ll} \sigma(\texttt{d}) = \textbf{true} & \\ \sigma(\texttt{d}') = \textbf{false} & \forall \texttt{d}' \in (\texttt{remove}(\texttt{C}) \cup \texttt{remove}(\texttt{C.m}) \cup \texttt{replace}(\texttt{C.m})) \cap \texttt{after}(\texttt{d}) \end{array} \tag{2}$$

Consequently, we have that $\sigma$ validates $\neg(\texttt{rm}(\texttt{d}, \texttt{C.m}) \vee \texttt{replace}(\texttt{d}, \texttt{C.m}))$ and $\neg\texttt{rm}(\texttt{d}, \texttt{C})$. We finally prove that the partial typing statement $\texttt{FCST}, \Gamma \vdash e : T$, the dependency constraint generation statement $\Gamma \vdash e : T \mid \Phi$ and $\sigma$ validating $\Phi$ hold by induction on $e$. For simplicity, when possible we will use the same notation as in the typing rules presented in Figure 5, and we will only present the most relevant cases in this demonstration:

- **Variable $e = \texttt{x}$.** This case is straightforward using the partial typing rule (PT-Var) and the first rule of Figure 3;
- **Field Access $e = e.\texttt{f}$.** By hypothesis, we have that $\texttt{FCST}, \Gamma \vdash e : \texttt{C}$and $\Gamma \vdash e : \texttt{C} \mid \Phi'$ hold, with $\Phi = \Phi' \wedge \texttt{decl}(\texttt{C.f})$ and $\sigma$ validating $\Phi'$. By construction of $\texttt{CST}$, we have that $\texttt{FCST}(\texttt{C.f}) = \texttt{CST}(\texttt{C.f}) = \texttt{C}'$. Hence, using the partial typing rule (PT-Access), we have that $\texttt{FCST}, \Gamma \vdash e.\texttt{f} : \texttt{C}'$ holds. Moreover, using the second rule of Figure 3, we have that $\Gamma \vdash e.\texttt{f} : \texttt{C}' \mid \Phi$ holds too. Finally, by construction of the constraint of $\texttt{decl}(\texttt{C.f})$, the fact that $\sigma$ corresponds to the variant and that $\texttt{CST}$ contains $\texttt{C.m}$, we have that $\sigma$ validates $\texttt{decl}(\texttt{C.f})$;
- **Method Call $e.\texttt{m}'(e_1, \ldots, e_n)$.** By hypothesis, we have that

  $$\begin{array}{lll} \texttt{FCST}, \Gamma \vdash e : \texttt{C} & \Gamma \vdash e : \texttt{C} \mid \Phi' & \sigma \text{ validates } \Phi' \\ \texttt{FCST}, \Gamma \vdash e_i : T_i & \Gamma \vdash e_i : T_i \mid \Phi_i & \sigma \text{ validates } \Phi_i \end{array}$$

  By construction of $\texttt{CST}$, we have that $\texttt{FCST}(\texttt{C.m}) = \texttt{CST}(\texttt{C.m}) = \texttt{C}_0(\texttt{C}'_1, \ldots, \texttt{C}'_n)$ and that $T_i <: \texttt{C}_i$ in $\texttt{FCST}$. Hence, using the partial typing rule (PT-Call), we have that $\texttt{FCST}, \Gamma \vdash e.\texttt{m}'(e_1, \ldots, e_n) : \texttt{C}_0$ holds. Moreover, using the fourth rule of Figure 3, we have that the following statement holds:

  $$\Gamma \vdash e.\texttt{m}(e_1, \ldots, e_n) : \texttt{C}_0 \mid \bigwedge_i (\Phi_i \wedge \texttt{sub}(T_i, \texttt{C}_i)) \wedge \Phi' \wedge \texttt{decl}(\texttt{C.m})$$

  As $\texttt{CST}$ contains $\texttt{C.m}$ and all subtyping statements $T_i <: \texttt{C}_i$, we have that $\sigma$ validates $\texttt{decl}(\texttt{C.m})$ and $\texttt{sub}(T_i, \texttt{C}_i)$ for all $i$. We can then conclude that $\sigma$ validates $\Phi$.

