



Automated refactoring of legacy JavaScript code to ES6 modules[☆]

Katerina Paltoglou, Vassilis E. Zafeiris^{*}, N.A. Diamantidis, E.A. Giakoumakis

Department of Informatics, Athens University of Economics and Business, 76 Patission Str., Athens 104 34, Greece

ARTICLE INFO

Article history:

Received 30 November 2020
Received in revised form 19 June 2021
Accepted 20 July 2021
Available online 28 July 2021

Keywords:

Refactoring
Code migration
Module object deconstructing
ES6 modules
AMD/CommonJS

ABSTRACT

The JavaScript language did not specify, until ECMAScript 6 (ES6), native features for streamlining encapsulation and modularity. Developer community filled the gap with a proliferation of design patterns and module formats, with impact on code reusability, portability and complexity of build configurations. This work studies the automated refactoring of legacy ES5 code to ES6 modules with fine-grained reuse of module contents through the *named import/export* language constructs. The focus is on reducing the coupling of refactored modules through deconstructing exported module objects to fine-grained module features and enhancing module dependencies by leveraging the ES6 syntax. We employ static analysis to construct a model of a JavaScript project, the Module Dependence Graph (MDG), that represents modules and their dependencies. On the basis of MDG we specify the refactoring procedure for module migration to ES6. A prototype implementation has been empirically evaluated on 19 open source projects. Results highlight the relevance of the refactoring with a developer intent for fine-grained reuse. The analysis of refactored code shows an increase in the number of reusable elements per project and reduction in the coupling of refactored modules. The soundness of the refactoring is empirically validated through code inspection and execution of projects' test suites.

© 2021 Elsevier Inc. All rights reserved.

1. Introduction

JavaScript followed, in the last decade, the wide reach of the web platform and turned into a general-purpose programming language for the mobile and server platform. Despite the growing scope and size of JavaScript applications, the language specification did not include, until recently, native features for streamlining encapsulation and modularity (Wirfs-Brock and Eich, 2020). Syntactic constructs for modules and classes, along with other features for writing complex applications (e.g., promises, arrow functions) were introduced in the ECMAScript 2015 (ES6) release (ECMA International, 2015). Since ES6 classes represent syntactic “sugar” over function constructors and prototype-based inheritance, they were implemented in browser engines soon after ES6 release. On the other hand, native support for ES6 modules in popular browsers was introduced after the first quarter of 2017 (MDN Web Docs, 2020), while it is still an experimental feature of the latest release of the Node.js runtime (OpenJS, 2020a).

ES6 modules support the development of complex applications through encapsulation of top-level variable and function

declarations in a JavaScript file in a new scope, the *module* scope, that differs from the global scope. Thus, top-level declarations do not “pollute” the global scope, preventing name collisions and enabling variable checking. Such incompatibilities in the execution model of ES6 code files with respect to ES5 or earlier versions, required non-trivial extensions to JavaScript engines that needed time to be tested and released.

As native support for modularity lagged with respect to the needs of the industry, the developer community filled the gap with the proliferation of design patterns, module formats and tools. The encapsulation of state and behavior and the avoidance of global scope “pollution” were handled by code patterns (e.g. Immediately Invoked Function Expressions – IIFEs) and module formats, mainly AMD (AMD JS, 2020) and CommonJS (OpenJS, 2020b), which introduced interfaces for declaring modules and their dependencies in the browser and the server respectively.

Since ES6 modules are currently supported in most JavaScript runtimes, developers are confronted with the need to migrate their legacy codebases to the standard module format. Besides simplification of code and build configurations, the migration provides performance improvements, as module loading and dependency management in ES6 are natively supported by the runtime, rather than external libraries. Moreover, special features of ES6 modules improve code maintainability: (a) named imports/exports enforce name consistency within the codebase, facilitate

[☆] Editor: Gabriele Bavota.

^{*} Corresponding author.

E-mail addresses: paltoglouk@aeub.gr (K. Paltoglou), bzafiris@aeub.gr (V.E. Zafeiris), nad@aeub.gr (N.A. Diamantidis), mgia@aeub.gr (E.A. Giakoumakis).

refactoring operations and allow for optimization of application size (tree-shaking) during deployment, (b) named exports act as live immutable bindings that enhance safety by preventing errors relevant to cyclic dependencies (Rauschmayer, 2015), (c) eager resolution of the reachable dependencies of a module (e.g. the application entypoint), provides fail-fast program behavior (Shore, 2004).

In this paper, we propose a fully-automated method for migration of an ES5 codebase to ES6 modules. The method applies to code that is either non-modular or uses the AMD/CommonJS module formats and operates on the Module Dependence Graph (MDG), a global model of the codebase. The MDG provides a representation of the code in terms of modules, exported module features and module dependencies, as inferred and optimized by the our method. The proposed migration represents a large and complex refactoring, since it affects the entire code-base and the transformation of each individual source file requires analysis of all incoming and outgoing dependencies. Besides handling the mapping among different module formats, our method focuses on reducing the coupling of refactored modules through: (a) selective destructuring of existing AMD/CommonJS module objects to fine-grained module features, (b) establishment of the minimum required dependencies among modules by leveraging the “named imports/exports” ES6 language construct and (c) encapsulation of global state and functions as cohesive module features.

As part of this work, we explore the relevance of the proposed refactoring and evaluate its impact on system modularity and code correctness through an empirical study that provides answers to the following research questions:

1. How prevalent is the presence of ES5 code, modular or not, in real-world applications?
2. Is there a developer intent for independently reusable features within modules?
3. What are the effects of the proposed refactoring to system modularity?
4. Does the proposed source code transformation preserve the external behavior of the analyzed system?

Among the main findings of our empirical study is the wide use of legacy module formats in open source projects and the frequent presence of code patterns in legacy code that reveal a need for individually reusable module features within modules. As concerning the effects on system modularity, our method contributes to fine-grained reuse among system modules and a reduction in module coupling. Finally, the successful execution of the benchmark projects’ test-suites on refactored code provides empirical evidence on the soundness of the proposed source code transformation.

The rest of the paper is organized as follows. Section 2 presents related work and Section 3, provides a brief review on modularity support in ES5 and compares it against the features of ES6. In Section 4, we specify our method for refactoring towards ES6 fine-grained modularity. Section 5 presents the details and results of an empirical evaluation of the method, while Section 6 discusses the threats to its validity. Finally, Section 7 provides a discussion of results and the paper is concluded in Section 8.

2. Related work

Refactoring involves reorganizing source code while preserving its external behavior (Mens and Tourwé, 2004). Fowler (1999) introduced a series of basic refactorings for handling common code smells in object-oriented code and specialized them on the JavaScript language in the second edition of his book (Fowler, 2018). Recent studies and systematic literature reviews focus

on the relationship between code smells and refactoring, especially refactoring towards mitigating architectural problems (Sousa et al., 2020; Soares et al., 2020), and their impact on software quality (Bavota et al., 2015; Lacerda et al., 2020; Fernandes et al., 2020; Agnihotri and Chug, 2020). Ongoing research expands the application of refactoring techniques to address diverse requirements, beyond design improvements in code or other software artifacts, e.g. UML diagrams (Baqais and Alshayeb, 2020). Specifically, several works focus on techniques for improving security (Abid et al., 2020) and maintainability metrics (Nasagh et al., 2020), improving performance through removing unused features (Bruce et al., 2020), reasoning about refactoring activities over long periods of time (Brito et al., 2020) as well as improving software accessibility (Paiva et al., 2020) and achieving business-oriented goals (Ivers et al., 2020).

This work belongs to the research area of refactoring methods and techniques that target the JavaScript language. Relevant works in this area study the automation of refactorings that handle the following concerns: (a) improvement of the code’s runtime performance, (b) improvement of software quality characteristics through generic refactorings or replacement of insecure language features and (c) migration of legacy code to newer language versions. These works are outlined in the rest of this section.

Ying and Miller (2013) leverage static analysis techniques for refactoring AJAX applications to shift from the XML to the JSON data format, in order to improve the runtime performance. Furthermore, static and dynamic analysis techniques are proposed towards automating dead code elimination in JavaScript applications (Obbink et al., 2018; Vazquez et al., 2018). Feldthaus et al. (2011) and Feldthaus and Møller (2013) improve modularity through providing a framework for specifying generic refactorings based on pointer analysis. Jensen et al. (2012) propose a static analysis technique for replacing `eval` with language features that enable reasoning about the program’s behavior.

Several approaches focus on refactoring JavaScript legacy code towards using language features provided in more recent versions of ECMAScript. The prevalence of callbacks in JavaScript applications and their equivalence to ES6 promises is studied in Gallaba et al., Brodu et al. (2015) and Gallaba et al. (2017). The identification of classes and their dependencies in legacy JavaScript is studied in Rostami et al. (2016) and Silva et al. (2017a). Extensions of these works provide migration rules for their refactoring towards using ES6 classes (Silva et al., 2017c,b). Paltoglou et al. (2018) propose a static analysis technique for refactoring non-modular ES5 web applications towards ES6 modules.

In this paper, we propose a method for automated refactoring of a code-base implemented in ES5 or earlier versions, towards fine-grained modularity through ES6 *named imports/exports*. Our work handles concerns related to code migration to a newer language version, as well as to the improvement of its internal quality attributes. As compared to earlier work on refactoring legacy ES5 code to the ES6 class syntax (Silva et al., 2017c,b), our work is complementary and a basic requirement for its application. The reason is that migration to ES6 modularity is a prerequisite for further application of migrations that incorporate modern language features to a code-base. Concerning our previous work on this area (Paltoglou et al., 2018), this paper contributes major extensions in terms of (a) the scope of the refactoring, since it additionally handles AMD/CommonJS codebases, (b) the effect of the refactoring, as it changes module structure and enables effective usage of ES6 modularity constructs and (c) the evaluation methodology that highlights the effectiveness and practicality of the approach.



Fig. 1. Declaration and access to global variables in ES5.

3. Modularity in JavaScript

In this section, we review the evolution of modularity from ES5 to ES6. We briefly present the CommonJS and AMD module formats that are widely used in ES5 and their mostly adopted implementations (Rauschmayer, 2014), followed by the ECMAScript Modules (ESM) feature standardized in ES6. Details on the projects containing the presented fragments are provided in Section 5. We focus on disparities concerning the encapsulation of module and global state, the granularity of exported definitions and the evaluation order of module dependencies. Moreover, we refer to the imposed limitations with respect to the scope of applicability of the provided features. In line with the overview of the features provided by ES6, we compare ES5 formats against ES6 and highlight their incompatibilities, which complicate automated migration of ES5 code to ES6.

