Influence of programming style in transformation bad smells: Mining of ETL Repositories

Nicolás Bonet[1a], Kelly Garcés[1a], Rubby Casallas[1a], María Elsa Correal[1b], and Ran Wei[2]

[1] Universidad de los Andes, Bogota D.C. Colombia, School of Engineering. Department of Systems and Computing Engineering[a] and Departament of Industrial Engineering[b]
{n.bonet2476, kj.garces971, rcasalla, mcorreal}@uniandes.edu.co
[2] The University of York. UK. Department of Computer Science
{ran.wei}@cs.york.ac.uk

References

1. Wimmer, M., Jouault, F., Cabot, J.: A catalogue of refactorings for model-to-model transformations. Journal of Object Technology 11(2), 2–1 (2012)

A Annexed

Table 1. Catalog of bad smells proposed in literature and detected by our approach

Acronym[1]	**Name**	**Problem**	**Algorithm**
DOE	Duplicated and complex OCL expressions	Some OCL expressions are duplicated or too complex. The latter results in less readable code and the former in less maintainable code	All calls to operations inside an OCL are counted and if the result is higher than a threshold, it is shown as a complex OCL. In this case we set it to 5 based on our academic experience
TOC	Trivial operations called once	Operations that are called once all over the transformation (for example, to create an element), make the code less understandable	The entire transformation is verified checking how many times each operation is used, in the case that the operation is called just once and the operation has less than a defined threshold of lines (excluding comments and whitespaces, we defined the threshold in 3 due to our dataset review) then the bad smell is identified
REB	Rule body is embedded into if blocks	The body of the output pattern is embedded into if (and else) blocks to separate the logic being applied in different cases, which makes transformation hard to understand	Each rule is verified and if the first element contained in the output pattern is an if, then the bad smell count is increased
TMB	Target metamodel attribute set in multiple bindings	A target feature is being set by at least 2 bindings, this makes the transformation complex and harder to understand	Check the body of each rule looking if each target binding is set more than once, if that case is found this bad smell type is increased by one
NIC	Nested if/else chains should be replaced with switch	Nested IF/ELSE chains are commonly used to convert values from the source model into values needed for the target model, e.g., convert Java to SQL data types. This is discouraged as the switch statement is less complex and more suitable for this kind of scenarios	The discoverer checks each if statement and if it finds an else or a chain of elseif this bad smell is increased by one
CSF	A chain of select/first in OCL is less efficient than using selectFirst	An OCL expression containing Select/First operation chains is difficult to read and unnecessarily expensively calculated	All OCL expressions are checked searching for the combination of the operations: select and first, if there is at least one, then this bad smell is increased by one