**Lemma 5.** *Suppose that all the variants of $L$ are generable. Consider the properties:*

*i. L is well partially-typed, and dependency-consistent.*
*ii. All the variants of L are well-typed IFJ programs.*

*Then: (i) implies (ii); and if L has no useless declarations and operations then (ii) implies (i).*

*Proof.* $(i) \Rightarrow (ii)$: let consider a generated variant of $L$ and its corresponding solution $\sigma$ (by Lemma 1). We prove it is typable by induction on its structure. For simplicity, when possible we will use the same notation as in the partial typing rules presented in Figure 6, and we will only present the most relevant cases in this demonstration:

– **Method** $\mathtt{C}_0\ \mathtt{m}(\overline{\mathtt{C}\ \mathtt{x}})$ $\{$**return** $e;\}$ **in class** $\mathtt{C}$**.** By hypothesis, there exists a module $\mathtt{d}$ used in the construction of the variant that declares this method, and a partial typing derivation of $\mathtt{FCST}, \mathtt{this} : \mathtt{C}; \mathtt{x_i} : \mathtt{C_i} \vdash e : \mathtt{C}'$ with $\mathtt{C}' <: \mathtt{C}_0$ in $\mathtt{FCST}$. As $L$ is supposed dependency-consistent, we can apply Lemma 3 to get a valid type derivation of $\mathtt{CST}, \mathtt{this} : \mathtt{C}; \mathtt{x_i} : \mathtt{C_i} \vdash e : \mathtt{C}'$ with $\mathtt{CST}$ being the CST of the variant. Moreover, as $\mathtt{ac}(L)$ is a tautology, we have, using the same reasoning as in the previous proofs, that $\sigma$ validates

$$\mathtt{d} \Rightarrow \neg\mathtt{rm}(\mathtt{d},\mathtt{C}) \Rightarrow \neg(\mathtt{rm}(\mathtt{d},\mathtt{C.m}) \vee \mathtt{replace}(\mathtt{d},\mathtt{C.m})) \Rightarrow \bigwedge_i \mathtt{decl}(\mathtt{C}_i) \wedge \mathtt{sub}(\mathtt{C}',\mathtt{C}_0)$$

Hence, by construction of the constraints $\mathtt{decl}(\mathtt{C}_i)$, we have that $\mathtt{CST}$ contains $\mathtt{C}_i$, and by construction of the constraint $\mathtt{sub}(\mathtt{C}',\mathtt{C}_0)$, we have that $\mathtt{C}' <: \mathtt{C}_0$ in $\mathtt{CST}$. We can thus conclude this case by applying the typing rule (T-Method).

$(ii) \Leftarrow (i)$. We prove the result by induction on the structure of $L$. For simplicity, when possible we will use the same notation as in the typing rules presented in Figure 5, and we will only present the most relevant cases in this demonstration:

– **Method** $\mathtt{C}_0\ \mathtt{m}(\overline{\mathtt{C}\ \mathtt{x}})$ $\{$**return** $e;\}$ **in class** $\mathtt{C}$ **in module** $\mathtt{d}$**.** As $L$ does not contains useless code, there exists a variant and a corresponding solution $\sigma$ (by Lemma 1) that contains that exact method. Note then that By Hypothesis, and using the definition of the typing rule (T-Method), the following statements hold, with $\mathtt{CST}$ beeing the CST of the variant:

$$\begin{array}{ll} \mathtt{CST}, \mathtt{C} \vdash \mathtt{C}_0\ \mathtt{m}(\mathtt{C}_1\ \mathtt{x}_1, \ldots, \mathtt{C}_n\ \mathtt{x}_n)\ \{\textbf{return}\ e\} & \\ \mathtt{CST}, \mathtt{this} : \mathtt{C}; \mathtt{x_i} : \mathtt{C_i} \vdash e : \mathtt{C}' & \mathtt{C}' <: \mathtt{C}_0 \text{ is in } \mathtt{CST} \\ \mathtt{CST}(\mathtt{C.m}) = \mathtt{C}_0(\mathtt{C}_1, \ldots, \mathtt{C}_n) & \mathtt{C}_i \in \mathrm{dom}(\mathtt{CST}) \end{array}$$