3.1. Modularity in ES5

Unlike static programming languages JavaScript did not provide prior to ES6, a formal mechanism for preserving modularity and encapsulation (Osmani, 2012). The community addressed this need through adopting design patterns, e.g. the Singleton pattern (Zakas, 2012), that inject data and functionality belonging to the global namespace into a single global variable.

Although each JavaScript file is meant to bundle coherent functionality, representing a conceptual module, the variables/-functions declared in the file's top-level scope are bound to the global object after the file is loaded. Global variables are, also, introduced in terms of using non-declared variables (implied

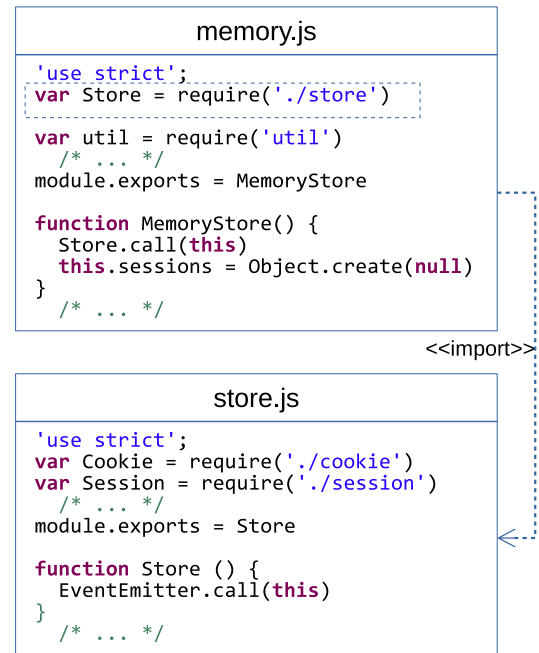


Fig. 2. Module definition and loading using the CJS format.

globals) (Crockford, 2008) in non-strict mode code (Zakas, 2012) and attaching properties to the global object. Fig. 1 depicts fragments from ES5 files of the Hangman project that are loaded in page hangman.html. Upon loading, content.js defines the global variable hintsleft and makes it available to hint.js.

The main problem arising from loading files with non-modular ES5 code is that it complicates program state management due to global variables. The presence of global variables in different program units hampers reasoning about the origin of their values, notably when these units are located in different files (Zakas, 2012). Furthermore, it undermines maintainability, since variables loaded from different files increase the likelihood of name conflicts leading to runtime errors or unexpected behavior (Crockford, 2008). Finally, the lack of a module loading mechanism increases the application's complexity, since the module loading order must be manually specified. The developer community handled these challenges with the CommonJS and AMD module formats, that support the declaration of modules on the server and the browser, respectively (Rauschmayer, 2015).

3.1.1. CommonJS

The CommonJS (CJS) proposal specifies an interface for defining and loading modules on the server (OpenJS, 2020b). A CJS module comprises a piece of reusable JavaScript code which makes objects available to other modules (Osmani, 2012). Fig. 2 presents code fragments from express-session Node.js package.

The top-level variables/functions of a CJS module are bound to the module scope, since CJS restricts the visibility of top-level declarations (OpenJS, 2020b). Similarly to non-modular ES5, global variables are introduced in terms of implied globals in non-strict mode code (Zakas, 2012; Crockford, 2008), and global object properties. CJS modules are synchronously loaded through require (Osmani, 2012), a function which returns the module object, i.e. the object containing the variables/functions that are visible beyond the scope of their declaring module (Rauschmayer, 2015). They behave as singleton objects, since their instances are cached at the time of their first load (OpenJS, 2020b). A variable/-function defined in the module scope becomes available to other

modules through binding to the `module.exports` object or its alias `exports`.¹ CommonJS does not pose limitations regarding the scope of applicability of the `require/export` features, since they can be used in the module's inner scopes (e.g. nested in conditional statements).

As depicted in Fig. 2, `memory.js` imports the function `Store` exported from `store.js`. The function is defined at the time it is exported from `store.js` due to variable/function declaration hoisting (Rauschmayer, 2015), i.e. a reference to `Store` is bound to the nearest scope in which it is defined.

Notice that CJS partially enhances encapsulation, since it does not prevent the introduction of implied global variables (Crockford, 2008). The global namespace “pollution” due to implied globals is mitigated through the `strict` mode directive (Zakas, 2012) that prohibits certain language features. Finally, software portability cannot be guaranteed, since the format targets at defining modules for server-side tasks. The use of design patterns (e.g. Universal Module Definition-UMD UMD, 2020) is needed in order to load CJS modules in the browser. These patterns produce overhead, in terms of memory space and time, in order to resolve the execution environment of the loaded module (Osmani, 2012).

3.1.2. Asynchronous Module Definition (AMD)

The AMD proposal specifies an interface for declaring and loading modules on browsers (AMD JS, 2020). Similarly to non-modular ES5, AMD supports global variables through either the declaration of top-level variables, the definition of implied globals in non-strict mode code (Zakas, 2012; Crockford, 2008) or property bindings to the global object.

AMD modules are loaded either *eagerly* or upon request with the help of the features provided by the format or script loaders, e.g. RequireJS, respectively. By eager loading we mean that upon request for loading an AMD module, all its dependencies are loaded before the execution of the requested module's code. They are instantiated through the function `define` provided by the format (Zakas, 2012), which returns the module object. However, AMD does not provide a feature for the actual loading of modules upon request, and, thus, script loaders (e.g. RequireJS Anon, 2020) are used for module loading and dependency management.

Similarly to the CJS format, the features provided by the AMD format and the respective script loaders can be used in the module's inner scopes. Moreover, AMD modules are loaded in an asynchronous manner. The module loading and execution order cannot be precisely assumed, which can lead to the introduction of runtime errors.² Finally, the problems encountered in defining AMD modules correspond to these encountered in defining CJS modules.

Fig. 3 presents AMD module fragments from the `tetrisJS` application, which are loaded with the RequireJS loader (Anon, 2020). The `TetrisGame.js` module depends on `Score.js`, which is referenced through the entry `app/Score` in the dependency array provided in `define`. The resolution of `Score.js` imports its module object, i.e. the function `Score`, in the scope of the module's definition function.

3.2. ES6 modules

The support for code modularity through the AMD/CJS formats is currently replaced by native language features standardized in ECMAScript 6. These features enable defining platform-independent modules which can be loaded without the need for adopting design patterns (e.g. UMD UMD, 2020).

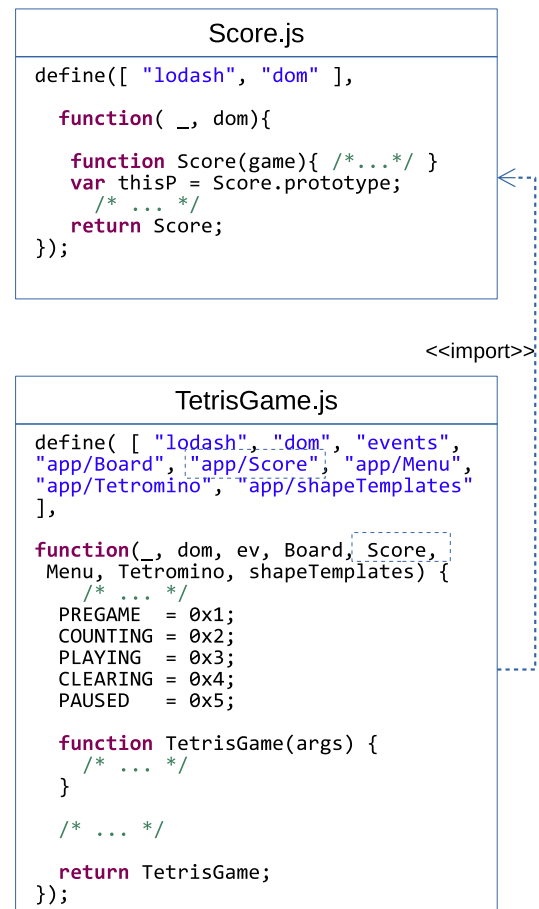


Fig. 3. Module definition and loading using the AMD format.

The variables/functions defined in the top-level scope of an ES6 module are restricted to the module scope (Rauschmayer, 2015). Global variables are introduced in terms of global object properties, since ES6 modules enforce *strict* mode which prohibits the introduction of implied global variables (Zakas, 2012).

ECMAScript 6 provides a declarative syntax for exporting and importing declarations through `export` and `import` statements. Unlike AMD/CJS modules, which export a single *module object*, i.e. a compound object with the module's exported declarations, an ES6 module can apply multiple strategies for managing dependencies. Specifically, it may use *named exports* to export multiple *module features*, i.e. independent variable/function declarations, to client modules. Moreover, an ES6 module may selectively import individual module features from another module through static *named imports*. Importing all module features as a compound object is, also, supported through *namespace imports*. Furthermore, each module may declare a specific feature as a *default export* that is imported in client modules through a *default import* statement. Finally, ES6 enables module loading at module execution time through *dynamic imports*. Fig. 4 presents the equivalent ES6 module of `store.js` (Fig. 2). The module declares the function `Store`, exported through a *named export*, while it imports `Session` from `session.js` through a static *named import*.

Although ES6 standardized features for platform-independent modules, their semantics pose a number of restrictions. Specifically, imported declarations introduce immutable bindings to the respective exported declarations in the module's scope (Rauschmayer, 2015). Moreover, ES6 exports and static imports are permitted only at the top-level scope of the module. The

¹ https://nodejs.org/api/modules.html#modules_exports

² <https://requirejs.org/docs/api.html#data-main>


```

store.js

import {Cookie} from './cookie.js'
import {Session} from './session.js'
/* ... */
function Store () {
  EventEmitter.call(this)
}
/* ... */
export {Store}

```

Fig. 4. Module definition of Fig. 2 in ES6.

establishment of these statements in the module's inner scopes introduces evaluation errors.