By Lemma 4, there exist valid derivations of

$$\begin{array}{ll} & \mathtt{FCST}, \mathtt{this} : \mathtt{C}; \mathtt{x_i} : \mathtt{C_i} \vdash e : \mathtt{C} \qquad \mathtt{this} : \mathtt{C}; \mathtt{x_i} : \mathtt{C_i} \vdash e : \mathtt{C} \mid \Phi \\ \text{with} & \sigma(\mathtt{d}) = \textbf{true} \\ & \sigma \text{ validating } \neg\mathtt{rm}(\mathtt{d},\mathtt{C}) \text{ and } \neg(\mathtt{rm}(\mathtt{d},\mathtt{C.m}) \vee \mathtt{replace}(\mathtt{d},\mathtt{C.m})) \text{ and } \Phi \end{array}$$

By construction of $\mathtt{CST}$, we have that $\mathtt{C}_i \in \mathrm{dom}(\mathtt{FCST})$ and $\mathtt{C}' <: \mathtt{C}_0$ in $\mathtt{FCST}$. Hence, we can apply the partial typing rule (PT-Method) to get a valid derivation of $\mathtt{FCST}, \mathtt{C} \vdash \mathtt{C}_0\ \mathtt{m}(\mathtt{C}_1\ \mathtt{x}_1, \ldots, \mathtt{C}_n\ \mathtt{x}_n)\ \{\textbf{return}\ e\}$. Moreover, using the second rule of Figure 4, we get a dependency generation derivation of

$$\begin{aligned}&\mathtt{d}, \mathtt{C} \vdash \mathtt{C}_0\ \mathtt{m}(\mathtt{C}_1\ \mathtt{x}_1, \ldots, \mathtt{C}_n\ \mathtt{x}_n)\ \{\textbf{return}\ e\}\\ &\qquad : \neg(\mathtt{rm}(\mathtt{d}, \mathtt{C.m}) \vee \mathtt{replace}(\mathtt{d}, \mathtt{C.m})) \Rightarrow (\textstyle\bigwedge_i \mathtt{decl}(\mathtt{C}_i) \wedge \Phi \wedge \mathtt{sub}(\mathtt{C}', \mathtt{C}_0))\end{aligned}$$

As $\mathtt{C}_i \in \mathrm{dom}(\mathtt{CST})$ and $\mathtt{C}' <: \mathtt{C}_0$ is in $\mathtt{CST}$, we can conclude that $\sigma$ validates $\mathtt{decl}(\mathtt{C}_i)$ and $\mathtt{sub}(\mathtt{C}', \mathtt{C}_0)$.

We now need to prove that $\mathtt{dc}(L)$ is a tautology. Let consider $\sigma'$ that validates $\Pi.\mathtt{FMandCK}$. By Lemma 1, there exists a variant that corresponds to $\sigma'$. now let consider $\mathtt{dep}(L)$: this formula is constructed as a conjunction of formula of the form

$$\begin{array}{ll}\mathtt{d} \Rightarrow \neg\mathtt{rm}(\mathtt{d}, \mathtt{C}) \Rightarrow \neg\mathtt{rm}(\mathtt{d}, \mathtt{C.f}) \Rightarrow \mathtt{decl}(\mathtt{C}') & \text{for fields}\\ \mathtt{d} \Rightarrow \neg\mathtt{rm}(\mathtt{d}, \mathtt{C}) \Rightarrow \neg(\mathtt{rm}(\mathtt{d}, \mathtt{C.m}) \vee \mathtt{replace}(\mathtt{d}, \mathtt{C.m})) \Rightarrow \Phi & \text{for methods}\\ \mathtt{d} \Rightarrow \neg\mathtt{rm}(\mathtt{d}, \mathtt{C}) \Rightarrow \neg\mathtt{modifyEC}(\mathtt{d}, \mathtt{C}) \Rightarrow \mathtt{decl}(\mathtt{C}') \wedge \neg\mathtt{sub}(\mathtt{C}', \mathtt{C}) & \text{for } \textbf{extends} \text{ clauses}\end{array}$$

Let consider a formula for methods where $\sigma'$ validates $\mathtt{d}$, $\neg\mathtt{rm}(\mathtt{d}, \mathtt{C})$ and $\neg(\mathtt{rm}(\mathtt{d}, \mathtt{C.m}) \vee \mathtt{replace}(\mathtt{d}, \mathtt{C.m}))$. By construction, we thus have that this method is part of the variant, and so, using the previous case analysis, we have that $\sigma'$ validates $\Phi$. This result can straightforwardly be extended to fields and **extends** clauses. Hence, $\sigma'$ validates $\mathtt{dep}(L)$. We can then conclude that $\mathtt{dc}(L)$ is a tautology.

*Main Theorem*

*Proof (Theorem 1).* This theorem is a direct consequence of Lemmas 2 and 5.

### A.3 Version of the Approach in Section 5.

We present in this section the full set of typing rules for type-non-uniform IF$\Delta$J Product Lines. We start our presentation with the rules for Partial Typing, and then present the more complex rules of Dependency.

**A.3.1 Product Line Partial Typing.** Many rules describing Partial Typing in the type-non-uniform case are identical to the type-uniform case. Hence, to keep our presentation simple, we only present in Figure 7 the rules that are different from the ones presented in Figure 6: only the rules for expressions and for attribute declaration are changed. These rules are, like introduced in the main document, a direct extension of the rules for the type-uniform case, with two modifications: first, the lookup in the FCST returns a set of types; second, expressions are typed with a *set* of possible types. The most complex rules are (PT-Access) and (PT-Call), as they have to directly manipulate the combinatorial complexity of matching sets of types. The other rules are very close, if not identical, to their type-uniform version.

$$\textsc{PT-Var}\quad \dfrac{\Gamma(\mathtt{x}) = \mathtt{C}}{\mathtt{FCST}, \Gamma \vdash \mathtt{x} : \{\mathtt{C}\}} \qquad\qquad \textsc{PT-Null}\quad \mathtt{FCST}, \Gamma \vdash \textbf{null} : \{\bot\}$$

$$\textsc{PT-Access}\quad \dfrac{\mathtt{FCST}, \Gamma \vdash e : \{T_j\}_{j\in J} \qquad J' = \{j \mid j \in J \wedge T_j.\mathtt{f} \in \mathrm{dom}(\mathtt{FCST})\} \qquad \{\mathtt{C}_k\}_{k\in K} = \bigcup_{j\in J'} \mathtt{FCST}(T_j.\mathtt{f})}{\mathtt{FCST}, \Gamma \vdash e.\mathtt{f} : \{\mathtt{C}_k\}_{k\in K}}$$

$$\textsc{PT-Call}\quad \dfrac{\begin{array}{c}\mathtt{FCST}, \Gamma \vdash e : \{\mathtt{C}_j\}_{j\in J} \qquad \mathtt{FCST}, \Gamma \vdash e_i : \{T_{i,j}\}_{j\in J_i} \\ S = \{\mathtt{C}_0(\mathtt{C}_1, \ldots, \mathtt{C}_m) \mid \exists j \in J \wedge \mathtt{C}_0(\mathtt{C}_1, \ldots, \mathtt{C}_m) \in \mathtt{FCST}(\mathtt{C}_j.\mathtt{m}) \wedge m = n\} \\ S' = \{\mathtt{C}_0 \mid \mathtt{C}_0(\mathtt{C}_1, \ldots, \mathtt{C}_n) \in S \wedge \forall 1 \leq i \leq n \exists j \in J_i,\ T_{i,j} <: \mathtt{C}_i\}\end{array}}{\mathtt{FCST}, \Gamma \vdash e.\mathtt{m}(e_1, \ldots, e_n) : S'}$$

$$\textsc{PT-Assign}\quad \dfrac{\begin{array}{c}\mathtt{FCST}, \Gamma \vdash e.\mathtt{f} : \{\mathtt{C}_j\}_{j\in J} \\ \mathtt{FCST}, \Gamma \vdash e' : \{T_j\}_{j\in J'} \qquad J'' = \{j \mid j \in J \wedge \exists j' \in J',\ T_{j'} <: \mathtt{C}_j\}\end{array}}{\mathtt{FCST}, \Gamma \vdash e.\mathtt{f} = e' : \{\mathtt{C}_j\}_{j\in J''}} \qquad \textsc{PT-New}\quad \dfrac{\mathtt{C} \in \mathrm{dom}(\mathtt{FCST})}{\mathtt{FCST}, \Gamma \vdash \textbf{new}\ \mathtt{C}() : \{\mathtt{C}\}}$$