Compared to ES5 module formats, ES6 modules support more effective elimination of unused code through the use of *named exports* that simplify the identification of unused module features. Moreover, ES6 module dependencies are eagerly evaluated (ECMA International, 2015), before the actual execution of the module's code. Specifically, the module features that are imported in an ES6 module are resolved during the module's evaluation stage, in contrast to ES5 formats where they are resolved at module execution time (AMD JS, 2020; OpenJS, 2020b). This enables the identification of software defects relevant to non-declared imported features at the module's loading stage instead of its execution stage, thus providing fail-fast behavior (Shore, 2004). Finally, the ES6 module format is platform-independent, enabling module reuse across different platforms (Grover and Kunduru, 2017). This, also, enhances testability (Zakas, 2012) through mitigating the need for simulating the module's execution environment.

4. Refactoring towards using ES6 modules

We propose a method for fully-automated migration of legacy ES5 code to ES6 modules that applies to ES5 projects that are non-modular or employ the AMD/CJS formats as implemented by RequireJS (Anon, 2020) and Node.js (OpenJS, 2020b), respectively. The method focuses on reducing the coupling of refactored modules, through destructuring existing module objects to fine-grained module features and optimization of module dependencies on the basis of the minimum required module features. Moreover, our method restricts the scope of global declarations, and confines them to the module scope as appropriate module features.

Fig. 5 illustrates a motivating example, in order to showcase the application and the merits of the proposed refactoring. It presents code from the `planck.js` 2D physics engine for game development and, specifically, code fragments from the `Solver.js` module and its dependencies. `Solver.js` is a core module of `planck.js` and uses declarations from 9 modules of the same project, e.g. it reads configuration properties from the `Settings.js` module and uses utility functions for vector calculations from the `Vec2.js` module. The code is based on the CJS module format, where the reusable declarations of a module are exported as part of a single module object (top part of Fig. 5). For instance, the `Settings.js` module exports 26 configuration properties, as part of the `Settings` object, `Vec2.js` exports 27 utility functions, as bound properties to the `Vec2` constructor function, and `Math.js` exports 8 functions and a constant as part of the module object. In CommonJS, each client of these modules, e.g., the `Solver.js` module, imports at once all their reusable features through the respective module object. Thus, the module clients depend on complex, coarse-grained objects instead of

the elementary features that they actually use. In the case of `Solver.js` and imported modules `Settings.js`, `Vec2.js` and `Math.js`, the former depends on their entire public API, although it uses 8, 4 and 1 of their exported features, respectively.

The bottom part of Fig. 5 presents an ES6 version of the code fragments from `Solver.js` and its dependencies, after the application of the proposed refactoring. Notice that the independently reusable features of each module, e.g. configuration properties in `Settings.js` or utility functions in `Vec2.js`, are refactored to standalone declarations and exported to client modules through ES6 *named export* statements. The module structure is simplified, while empty module objects are eliminated, as is the case for the `Settings` object. Moreover, all declarations that a module requires from other modules are introduced to its scope through static *named import* statements. The `Solver.js` module, now imports the minimum required declarations from its dependencies.

Fine-grained module design, in terms of multiple independently reusable module features, accompanied with fine-grained module dependencies through named imports, boosts refactorings, code optimizations and error tracing. Specifically, it streamlines the renaming of module features and their relocation among modules. It, also, supports module bundling tools in effective tree-shaking, i.e., identification and elimination of unused code from deployment artifacts (Rollup, 2020; Webpack, 2019). Moreover, it enforces a static module structure which, in turn, paves the way for static variable checking and type inference (Rauschmayer, 2015), enabling, thus, early error detection. Named imports, also, enable fail-fast program behavior, since missing imports are resolved at module evaluation stage (probably at application startup), rather than at a later execution stage (Shore, 2004).

Besides *named imports* or *named exports*, the proposed method uses ES6 *default import* statements for the migration of module dependencies on external libraries, since they support backwards compatibility with ES5 module formats. As concerning ES6 *namespace imports*, i.e. import of all module features as a compound object, they are not used by our method, since they introduce implicit dependencies among modules. Specifically, in a namespace import, the client module does not explicitly specify the module features that it depends upon. Instead, it depends on all the features of the imported module, including currently unused features or features added in subsequent revisions.

In addition to the benefits provided by fine-grained modularity, the proposed refactoring improves code portability and testability (Zakas, 2012), due to the migration of AMD/CJS modules to platform-independent modules and global variable encapsulation. The automation of the proposed refactoring is a challenging task due to incompatibilities of the ES6 module format against ES5 formats, such as read-only bindings for imported module features, top-level declaration of static imports, hoisting and eager loading of static module dependencies, as opposed to on-demand loading in CommonJS etc.

In the rest of this section, we specify our method for automated migration to ES6 modules as two distinct processing stages: (a) identification of refactoring candidates, i.e., code fragments that need to be refactored and (b) application of the source code transformation. The specification of our method, also, defines the refactoring preconditions that preclude transformations leading to erroneous code or influencing the analyzed program's external behavior.

4.1. Refactoring candidate identification procedure

The identification of refactoring candidates involves: (a) the resolution of the structure of each module, in terms of module features that can be individually exported and reused by other



Fig. 5. Motivating example from the project planck.js.

modules, and (b) the resolution of fine-grained module dependencies, on the basis of the minimum required module features that have to be imported by each module after refactoring.

The outcome of this procedure is a directed graph, the Module Dependence Graph (MDG), that models the desired module structure after refactoring and supports the source code transformation procedure. The MDG concept, initially introduced in our previous work (Paltoglou et al., 2018), is redefined here to provide a unified representation of the module structure in modular and non-modular JavaScript code. Specifically, it comprises a pair $MDG = (M, D)$, where M represents the set of graph

nodes. MDG nodes correspond to the ES6 modules that will be introduced after refactoring. The set M is initially populated with one node for each JavaScript file of the analyzed project. The set D represents graph edges. Each edge models a dependency between two modules of the refactored system.

4.1.1. Resolution of module structure

This step involves the identification of the module features for each module in M . Let $m_i \in M$ be a module and n_i be its module name used in AMD/CJS or its module path in the case of non-modular ES5. We define m_i as a pair $m_i = (n_i, F_i)$, where

F_i represents the set of module features of m_i . Module features comprise (a) properties of the module object – in AMD/CJS – that can be extracted to module scope and get individually exported, (b) the module object per se, in case that it is not stripped of all its content, (c) global declarations that are assigned to modules in order to restrict their scope.

Module object destructuring. The identification of object properties that can be extracted as standalone module features and their subsequent extraction from the AMD/CJS module object is termed in this work as *module object destructuring*. The identification procedure involves recovery of the module structure from its AST representation and is based on the analysis of statements that apply property bindings to the module object.

Property binding statements comprise assignments whose left-hand side expression is an object property access through dot/bracket notation, for instance `foo.bar = <expression>`. Properties defined through bracket notation are excluded from further analysis, as explained in 4.1.4. The *bound properties* B_i of module m_i , henceforth, refer to all property bindings on the module object e_i with the dot notation. In case that the module object is instantiated through object literal notation, all properties appearing in the property-value pairs are, also, appended to the set B_i .

The module object e_i exported by module m_i may be constructed within m_i or imported from another module. Its instantiation may be declarative, through a function declaration or object literal notation, or may involve the evaluation of an expression (e.g. function invocation). A module object instantiated as (a) an empty object/function with properties attached at runtime, or (b) through object literal notation, is characterized by our method as a *namespace object*. The reason is that such objects serve as containers for data/ function properties that need to be exported by an AMD/CJS module. Namespace objects would be stripped of all their properties after refactoring, and, thus, are eliminated from the set of module features. Fig. 6 presents a module from `goajs` project that follows the namespace object pattern. The module exports mathematical constants and functions and its original version is depicted in sub-figure (a). Sub-figures (b) and (c) provide syntactic variants that use an empty object or object literal notation.

Let e_i be the module object of a module m_i and F_i the identified module features. The module object destructuring algorithm computes the set F_i for a module m_i and can be summarized as following:

- Step 1.** Evaluate module object destructuring preconditions (see Section 4.1.4). If any precondition fails then return $F_i = \{e_i\}$, i.e., the module object cannot be destructured.
- Step 2.** Identify the set B_i of bound properties to e_i .
- Step 3.** If e_i is a *namespace object* return $F_i = B_i$. Otherwise, return $F_i = B_i \cup \{e_i\}$, i.e., the module object would be available as a module feature after destructuring

Resolution of global declarations. Global declarations involve explicitly declared and implied global variables, properties bound to platform specific global objects (e.g., `window`, `global`) and top-level declarations in non-modular ES5 source files. We identify the global variables of a module m_i through AST traversal and search for (a) top-level `var` statements (non-modular ES5/AMD) and (b) global object property definitions. Implied globals are identified through analysis of AST nodes that represent assignment expressions. Specifically, the variable in the left-hand side of an assignment is sought in the sets of the variables/functions defined within the scope hierarchy formed by the assignment's scope and its surrounding scopes. The variable comprises an implied global if it is not declared in the hierarchy. Finally, top-level

declarations in non-modular ES5 are identified by locating function and variable declaration nodes which do not have function declaration ancestors.

Let m_i be a module and G_i , O_i , T_i refer, respectively to its global variables, global object properties and top-level declarations identified in its AST. The resolution of module features from global declarations involves the following steps:

- Step 1.** Evaluate global declaration preconditions (see Section 4.1.4) on G_i , O_i , T_i . In case of any violation return $F_i = \emptyset$.
- Step 2.** For each identified global object property $o_k \in O_i$ search for its presence in the feature set of other modules
 - (a) If o_k is also defined in module m_j , allocate o_k to the module with lower number of imports, so as to normalize coupling among modules,
 - (b) If not found, then $F_i = F_i \cup \{o_k\}$.
- Step 3.** Return $F_i = F_i \cup G_i \cup T_i$.

In case that global declarations' resolution returns $F_i = \emptyset$ for any module m_i , then the refactoring is abandoned.

4.1.2. Resolution of fine-grained module dependencies

The resolution of the fine-grained module dependencies involves the identification of the module features that each module needs to import from other modules of the same project. Moreover, it identifies dependencies to external library modules that are treated differently for compatibility reasons.

Let $d_j \in D$ be a module dependency from module m_i to a project module m_t . A module dependency represents the import and use in a module m_i of a module feature $f_{t,k} \in F_t$, i.e., exported from module m_t . A module dependency is, also, characterized by the type of use $u_j \in \{R, W, C, L\}$ of $f_{t,k}$ within m_i . If $f_{t,k}$ is only read or called in m_i , then u_j has a value of R or C , respectively. If $f_{t,k}$ is, also, defined in m_i then $u_j = W$. Thus, a module dependency is defined as a tuple $d_j = (m_i, m_t, f_{t,k}, u_j)$. In case that imported module m_t corresponds to an external library, the dependency is represented as $d_j = (m_i, m_t, \emptyset, L)$.

In CommonJS, we identify imported modules through AST traversal and search for `require` invocations. A typical `require` invocation for a module m_j is equivalent to an ES6 namespace import, i.e. the module object exported by m_j is returned and assigned to a variable that provides access to the entire feature set F_j of m_j . A specific property of the imported module m_j may also be directly assigned to a variable, through property access on `require` result. This coding idiom emulates an ES6 named import in CommonJS and the receiving variable provides access to a single feature $f_k \in F_j$. Finally, a module may be imported only for its side effects, i.e., the `require` result is not assigned to any variable.

In the case of AMD, we identify imported modules by analyzing the parameters of the `define` module declaration function. Specifically, the first – optional – parameter of the `define` invocation provides a list with the relative paths or identifiers of imported modules. Each module m_j is resolved and the module object is provided as parameter to the module definition function (second parameter of the `define` invocation). AMD imports are namespace imports, since parameters of the module definition function provide access to the imported modules' entire feature set.

The identification of module dependencies for each module m_i is based on the following algorithm:

- Step 1.** For each imported system module m_j (AMD/CJS) identify the subset of its module features $A_j \subseteq F_j$ used in m_i
 - (a) Let u_k be the type of use of each $a_k \in A_j$,

(a) Function object

```

1  function MathUtils() {}
2
3  MathUtils.DEG_TO_RAD
4    = Math.PI/180.0;
5  /* ... */
6  MathUtils.radFromDeg =
7    function(degrees){/* ... */};
8
9  MathUtils.degFromRad =
10   function(radians){/* ... */};
11 /* ... */
12 module.exports = MathUtils;

```

(b) Empty object

```

var MathUtils = {};
MathUtils.DEG_TO_RAD
  = Math.PI/180.0;
/* ... */
MathUtils.radFromDeg =
  function(degrees){/* ... */};
MathUtils.degFromRad =
  function(radians){/* ... */};
/* ... */
module.exports = MathUtils;

```

(c) Object literal

```

1  var MathUtils = {
2
3    DEG_TO_RAD: Math.PI/180.0,
4    /* ... */
5    radFromDeg: function(degrees){/* ... */},
6    degFromRad: function(radians){/* ... */},
7    /* ... */
8  };
9  module.exports = MathUtils;

```

Fig. 6. Namespace object variants for MathUtils module in goojs project.

- (b) For each $a_k \in A_j$ add a module dependency, i.e. $D = D \cup \{(m_i, m_j, a_k, u_k)\}$.

Step 2. For each imported library module m_j (AMD/CJS) add a module dependency, i.e. $D = D \cup \{(m_i, m_j, \emptyset, L)\}$.

Step 3. For each access to a global declaration $g_k \notin F_i$

- Let u_k be the type of use of g_k . Identify the module m_j that encapsulates the global declaration, i.e., $g_k \in F_j$,
- Add a new module dependency $D = D \cup \{(m_i, m_j, g_k, u_k)\}$.

4.1.3. MDG construction algorithm

On the basis of the aforementioned definitions and algorithms the construction of the MDG data structure for an ES5 project is summarized below:

Step 1. Initialize sets $M = \emptyset$, $D = \emptyset$

Step 2. For each source file f_i with name n_i

- Evaluate module format preconditions on f_i (see Section 4.1.4). In case of any violation, refactoring of f_i is abandoned.
- Resolve the module structure F_i . If $F_i \neq \emptyset$ then $M = M \cup \{(n_i, F_i)\}$ else refactoring of f_i is abandoned.

Step 3. For each module $m_i \in M$

- Resolve module dependencies D_i against modules $M' = M \setminus \{m_i\}$,
- $D = D \cup D_i$

Step 4. Return (M, D) .

Fig. 7 presents a fragment of the MDG constructed during this stage for the project `planck.js`. The fragment focuses on module `Math`, its outgoing dependencies and a subset of its incoming dependencies from other modules. The module's outgoing dependencies correspond to the module features that are imported and used in the module, while incoming dependencies correspond to the module's features that are exported to client modules. For instance, `Math` imports the feature `create` from the module `create`, while it exports the feature `isFinite` to the client modules `Fixture` and `Body`.

4.1.4. Refactoring preconditions

Refactoring preconditions are evaluated by the refactoring identification algorithm on each module of the analyzed system. They prevent evaluation errors in the refactored code and ensure that the external behavior of the system is preserved. The identified preconditions, depending on their scope of application, are grouped in three categories: (a) global declarations' preconditions, (b) module object destructuring preconditions, (c) module format preconditions.

Global declarations' preconditions involve defining global variables and top-level functions with the same name in different modules:

- Each global variable should be introduced in a single location. Conflicting declarations across different files are a potential source of errors.
- In non-modular ES5 or AMD, the names of top-level functions should be unique across all JavaScript files. Different top-level functions with the same name in multiple modules lead to function overriding; the binding in the module's namespace corresponds to the most recently loaded function.

Module object destructuring preconditions involve extracting the properties of the module object that can be independently exported, while preserving the program's external behavior:

- Object properties should be defined with dot notation. Properties defined with bracket notation may lead to inaccurate conclusions regarding their declaration module, in cases that the object is indexed by a variable.
- Module objects should not be fully referenced (e.g. provided in function invocations). Detaching object properties from these objects may lead to errors originating from name aliasing.
- Module objects should not be modified. Detaching properties from such objects may lead to errors related to violations of the ES6 specification (Rauschmayer, 2015).

Module format preconditions involve restrictions that are imposed by the language specification. They, also, involve restrictions that are not imposed by ES5 formats, but are introduced in the ES6 specification. The violation of these restrictions may introduce defects in the refactored code or affect the program's external behavior:

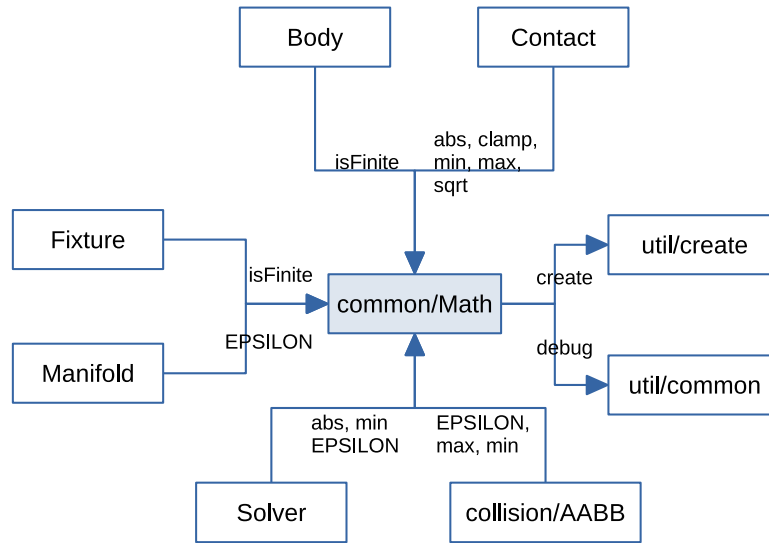


Fig. 7. Module Dependence Graph fragment from planck.js project.

1. Function definitions that reference `this` should be declared in files that enforce strict mode. In case that a function is neither an object method nor a constructor, the value of `this` is determined by the module's code mode. In *non-strict* mode, `this` references the global object. In *strict* mode `this` is undefined, as its value is not updated during the function invocation (MDN Web Docs, 2020). Since ES6 modules enforce *strict* mode (Rauschmayer, 2015), the transformation of an ES5 source file in *non-strict* mode into an ES6 module would lead to runtime errors originating from the value of `this`.
2. Invocations of `require` in CJS modules should be located in the module's top-level scope. Notice that in ES6, the introduction of static imports and exports is allowed only at the module scope. In-place transformation of `require` calls to ES6 static imports would lead to evaluation errors, if they are nested in block statements (Rauschmayer, 2015). This, also, applies to the replacement of return statements with export statements in AMD.

4.2. Source code transformation procedure

The source code transformation for migration to ES6 modules employs the MDG constructed in the refactoring candidate identification stage. Each MDG node, corresponding to a module m_i , is individually refactored in the following steps:

Step 1. If m_i is an AMD module, clear its dependencies to the AMD API:

- (a) Replace the invocation of the `define` function with a list of all statements included in the body of the module definition function.
- (b) Replace the return statement included in the statement list with a variable assignment of the form:


```
var mod_<filename> = <return_expr>;
```

 where `<return_expr>` is the expression returned by the replaced statement.

The application of this step to an AMD module is illustrated in Fig. 8. The affected statements are enclosed in a dashed frame and are marked with the step number (1). In the rest of this section, numbers inside parentheses will be used to refer to similarly marked parts of Figs. 8, 9.

Step 2. If m_i is a CJS module, clear dependencies to the CommonJS API:

- (a) Declare an empty placeholder variable for the module object with the following naming scheme:


```
var mod_<filename> = null;
```
- (b) Replace all references to `module.exports` or its alias `exports` with the above variable.
- (c) Eliminate all `require` statements.

Fig. 9 presents the application of this step to the `Math.js` module of the planck.js project. The newly introduced variable `mod_Math`, replaces `module.exports` and holds the reference to the module object in refactored code (2). Variable `math` remains an alias to the module object.

Step 3. Apply module object destructuring by processing each identified feature $f_j \in F_i$ for the module m_i . Each feature f_j is declared and initialized in ES5 through a property-value pair in object literal notation or through an assignment statement that binds a property to the module object. In order to extract f_j from the module object:

- (a) Introduce a variable declaration for f_j that has the form:


```
var <feature> = <feature_expr>
```

 where `<feature>` corresponds to the name of f_j . In case of name conflicts, the feature name is prefixed with the module name, i.e., `<file_name>_`. The term `<feature_expr>` corresponds to the expression used for initialization of f_j during its declaration.
- (b) Replace references to f_j in module m_i with the feature name.
- (c) Remove the property-value pair or the property binding assignment statement, originally used for f_j declaration.