$$\textsc{PT-Cast}\quad \dfrac{\mathtt{FCST}, \Gamma \vdash e : \{T_j\}_{j\in J} \qquad \{j \mid j \in J \wedge (T_j <: \mathtt{C} \vee \mathtt{C} <: T_j)\} \neq \emptyset}{\mathtt{FCST}, \Gamma \vdash (\mathtt{C})e : \{\mathtt{C}\}}$$

$$\textsc{PT-Field}\quad \dfrac{\mathtt{C}' \in \mathrm{dom}(\mathtt{FCST}) \qquad \mathtt{C}' \in \mathtt{FCST}(\mathtt{C.f})}{\mathtt{FCST}, \mathtt{C} \vdash \mathtt{C}'\ \mathtt{f}} \qquad \textsc{PT-Method}\quad \dfrac{\begin{array}{c}\mathtt{FCST}, \mathtt{this} : \mathtt{C}; \mathtt{x_i} : \mathtt{C_i} \vdash e : \{\mathtt{C_j}\}_{\mathtt{j}\in\mathtt{J}} \qquad \exists j \in J, \mathtt{C}_j <: \mathtt{C}_0 \\ \mathtt{C}_0(\mathtt{C}_1, \ldots, \mathtt{C}_n) \in \mathtt{FCST}(\mathtt{C})(\mathtt{m}) \qquad \mathtt{C}_i \in \mathrm{dom}(\mathtt{FCST})\end{array}}{\mathtt{FCST}, \mathtt{C} \vdash \mathtt{C}_0\ \mathtt{m}(\mathtt{C}_1\ \mathtt{x}_1, \ldots, \mathtt{C}_n\ \mathtt{x}_n)\ \{\textbf{return}\ e\}}$$

**Fig. 7.** IF$\Delta$J Partial Typing Rules for the version of the approach in Section 5

**A.3.2 Product Line Dependency.** Similarily to the Partial Typing, most of the rules describing dependency in the type-non-uniform case are identical to the type-uniform case, except for expressions an few other declarations. Hence, we present in Figure 8 only the rules that are different from the ones presented in Figures 3 and 4. As introduced in Section 5, the statements used for our dependency analysis have the form $\Gamma \vdash e : [\Phi_i \mapsto T_i]_{i\in I}$ where the $T_i$ are the possible types of $e$, and $\Phi_i$ is the condition (i.e. which module must or must not be activated) for $e$ to have the type $T_i$ in the final product. Moreover, to simplify the presentation of our rules, we use the two following notations. First, we extend the FCST to specify in which module and class a specific type is associated to an attribute: e.g. $\mathtt{FCST}(\mathtt{C.f}) = [\mathtt{d}_1, \mathtt{C} \mapsto \mathtt{C}_1; \mathtt{d}_2, \mathtt{C}' \mapsto \mathtt{C}_2]$ means that the module $\mathtt{d}_1$ declares the field $\mathtt{f}$ in class $\mathtt{C}$ to be of type $\mathtt{C}_1$, while the module $\mathtt{d}_2$ declare it in a superclass $\mathtt{C}'$ and to have the type $\mathtt{C}_2$. Second, we define the predicate $\mathtt{decl}(\mathtt{d}, \mathtt{C}, \mathtt{C}'.\mathtt{a})$ to ensure that the attribute $\mathtt{a}$ defined in class $\mathtt{C}$ by $\mathtt{d}$ is accessible from class $\mathtt{C}'$ in the final variant:

$$\mathtt{decl}(\mathtt{d}, \mathtt{C}', \mathtt{C.a}) \triangleq \mathtt{d} \wedge \neg\mathtt{rm}(\mathtt{d}, \mathtt{C}'.\mathtt{a}) \wedge \mathtt{sub}(\mathtt{C}, \mathtt{C}'))$$

Like for the Partial Typing, the most complex rules in Figure 8 are for the field lookup[4] and method call, due to the complexity to match sets of type using

[4] Note that the rule presented in Section 5 is slightly different from rule (D-Access) which uses a more explicit notation fitting with the other rule of Figure 8.