The declaration of module features in the ES5 version of `Math.js` (left part of Fig. 9), is based on property binding assignments to the `math` alias of the module object. All property bindings can be extracted as accordingly named variable declarations (3), with the exception of `isFinite` that introduces name collision with a global JavaScript function.

in m_i , i.e., $u_j = W$, the corresponding mutator function is, also, imported. The introduced named imports follow the pattern:

```
import {<feature> [as <prefix>_<feature>]} from <file>;
```

- (b) For each imported feature f_j , replace all references to it with the feature name. In case of assignment statements that define the value of f_j , they are replaced with an invocation of the feature's mutator function.
- (c) For each outgoing MDG edge (m_i, m_k, \emptyset, L) to an external library module, introduce a default import, i.e.,

```
import <lib> from <file>;
```

where $\langle \text{lib} \rangle$ corresponds to the variable name used to reference the library module in ES5 code.

Fig. 8 presents the import of two library modules through default import statements (5). In Fig. 9, the `require` statements are replaced by two named imports that introduce the minimum required module features to the namespace of `Math.js` (5).

5. Empirical evaluation

We have designed and performed an experimental study on the proposed method for automated refactoring of ES5 software projects towards ES6 fine-grained modularity. The goal of this study is to analyze the proposed method for the purpose of evaluation with respect to effectiveness and practicality from the perspective of a developer, that considers the implications of the refactoring and its potential for automation, in the context of a set of open source projects.

5.1. Research questions

The experimental study aims at answering the following research questions:

RQ1. *How prevalent is the presence of ES5 code, modular or not, in real-world applications?* At first, we investigate to which extent ES5 code, that is either non modular or supports the AMD/CJS module formats, is still prevalent in web applications. The results of this study will provide empirical evidence on the scope and practicality of the proposed method.

RQ2. *Is there a developer intent for independently reusable features within modules?* Our purpose is to explore at which extent the proposed refactoring to ES6 fine-grained modularity is aligned with the semantics of the analyzed modules and the developer's intent for independently reusable features within modules. The results will support the practicality of the proposed refactoring for effective usage of ES6 language constructs after migration of legacy code.

RQ3. *What are the effects of the proposed refactoring to system modularity?* Our aim is to study the level of granularity, in terms of individually exported module features, that can be introduced in project modules and the potential for encapsulation of module features within their declaring modules. The research question, also, evaluates potential reduction to module coupling due to fine-grained dependencies introduced by ES6 named imports. The results of this research question will highlight the effectiveness of the proposed method for increasing the granularity of reusable module features and reducing the internal coupling of system modules.

RQ4. *Does the proposed source code transformation preserve the external behavior of the analyzed system?* The research question seeks to empirically investigate the soundness of the proposed refactoring, i.e., whether the application of the refactoring does not introduce syntax errors or alter the external behavior of refactored code. The results of this study will further support the practicality of the proposed method.

5.2. Context selection

The context of the experimental study comprises 19 JavaScript projects. The selection of these projects is based on the following criteria: (a) they are open source, for study replication reasons, (b) they are implemented in ES5 or below, (c) they are either non-modular or use the AMD/CJS module formats as implemented in RequireJS/Node.js respectively, (d) their test suite, if available, can be successfully executed, (e) to the extent possible, they are used as benchmarks in earlier work in this area, e.g., Vazquez et al. (2018).

Table 1 presents the projects used in our evaluation, along with relevant implementation details. Column 2 provides the module format used in each project. Notice that the first four projects are non-modular. Column 3 (SLOC) includes the total lines of JavaScript production code, computed with CLOC.³ In case that the codebase of a project includes external libraries, library code is excluded from further analysis. The reason is that these libraries make use of dynamic features of the language (e.g. on-the-fly creation and deletion of object properties, dynamic code generation). These features undermine the results of static analysis (Feldthaus and Möller, 2013) and, thus, their refactoring may introduce errors or influence the program's external behavior. Column 4 provides the test coverage of each project, in terms covered statements by its test suite. Test coverage is not reported for projects without a test suite. Column 5 specifies the tag/commit of the project's revision history that was used in this study and Column 6 provides the projects' Github repository names.

The proposed method is implemented as an open source Node.js tool,⁴ which employs jscodeshift⁵ for Abstract Syntax Tree (AST) traversal and transformation, JSHint⁵ for the resolution of implied globals and Tern⁵ for the dataflow analysis required in the module dependency identification step (4.1.2). The JavaScript code that is embedded in the HTML files of the analyzed web applications is extracted with the help of the jsoup HTML parser⁵. The implemented prototype requires limited user interaction, since the user only needs to specify the application to be refactored and the format that is used for its implementation. The evaluation of the prototype is performed on a workstation, equipped with an Intel-Core i7-7700 CPU @3.60 GHz and 16GB RAM.

5.3. Evaluation results

5.3.1. RQ1. How prevalent is the presence of ES5 code, modular or not, in real-world applications?

The first research question explores how relevant is the proposed method with the needs of current web applications. Since the ES6 (ES2015) specification has been fully implemented in most JavaScript runtimes, a question naturally arises as to what extent web applications still use the legacy module formats and have not yet migrated to ES6. To the best of our knowledge, there is no published study that explores the use of AMD/CJS formats and the current adoption of ES6 modules.

In order to answer RQ1, we analyzed the code-base of all Github repositories that use JavaScript as their main language and have received at least a minimum of developer engagement in the recent years. We acquired access to the source code of the main branch and the public activity timeline of Github repositories, through the Github Activity Data⁶ and GithubArchive⁷ public

³ <https://github.com/AIDanial/cloc>

⁴ <https://github.com/katerinapal/es52es6>

⁵ jscodeshift: <https://github.com/facebook/jscodeshift>, JSHint: <https://github.com/jshint/jshint>, Tern: <https://github.com/ternjs/tern>, jsoup: <https://jsoup.org>.

⁶ <https://console.cloud.google.com/marketplace/product/github/github-repos>

⁷ <https://www.gharchive.org/>

Table 1
JavaScript projects used in the experimental study.

Project	Module Format	SLOC	Test Coverage	Tag/Commit	Github Repository
UltraTetris	-	326	-	master@5ad237e	silviolucenajunior/UltraTetris
Hangman		346	-	master@b950842	aurobindodebnath/Hangman
TicTacToe		919	-	master@346ebe8	seanpr/TicTacToe
uki		2360	-	master@6cd2e47	crcx/uki
GluttonousSnake	AMD	378	-	master@c6b49cb	yyfer/GluttonousSnake
astix		481	-	master@f0ecc39	migace/astix
game-of-life		887	-	master@1d83874	devlysh/game-of-life
tetrisJS		1370	-	master@27712a0	marneborn/tetrisJS
dynablaster-js-port		3679	-	master@5b5052e	gardziej/dynablaster-js-port
backbone-tableview	CommonJS	279	100.0%	master@0c26357	mbrevda/backbone-tableview
easystarjs		369	100.0%	v0.4.3	prettymuchbryce/easystarjs
geojsonhint		473	100.0%	v.2.0.0	mapbox/geojsonhint
express-session		612	99.0%	v1.15.6	expressjs/session
underscore.string		873	99.3%	3.3.4	esamattis/underscore.string
messy		1694	91.8%	6.11.0	papandreou/messy
virtual-dom		1978	97.5%	v2.1.1	Matt-Esch/virtual-dom
recipe-parser		4134	41.0%	master@626f124	catesandrew/recipe-parser
planck.js		10390	40.5%	v0.2.7	shakiba/planck.js
goojs		57373	48.0%	v0.16.8	GooTechnologies/goojs

BigQuery datasets. Among approximately 4 million repositories that include JavaScript code, we narrowed down our analysis to 14,167 repositories on the basis of the following criteria:

- the primary language of the repository, in terms of bytes of code, is JavaScript or HTML. We excluded repositories that include Java, Objective-C or C# code, in order to avoid analyzing cross-platform mobile applications, where migration to ES6 is constrained by application development frameworks.
- each repository has received at least 10 commits and 5 stars since 01.01.2018. Our purpose is to exclude abandoned projects that would bias our results in favor of legacy module formats.

In a next stage, we collected all JavaScript files from the aforementioned repositories and classified them as AMD, CJS or ES6 modules. The classification is based on regular expressions that match the syntax of each module format in the source code of each file. Regular expressions are appropriately formulated so that each file is categorized to a single module format and are executed against a BigQuery table with JS file contents. We excluded from our analysis build tool configuration files (e.g. webpack, rollup), as well as source files located at the installation folder of package dependencies (e.g., `node_modules`).

The results of JS content classification to module formats are available in Table 2. The table provides the number of JS files (Column 2), among all 14,167 repositories, that declare AMD, CommonJS, ES6 modules or do not use any standard module format. Column 3 estimates the share of all files that fall under these categories. Classification results show that a large share of JS content (31.3%) is either non-modular or does not use a standard module format. Moreover, about one-third of all JS files (35.4%) use the standard ES5 module formats, AMD (6.7%) and CJS (28.7%). Finally, the adoption of the ES6 module format characterizes another one-third of all files (33.3%).

AQ 1. *Legacy module formats and non-modular ES5 code are still widely used in actively maintained open source projects. This stresses the need for methods and tools for their automated migration to ES6 modules in order to reap the benefits of new language features.*

5.3.2. RQ2. *Is there a developer intent for independently reusable features within modules?*

The second research question is related to the practicality of the proposed refactoring. Its purpose is to investigate whether the

Table 2
Classification of JS files to module formats.

Module format	Files	%
AMD	65,457	6.7
CommonJS	281,874	28.7
ES6	327,367	33.3
Other	307,523	31.3
Total	982,221	100.0

extracted fine-grained module features correspond to elements with similar semantics in the original module design, i.e., elements that belong to the module scope and may be reused independently from each other. Thus, the language constructs (ES6 named imports/exports) that are automatically applied through refactoring, could have been employed by the programmer if they were supported by ES5.

In order to answer RQ2, we conducted an exploratory study on frequently occurring patterns on the design of module objects. The context of this study comprised the subset of benchmark projects that use the AMD/CJS formats. Non-modular ES5 projects were excluded, since their source files do not declare module objects. The study involved, at first, the identification of common patterns through code inspection on a sample of modules from each benchmark project. Then, we estimated the frequency of each pattern, across all modules, through a static analysis tool implemented for that purpose. Besides the *namespace object* pattern, already described in 4.1.1, code inspection revealed several module object patterns that can be grouped in two categories: (a) *factory* objects and (b) *utility* objects. The categorization focuses on the purpose of module objects and abstracts variations in their design.

A *factory object* is a function that implements a template for object creation and initialization. The function is usually a *function constructor*, a factory method or a generic object configuration function. Its purpose is to initialize a new object with attributes and methods through statements in the function body (property bindings and property assignments on `this` reference). Instance methods are, also, attached through property bindings to the factory object's prototype object. Moreover, the module code, usually, includes statements that apply property bindings directly on the factory object. These bindings are extracted as independent module features by the proposed refactoring. They represent static members of the class declaration that is emulated


```

1  /* ... */
2  module.exports = Vec2;
3  /* ... */
4  function Vec2(x, y) {
5      if (!(this instanceof Vec2)) {
6          return new Vec2(x, y);
7      }
8      if (typeof x === 'undefined') {
9          this.x = 0;
10         this.y = 0;
11     }
12     /* ... */
13 }
14
15 Vec2.zero = function() { /*...*/ };
16
17 Vec2.neo = function(x, y) { /*...*/ };
18
19 Vec2.clone = function(v) { /*...*/ };
20
21 Vec2.prototype.toString = function() {
22     return JSON.stringify(this);
23 };
24
25 Vec2.isValid = function(v) { /*...*/ };
26
27 Vec2.assert = function(o) { /*...*/ };
28 /* ... */

```

Fig. 10. Factory object in Vec2 module of planck.js project.

Table 3
Common patterns of module objects and their frequency.

Project	Module Objects		
	Factory	Namespace	Utility
GluttonousSnake	1 (25.0%)	3 (75.0%)	0 (0.0%)
astix	0 (0.0%)	8 (100%)	0 (0.0%)
game-of-life	4 (66.6%)	1 (16.7%)	1 (16.7%)
tetrisJS	5 (50.0%)	5 (50.0%)	0 (0.0%)
dynablaster-js-port	38 (80.9%)	5 (10.6%)	4 (8.5%)
backbone-tableview	1 (8.3%)	1 (8.3%)	10 (83.4%)
easystarjs	2 (66.7%)	1 (33.3%)	0 (0.0%)
geojsonhint	0 (0.0%)	0 (0.0%)	3 (100%)
express-session	4 (80.0%)	1 (20.0%)	0 (0.0%)
underscore.string	1 (1.0%)	5 (5.2%)	90 (93.8%)
messy	10 (62.5%)	2 (12.5%)	4 (25.0%)
virtual-dom	7 (22.6%)	1 (3.2%)	23 (74.2%)
recipe-parser	1 (6.7%)	2 (13.3%)	12 (80.0%)
planck.js	37 (74.0%)	4 (8.0%)	9 (18.0%)
goajs	433 (81.8%)	74 (14.0%)	22 (4.2%)
Total	544 (65.1%)	113 (13.5%)	178 (21.3%)

within the module, and can be reused independently from the instances created by the factory object.

Fig. 10 presents a factory object declared in Vec2 module of planck.js project. The module object acts as a function constructor or as a factory method, depending on being invoked with the new operator or not. An instance method of Vec2 objects is declared in lines 21–23. On the other hand, direct property bindings to Vec2 object, e.g., zero, neo, represent static utility methods that can be extracted and exported as individual module features.

In the case of *utility objects*, the module object is a non-empty function that serves a domain-specific functionality and does not include bindings to the this reference. Furthermore, its prototype object is not extended within the module scope. Bound properties to this module object are, also, extracted and exported as separate module features after refactoring. However, due to the generic structure of *utility objects*, we cannot generalize on the role of extracted module features in module design. Thus, we assume that the individual reuse of these module features might not express the original intent of the developer.

Table 3 presents static analysis results on the classification of all module objects in AMD/CJS projects to module pattern

Table 4
Exported and encapsulated module features after the refactoring.

Project	Modules	Module Features		
		Total	Exported	Encapsulated
UltraTetris	5	15	12 (80.0%)	3 (20.0%)
Hangman	9	44	27 (61.4%)	17 (38.6%)
TicTacToe	2	12	5 (41.7%)	7 (58.3%)
uki	3	39	5 (12.8%)	34 (87.2%)
GluttonousSnake	4	10	8 (80.0%)	2 (20.0%)
astix	10	49	42 (85.7%)	7 (14.3%)
game-of-life	7	23	16 (69.6%)	7 (30.4%)
tetrisJS	11	41	20 (48.8%)	21 (51.2%)
dynablaster-js-port	49	259	61 (23.6%)	198 (76.4%)
backbone-tableview	12	19	19 (100%)	0 (0.0%)
easystarjs	3	11	11 (100%)	0 (0.0%)
geojsonhint	3	3	3 (100%)	0 (0.0%)
express-session	5	9	9 (100%)	0 (0.0%)
underscore.string	97	191	166 (86.9%)	25 (13.1%)
messy	16	111	111 (100%)	0 (0.0%)
virtual-dom	31	45	44 (97.8%)	1 (2.2%)
recipe-parser	16	65	53 (81.5%)	12 (18.5%)
planck.js	55	285	276 (96.8%)	9 (3.2%)
goajs	540	2045	1734 (83.8%)	332 (16.2%)
Total	878	3276	2601 (79.4%)	675 (20.6%)

categories. Specifically, Columns 2–4 provide for each project the number of module objects that fall, respectively, into the factory, namespace and utility object categories. The results show that the majority of module objects, i.e., 65.1%, are classified as factory objects. Moreover, a smaller but significant share of module objects (13.5%) are classified as namespace objects, while utility objects correspond to 21.3% of all module objects. Thus, a total of 78.6% of module object declarations follow the *namespace* or *factory* object pattern, i.e., their module scope emulates, either a container for features or a class. In each one of these declarations, the module object's bound properties represent features that reside semantically at the module scope and can be individually reused as ES6 named exports. Bound properties in the rest of module declarations, i.e., in *utility* objects, may also be reusable as individual features, but their relevance must be inspected on a case-by-case basis.

AQ2. Bound properties in 78.6% of module object declarations correspond to namespace elements or static features of an emulated class, revealing a developer intent for their individual reuse at module level.

5.3.3. RQ3: What are the effects of the proposed refactoring to system modularity?

In order to answer RQ3, we evaluate the module structure after refactoring, where module objects are destructured and replaced by individually reusable module features. We investigate the reusability of module features across modules, as well as the impact of fine-grained dependencies to module coupling.

The effect of the proposed refactoring on the granularity of independently reusable elements within a project, can be inferred by Columns 2, 3 of Table 4. Column 2 (Modules) provides the number of module objects declared within each project, while Column 3 (Total Module Features) sums up all module features, i.e., (a) bound properties that are extracted from each module object and (b) global declarations that are allocated and accessed through modules after refactoring. Initially, the units of reuse were coarse-grained module objects, each one declared in its own source file. The proposed method sets the unit of reuse at the granularity of module features, with multiple feature declarations and exports per source file. The figures in Columns

2 and 3 show that this shift in granularity results to an increase in the number of reusable elements per project by a factor of 4 on average. This increase in fine-grained reusable elements reduces module coupling, as will be demonstrated hereafter, while supporting application bundling tools in effective tree-shaking, i.e., identification and elimination of unused code from deployment artifacts (Rollup, 2020; Webpack, 2019).

The module features that are, actually, reusable within each project are provided in Column 4 (Exported Module Features). These module features are required and used by at least one module, besides their declaring module. After refactoring, they are declared as *named exports* and are reused in client modules through corresponding *named import* statements. Moreover, the column provides, inside parentheses, the respective fraction of reusable module features against total module features of the project (Column 2). According to figures in Column 4, 79.4% of all module features are reusable and should be exported, while, in 13 out of 19 projects, these module features exceed 80%.

However, a significant part of module features, especially in non-CommonJS projects, may not be exported but encapsulated in their declaring modules. Notice that CJS projects implement Node.js library packages and, thus, their encapsulated features are estimated by integrating in the static analysis their unit tests, in order to act as client code of their public API. In case that unit tests do not provide sufficient coverage of the public API, encapsulation of module features can be disabled. As revealed by results in Column 5, 20.6% of all module features and 14.3%–87.2% in non-modular ES5 and AMD projects are not reused outside their declaring module. In any case, the results provide evidence on the poor support for module encapsulation in ES5. The proposed refactoring omits ES6 named export statements for these features, restricting their scope and narrowing the public API of modules. Thus, it improves modularity through better encapsulation.

As concerning the effect of refactoring on module coupling, we investigate it by comparing the average module coupling of each project before and after the application of refactoring. Module coupling estimation is based on the incoming dependencies (*Fan-in*) and the outgoing dependencies (*Fan-out*) of a module. Moreover, we estimate *Module Instability* in analogy to the well-established *Instability* metric for components or packages (Martin, 2003). We draw an analogy between an ES6 module and a Java package⁸ in order to estimate module *Fan-in* and *Fan-out*. A Java package groups classes (and other types or packages) and makes them accessible to other packages through the public class modifier. Its outgoing dependencies comprise classes of other packages that are imported by its own classes, while incoming dependencies comprise classes of other packages that import one or more of its own classes. In a similar vein, an ES6 module groups module features, that are declared in its source file, and become accessible to other modules through export statements. An ES6 module has outgoing dependencies to module features that are imported individually or in bulk (named or default imports) from other modules. Incoming dependencies of a module comprise the module features of other modules that import one or more of its own features.

Let m_i be a module, F_i its module features and D_i its set of outgoing dependencies in the MDG, according to the notation introduced in Section 4.1. Moreover, let $F_{i,out}$ be the set of module features that m_i depends on and imports into its source file (external library modules are excluded). The following equation provides a formal specification for $F_{i,out}$:

$$F_{i,out} = \bigcup \{f_{t,k} \mid \exists d \in D_i \text{ s.t. } d = (m_i, m_t, f_{t,k}, _) \wedge m_t \in M \wedge f_{t,k} \in F_t\}$$

⁸ Or an equivalent construct of another programming language, e.g. C# namespace.