$$
\textsc{D-Var}\ \frac{\mathit{\Gamma}(\mathtt{x}) = \mathtt{C}}{\mathit{\Gamma} \vdash \mathtt{x} : [\mathbf{true} \mapsto \mathtt{C}]}
\qquad
\textsc{D-Null}\ \mathit{\Gamma} \vdash \textsf{null} : [\mathbf{true} \mapsto \bot]
\qquad
\textsc{D-New}\ \mathit{\Gamma} \vdash \textsf{new}\ \mathtt{C}() : [\mathtt{decl}(\mathtt{C}) \mapsto \mathtt{C}]
$$

$$
\textsc{D-Access}\ \frac{\begin{array}{c}\mathit{\Gamma} \vdash e : [\mathit{\Phi}_j \mapsto T_j]_{j \in J} \\ J' = \{j \mid j \in J \wedge T_j.\mathtt{f} \in \mathrm{dom}(\mathtt{FCST})\} \qquad [\mathtt{d}_{j,k}, \mathtt{C}_{j,k} \mapsto \mathtt{C}_{j,k}]_{k \in K_j} = \mathtt{FCST}(T_j.\mathtt{f})\end{array}}{\mathit{\Gamma} \vdash e.\mathtt{f} : [(\mathit{\Phi}_j \wedge \mathtt{decl}(\mathtt{d}_{j,k}, \mathtt{C}_{j,k}, T_j.\mathtt{f})) \mapsto \mathtt{C}_{j,k}]_{j \in J', k \in K_j}}
$$

$$
\textsc{D-Call}\ \frac{\begin{array}{c}\mathtt{FCST}, \mathit{\Gamma} \vdash e : [\mathit{\Phi}_j \mapsto \mathtt{C}_j]_{j \in J} \qquad \mathit{\Gamma} \vdash e_i : [\mathit{\Phi}_j \mapsto T_j]_{j \in J_i} \qquad J' = \{j \mid j \in J \wedge \mathtt{m} \in \mathrm{dom}(\mathtt{FCST}(\mathtt{C}_j))\} \\ [\mathtt{d}_i, \mathtt{C}_i \mapsto \mathtt{C}_{0,i}(\mathtt{C}_{1,i}, \ldots \mathtt{C}_{m_i,i})]_{i \in I_j} = \mathtt{FCST}(\mathtt{C}_j.\mathtt{m}) \qquad I'_j = \{i \mid i \in I_j \wedge m_i = n\}\end{array}}{\mathit{\Gamma} \vdash e.\mathtt{m}(e_1, \ldots, e_n) : [(\mathtt{decl}(\mathtt{d}_i, \mathtt{C}_i, \mathtt{C}_j.\mathtt{m}) \wedge \bigwedge_{1 \leq k \leq n} \bigvee_{\substack{\{j \mid j \in J_k \wedge \\ T_j <: \mathtt{C}_{1,i}\}}} \mathtt{sub}(T_j, \mathtt{C}_{k,i})) \mapsto \mathtt{C}_{0,i}]_{j \in J', i \in I'_j}}
$$

$$
\textsc{D-Assign}\ \frac{\begin{array}{c}\mathit{\Gamma} \vdash e.\mathtt{f} : [\mathit{\Phi}_j \mapsto \mathtt{C}_j]_{j \in J} \qquad \mathit{\Gamma} \vdash e' : [\mathit{\Phi}_j \mapsto T_j]_{j \in J'} \\ S = \{(j, j') \mid j \in J \wedge j' \in J' \wedge T_{j'} <: \mathtt{C}_j\} \qquad K = \{j \mid (j, j') \in S\} \qquad S_j = \{j' \mid (j, j') \in S\}\end{array}}{\mathit{\Gamma} \vdash e.\mathtt{f} = e' : [\bigvee_{j' \in S_j} (\mathit{\Phi}_{j'} \wedge \mathtt{sub}(T_{j'}, \mathtt{C}_j)) \mapsto \mathtt{C}_j]_{j \in K}}
$$

$$
\textsc{D-Cast}\ \frac{\mathit{\Gamma} \vdash e : [\mathit{\Phi}_j \mapsto T_j]_{j \in J} \qquad J_1 = \{j \mid j \in J \wedge T_j <: \mathtt{C}\} \qquad J_2 = \{j \mid j \in J \wedge \mathtt{C} <: T_j\}}{\mathit{\Gamma} \vdash (\mathtt{C})e : [(\bigvee_{j \in J_1} \mathit{\Phi}_j \wedge \mathtt{sub}(T_j, \mathtt{C})) \vee (\bigvee_{j \in J_2} \mathit{\Phi}_j \wedge \mathtt{sub}(\mathtt{C}, T_j)) \mapsto \mathtt{C}]}
$$

$$
\textsc{D-Method}\ \frac{\mathtt{this} : \mathtt{C}; \mathtt{x_i} : \mathtt{C_i} \vdash e : [\mathit{\Phi}_\mathtt{j} \mapsto \mathtt{T_j}]_{\mathtt{j \in J}} \qquad J' = \{j \mid j \in J \wedge T_j <: \mathtt{C}_0\}}{\begin{array}{c}\mathtt{d}, \mathtt{C} \vdash \mathtt{C}_0\ \mathtt{m}(\mathtt{C}_1\ \mathtt{x}_1, \ldots, \mathtt{C}_n\ \mathtt{x}_n)\ \{\textsf{return}\ e\} \\ : \neg(\mathtt{rm}(\mathtt{d}, \mathtt{C.m}) \vee \mathtt{replace}(\mathtt{d}, \mathtt{C.m})) \Rightarrow (\bigwedge_i \mathtt{decl}(\mathtt{C}_i) \wedge (\bigvee_{j \in J'} \mathit{\Phi}_j \wedge \mathtt{sub}(T_j, \mathtt{C}_0))\end{array}}
$$

$$
\frac{\begin{array}{c}MD = \mathtt{C}_0\ \mathtt{m}(\mathtt{C}_1\ \mathtt{x}_1, \ldots, \mathtt{C}_n\ \mathtt{x}_n)\ \{\textsf{return}\ e\} \qquad \mathtt{FCST}(\mathtt{C.m}) = [\mathtt{d}_j, \mathtt{C}_j \mapsto \mathtt{C}_{0,j}(\mathtt{C}_{1,j}, \ldots, \mathtt{C}_{n_j,j})]_{j \in J} \\ S = \{(\mathtt{d}_j, \mathtt{C}_j) \mid j \in J \wedge n_j = n \wedge \forall 1 \leq i \leq n,\ \mathtt{C}_{i,j} = \mathtt{C}_i\} \\ \mathtt{FCST} ::= \mathtt{FCST}[\mathtt{C.original} \mapsto [\mathtt{d}', \mathtt{C}' \mapsto \mathtt{C_0}(\mathtt{C_1}, \ldots, \mathtt{C_n})]_{(\mathtt{d}', \mathtt{C}') \in \mathtt{S}}] \qquad \mathtt{d}, \mathtt{C} \vdash MD : \mathit{\Phi}\end{array}}{\mathtt{d}, \mathtt{C} \vdash \textsf{modifies}\ MD : \mathit{\Phi}}
$$

**Fig. 8.** Dependency generation for the version of the approach in Section 5.

inheritance. Basically, rule (D-Access) computes the set of possible types of $e$ with their application conditions: $[\mathit{\Phi}_j \mapsto T_j]_{j \in J}$. Then it extracts from that set the types that have a field $\mathtt{f}$ with $J' = \{j \mid j \in J \wedge T_j.\mathtt{f} \in \mathrm{dom}(\mathtt{FCST})\}$. And it finally generates the guarded set of types for $e.\mathtt{f}$ by taking the type of $T_j.\mathtt{f}$ for all $j \in J'$, and guarding it with a formula that ensures that this field is declared with the right type in the final variant. Rule (D-Call) performs similar operations, with the additional complexity of matching $n$ pairs of types using inheritance.

In addition to the rules for expression, we also redefine the rule for method declaration (to manage the multiple guarded types of the inner expression), and the rule for method modification. In this last rule we need to add a local binding in the FCST to define the `original` method, which must deal with the new extended structure of the FCST.

**A.3.3 Soundness and Completeness.** The properties of these modified algorithm is stated simiarily as in Theorem 1:

**Theorem 2.** *Let L be a product line. Consider the properties:*

*i. L is well partially-typed, applicability- and dependency-consistent.*
*ii. All the variants of L can be generated and are well-typed IFJ programs.*

*Then: (i) implies (ii); and if L has no useless operations then (ii) implies (i).*

*Proof.* This result can be proved by following the same approach as in Section A.2.4.