The set $F_{i,in}$ of module features that depend on module m_i contains all module features whose declaration references at least one module feature from m_i . Let m_s be a module with outgoing dependencies to m_i in the MDG, i.e., it imports one or more features from m_i . Moreover, let $A_{i,s} \subseteq F_s$ be the module features of m_s that access at least one imported feature from module m_i . We compute $F_{i,in}$ as the union of all sets $A_{i,s}$, where m_s is a module that imports features from m_i . A formal specification of $F_{i,in}$ is provided below:

$$F_{i,in} = \bigcup A_{i,s}, \text{ for each } m_s \in M \text{ s.t. } (\exists d \in D_s \mid d = (m_s, m_i, f_{i,k}, _) \wedge f_{i,k} \in F_i)$$

We estimate metrics *Fan-out* (FO_i), *Fan-in* (FI_i) and *Module Instability* (I_i) for each module m_i , on the basis of $F_{i,out}$, $F_{i,in}$:

$$FO_i = |F_{i,out}|, FI_i = |F_{i,in}|$$

$$I_i = \frac{|F_{i,out}|}{|F_{i,out}| + |F_{i,in}|}$$

Table 5 presents module coupling and instability results for each project, before and after refactoring. Each value is computed as the average of a given metric over all modules in the production code of the respective project. The standard deviation is provided inside parentheses. Columns 2–3 present the average *Fan-out* (FO_i), for the modules of each project, before (ES5) and after refactoring (ES6). The results show a significant decrease of module *Fan-out* in most projects, marginal decrease in one project and no change in 5 of them. The improvement of module *Fan-out* is due to the fine-grained reuse of module features after refactoring, where each module imports the minimum required features from its dependencies. On the other hand, the ES5 version of modules, usually, import entire modules and depend on all features bound to the respective module objects. The average *Fan-in* (FI_i) of modules for each project is presented in Column 4. The values for *Fan-in* are not affected by the refactoring. The reason is that the incoming dependencies of a module remain unchanged, despite its import in client modules as individual module features rather than as a coarse-grained module object.

The effect of the refactoring on *Module Instability* is analyzed in Columns 5–6 that provide the average instability of modules across all projects, before and after refactoring. Its average improvement is provided in Column 7. Following the decrease of module *Fan-out*, the instability of modules improves by 9–60% in 10 projects and 1%–2% in 3 projects. However, in projects Tic-TacToe and backbone-tableview module instability remains the same despite the improvement in *Fan-out*. These projects contain modules with either (a) $FI_i = 0$, i.e. their client modules do not define module features that depend on the exported features, or (b) $FO_i = 0$, i.e. they do not define module features that depend on the imported features. These modules are not considered in the calculation of I_i after refactoring.

AQ 3. The proposed method increases the number of reusable elements per refactored project by a factor of 4 on average, and enables fine-grained reuse through declaration of multiple reusable elements per module. It enforces module encapsulation by restricting the scope of module features that are not used by other modules and reduces module coupling in terms of decreased module *Fan-out* and module instability.

5.3.4. RQ4: Does the proposed source code transformation preserve the external behavior of the analyzed system?

To evaluate how effectively the proposed method preserves the external behavior of each analyzed project, we executed the project's test suite after the proposed refactoring. Its successful execution provides an indicator for the reliability of the applied

Table 5
Module coupling and instability metrics before and after refactoring.

Project	Fan-out (FO_i)		Fan-in (FI_i)	Module Instability (I_i)		
	ES5	ES6		ES5	ES6	ΔI_i %
UltraTetris	3.20 (0.45)	0.80 (1.30)	0.80 (0.45)	0.80 (0.11)	0.35 (0.49)	60 (55)
Hangman	21.33 (2.29)	2.56 (3.57)	9.11 (13.15)	0.79 (0.25)	0.37 (0.42)	58 (44)
TicTacToe	2.50 (3.54)	1.50 (2.12)	0.50 (0.71)	0.50 (0.71)	0.50 (0.71)	0 (0)
uki	10.00 (1.00)	2.67 (3.79)	1.00 (1.00)	0.91 (0.09)	0.44 (0.51)	53 (50)
GluttonousSnake	1.75 (2.87)	1.75 (2.87)	0.75 (0.50)	0.46 (0.54)	0.46 (0.54)	0 (0)
astix	12.80 (13.61)	10.40 (10.95)	2.60 (3.10)	0.71 (0.33)	0.64 (0.44)	25 (46)
game-of-life	0.71 (1.11)	0.71 (1.11)	1.43 (1.62)	0.32 (0.43)	0.32 (0.43)	0 (0)
tetrisJS	10.27 (5.50)	1.18 (2.04)	1.91 (2.30)	0.77 (0.31)	0.37 (0.39)	51 (42)
dynablaster-js-port	3.27 (3.83)	2.02 (2.76)	2.73 (4.30)	0.52 (0.39)	0.44 (0.37)	11 (19)
backbone-tableview	2.08 (2.94)	1.42 (2.31)	1.08 (1.00)	0.39 (0.41)	0.39 (0.41)	0 (0)
easystarjs	0.67 (1.15)	0.67 (1.15)	0.67 (0.58)	0.33 (0.58)	0.33 (0.58)	0 (0)
geojsonhint	0.50 (0.58)	0.50 (0.58)	0.50 (0.58)	0.50 (0.50)	0.50 (0.50)	0 (0)
session	1.40 (1.67)	1.00 (1.22)	8.80 (6.38)	0.24 (0.43)	0.23 (0.43)	9 (21)
underscore.string	2.63 (9.84)	1.82 (7.58)	1.02 (3.61)	0.71 (0.34)	0.70 (0.34)	1 (7)
messy	2.94 (4.95)	2.69 (4.88)	6.31 (10.68)	0.40 (0.43)	0.39 (0.42)	1 (3)
virtual-dom	2.30 (3.88)	1.73 (3.01)	1.45 (1.79)	0.49 (0.43)	0.48 (0.42)	2 (5)
recipe-parser	35.62 (23.87)	1.06 (1.61)	13.88 (29.81)	0.71 (0.40)	0.28 (0.38)	52 (44)
planck.js	70.95 (55.75)	10.22 (7.97)	24.42 (55.98)	0.73 (0.39)	0.60 (0.40)	20 (27)
goojs	26.40 (44.47)	16.20 (39.25)	10.74 (43.94)	0.70 (0.37)	0.63 (0.41)	19 (34)

refactoring (Gligoric et al., 2013). For projects without a test suite, we empirically evaluated the correctness of the refactored code through manual code inspection and system usage.

Evaluation results provide empirical evidence on the soundness of the code transformation, since the test suites and the manual inspection did not reveal changes to the external behavior of refactored code. Special attention needs to be attributed to the `goojs` project, where the transformation was applied despite the violation of a precondition. Specifically, `goojs` modules contain module dependencies that are established in inner scopes (e.g. inside conditional statements). Therefore, their code violates the module format precondition with respect to nested module dependencies (4.1.4). Nevertheless, these dependencies did not exhibit side-effects (e.g. modifications to the global scope) and the establishment of these dependencies at the module scope after refactoring did not alter the external behavior of the system.

AQ 4. *The proposed transformation preserves each analyzed system's external behavior, even in case that it violates the precondition relevant to module dependencies established in inner scopes.*

6. Threats to validity

In this section, we discuss threats to the validity of empirical evaluation results. We categorize and prioritize the threats to validity according to Wohlin et al. (2000): internal, external, construct and conclusion validity.

Internal validity threats refer to factors that affect the results of this study but are ignored or cannot be controlled. A potential threat relates to the quality of the benchmark projects' test suites that affects the conclusions of RQ4. Since the statement coverage of test suites in CJS projects ranges between 40.05%–100% and in 7 of them exceeds 90% (Table 1), we expect that they are capable of revealing the majority of flaws of the proposed method.

External validity threats affect the generalization of empirical evaluation results. A potential threat concerns the extent at which the selected sample of 19 projects is representative of the population of legacy ES5 projects. We deem that the sample has adequate size and diversity in order to mitigate this threat, since it includes both standalone applications and library projects, with various sizes and coming from different domains. The general applicability of the proposed method is, also, affected by some limitations on handling specific features of the JavaScript language. Specifically, our method does not analyze

code which uses dynamic code generation and name aliasing, e.g. resolving modules with variant file paths generated using `__filename` or `__dirname`, properties defined programmatically through the `Object.defineProperty()` method. Static analysis introduces significant challenges in these cases (Jensen et al., 2012) that may undermine the safety of the proposed refactoring. Finally, the method analyzes to a limited extent language features that are forbidden in strict mode (MDN Web Docs, 2020), in order to prevent the introduction of defects due to incompatibilities between the ES5 and ES6 specifications. Among benchmark projects, `GluttonousSnake`, `astix`, `game-of-life`, `UltraTetris`, `tetrisJS` and `goojs` make limited use of such features. In order to include them in this study, we manually replace them with equivalent features, permitted in strict mode.

Threats to *construct validity* concern the correspondence between theoretical constructs and the observations of the empirical evaluation. A potential threat relates to RQ1 and the regex classifiers that we employed in order to categorize JS files to module formats. In order to mitigate this threat we built a test suite for these regular expressions that is publicly available in Github.⁹ Another threat relates to the correct estimation of metrics that support the conclusions of RQ2 and RQ3. In order to contain this threat with performed manual validation of metric calculations for small projects that are included in this study.

Conclusion validity concerns the statistical significance of the results on cause–effect relations studied in the experiment. Since, we conduct an exploratory study, where experimental results and observations shape the findings of the study rather than confirm the validity of certain hypotheses, we evaluate *reliability* as a counterpart of conclusion validity (Wohlin et al., 2000). Threats to *reliability* concern the reproducibility of this study and we mitigate them through publishing the source code and all required artifacts for its replication.

7. Results and discussion

The proposed method automates a large refactoring that migrates an ES5 codebase to ECMAScript 6. The refactoring focuses on the application and effective use of the ES6 language constructs for modularity. This migration represents an important milestone to the evolution of a legacy JavaScript project, since the use of the ES6 module system is a basic requirement for further

⁹ <https://github.com/softeng-aueb/es-module-classifier>

adoption of language features (e.g. class syntax, promises API) introduced in ES6 or beyond, that boost developer productivity and improve code maintainability.

The scope of the proposed refactoring method is not limited to plain syntax transformation, that is supported to an extent by some module bundlers (Rollup, 2020). It additionally alters the design of refactored modules to enable fine-grained reuse of module contents, through destructuring of module objects to multiple individually reusable module features. Module dependencies are, then, redefined and established on the basis of these module features. This improves maintainability, since coupling between the refactored modules is reduced; each refactored module imports the module features that are used by its features, instead of entire module objects. It, also, enables encapsulation, since unused module features are restricted to the scope of their declaration module.

Our method achieves fine-grained reuse of module contents through effective use of the *named* imports/exports language constructs, instead of their *default* or *namespace* equivalents that establish dependencies on the basis of coarse-grained module objects. Although the latter are used for backwards compatibility with ES5 module formats, named imports/exports provide significant benefits.¹⁰ Specifically, they enforce name consistency across modules, which improves code understandability and allows for effective “tree-shaking” during code deployment. Moreover, they facilitate other refactoring operations, such as rename refactorings and move feature or extract module refactorings for remodularization. Finally, the evaluation part of this study contributes to the ongoing debate on *named* vs. *default* imports/exports¹¹ by highlighting the former’s positive effects on module coupling and reusability.

8. Conclusions — future work

We have proposed a method for automated refactoring of legacy ES5 code, that is either non-modular or uses the AMD/CJS formats, to ES6 modularity with extensive use of the *named import/export* language constructs. Besides migration of ES5 module formats to ES6 modules, the proposed method improves internal quality attributes of refactored code. Specifically, it enforces fine-grained reuse of module contents, through destructuring of module objects to individually reusable module features. On the basis of these features, it optimizes module dependencies by leveraging the ES6 syntax. The core of our method is a static analysis technique that constructs a global model of the analyzed project, the Module Dependence Graph (MDG). The MDG mirrors inferred or explicitly declared modules and their dependencies. On the basis of MDG, we specify the source code transformation for migration to ES6.

The proposed method has been implemented in the Node.js platform for the purpose of evaluating its effectiveness and practicality on a collection of open source projects. Evaluation results reveal a developer intent for fine-grained reuse, since 78.6% of extracted features have semantics that correspond to reusable elements at module scope (namespace properties, static features of emulated classes). The analysis of refactored code shows an increase in the number of reusable elements per project by a factor of four and an improvement in the coupling of refactored modules in terms of decreased module *Fan-out* and module instability. Finally, the execution of projects’ test-suites on refactored

code provides evidence on the soundness of the source code transformation.

Our future work will investigate the potential for further improvements to the modularity of an ES6 codebase through automated refactorings. Specifically, we intend to extend our research across two directions: (a) improvement of module cohesion through redistribution of fine-grained module features among modules, extraction of module features to new modules or consolidation of the contents of different modules, and (b) modernization of the internal structure of modules through the introduction of the ES6 class syntax.

CRedit authorship contribution statement

Katerina Paltoglou: Conceptualization, Methodology, Software, Writing, Validation. **Vassilis E. Zafeiris:** Conceptualization, Writing, Methodology, Validation. **N.A. Diamantidis:** Supervision, Conceptualization. **E.A. Giakoumakis:** Supervision, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The research work was supported by the Hellenic Foundation for Research and Innovation (HFRI), Greece under the HFRI PhD Fellowship grant (Fellowship Number: 696). The authors would, also, like to thank the anonymous reviewers for their useful comments that improved the quality of this work.

References

- Abid, C., Alizadeh, V., Kessentini, M., do Nascimento Ferreira, T., Dig, D., 2020. 30 years of software refactoring research: A systematic literature review. *arXiv:2007.02194*.
- Agnihotri, M., Chug, A., 2020. A systematic literature survey of software metrics, code smells and refactoring techniques. *J. Inf. Process. Syst.* 16, 915–934.
- AMD JS, 2020. Asynchronous module definition (AMD) API. <https://github.com/amdjs/amdjs-api>, Accessed: 2020.
- Anon, 2020. Requirejs module loader. <https://requirejs.org>, Accessed: 2020.
- Baqais, A.A.B., Alshayeb, M., 2020. Automatic software refactoring: A systematic literature review. *Softw. Qual. J.* 28, 459–502.
- Bavota, G., De Lucia, A., Di Penta, M., Oliveto, R., Palomba, F., 2015. An experimental investigation on the innate relationship between quality and refactoring. *J. Syst. Softw.* 107, 1–14.
- Brito, A., Hora, A.C., Valente, M.T., 2020. Refactoring graphs: Assessing refactoring over time. In: 2020 IEEE 27th International Conference on Software Analysis, Evolution and Reengineering (SANER), pp. 367–377.
- Brodu, E., Frénot, S., Oblé, F., 2015. Toward automatic update from callbacks to promises. In: Proceedings of the 1st Workshop on All-Web Real-Time Systems. AWeS ’15, ACM.
- Bruce, B.R., Zhang, T., Arora, J., Xu, G.H., Kim, M., 2020. Jshrink: In-depth investigation into debloating modern java applications. In: Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering. ESEC/FSE 2020, ACM, pp. 135–146.
- Crockford, D., 2008. JavaScript: The Good Parts. O’Reilly Media, Inc..
- ECMA International, 2015. ECMAScript 2015 language specification. <https://www.ecma-international.org/ecma-262/6.0/>.
- Feldthaus, A., Millstein, T., Möller, A., Schä, M., Tip, F., 2011. Tool-supported refactoring for javascript. In: Proceedings of the 2011 ACM Int. Conf. on Object Oriented Programming Systems Languages and Applications. OOPSLA ’11, ACM, pp. 119–138.
- Feldthaus, A., Möller, A., 2013. Semi-automatic rename refactoring for javascript. *SIGPLAN Not.* 48, 323–338.
- Fernandes, E., Chávez, A., Garcia, A., Ferreira, I., Cedrim, D., Sousa, L., Oizumi, W., 2020. Refactoring effect on internal quality attributes: What haven’t they told you yet? *Inf. Softw. Technol.* 106347.

¹⁰ <https://esdiscuss.org/topic/moduleimport>

¹¹ N.Zakas, <https://humanwhocodes.com/blog/2019/01/stop-using-default-exports-javascript-module/>, A. Rauschmayer, <https://medium.com/@rauschma/note-that-default-exporting-objects-is-usually-an-anti-pattern-if-you-want-to-export-the-cf674423ac38>.

- Fowler, M., 1999. *Refactoring: Improving the Design of Existing Code*, first ed. Addison-Wesley Longman Publishing Co. Inc., Boston, MA, USA.
- Fowler, M., 2018. *Refactoring: Improving the Design of Existing Code*, second ed. Addison-Wesley Longman Publishing Co. Inc., Boston, MA, USA.
- Gallaba, K., Hanam, Q., Mesbah, A., Beschastnikh, I., 2017. Refactoring asynchrony in javascript. In: 2017 IEEE Int. Conf. on Software Maintenance and Evolution (ICSME), pp. 353–363.
- Gallaba, K., Mesbah, A., Beschastnikh, I., 2015. Don't call us, we'll call you: Characterizing callbacks in javascript. In: 2015 ACM/IEEE Int. Symp. on Empirical Software Engineering and Measurement (ESEM).
- Gligoric, M., Behrang, F., Li, Y., Overbey, J., Hafiz, M., Marinov, D., 2013. Systematic testing of refactoring engines on real software projects. In: *Proceedings of the 27th European Conf. on Object-Oriented Programming. ECOOP'13*, Springer-Verlag, pp. 629–653.
- Grover, D., Kunduru, H.P., 2017. *ES6 for Humans: The Latest Standard of JavaScript ES2015 and beyond*, first ed. Apress, USA.
- Ivers, J., Ozkaya, I., Nord, R.L., Seifried, C., 2020. Next generation automated software evolution refactoring at scale. In: *Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, pp. 1521–1524.
- Jensen, S.H., Jonsson, P.A., Möller, A., 2012. Remediating the eval that men do. In: *Proceedings of the 2012 Int. Symp. on Software Testing and Analysis. ISTA 2012*, ACM, pp. 34–44.
- Lacerda, G., Petrillo, F., Pimenta, M., Gué, Y.G., 2020. Code smells and refactoring: A tertiary systematic review of challenges and observations. *J. Syst. Softw.* 167.
- Martin, R.C., 2003. *Agile Software Development: Principles, Patterns, and Practices*. Prentice Hall PTR, Upper Saddle River, NJ, USA.
- MDN Web Docs, 2020. Javascript reference. <https://developer.mozilla.org/en-US/docs/Web/JavaScript/Reference>.
- Mens, T., Tourwé, T., 2004. A survey of software refactoring. *IEEE Trans. Softw. Eng.* 30, 126–139.
- Nasagh, R.S., Shahidi, M., Ashtiani, M., 2020. A fuzzy genetic automatic refactoring approach to improve software maintainability and flexibility. *Soft Comput.* 1–31.
- Obbink, N.G., Malavolta, I., Scoccia, G.L., Lago, P., 2018. An extensible approach for taming the challenges of javascript dead code elimination. In: 2018 IEEE 25th International Conference on Software Analysis, Evolution and Reengineering (SANER), pp. 291–401.
- OpenJS, 2020a. Node.js v12.6.0 documentation: ECMAScript modules. <https://nodejs.org/api/esm.html>, Accessed: 2020.
- OpenJS, 2020b. Node.js documentation: CommonJS modules. <https://nodejs.org/api/modules.html>.
- Osmani, A., 2012. *Learning JavaScript Design Patterns*. O'Reilly Media, Inc..
- Paiva, D.M.B., Freire, A.P., de Mattos Fortes, R.P., 2020. Accessibility and software engineering processes: A systematic literature review. *J. Syst. Softw.* 110819.
- Paltoglou, A., Zafeiris, V.E., Giakoumakis, E.A., Diamantidis, N., 2018. Automated refactoring of client-side javascript code to es6 modules. In: 2018 IEEE 25th Int. Conf. on Software Analysis, Evolution and Reengineering. SANER, IEEE, pp. 402–412.
- Rauschmayer, A., 2014. *Speaking JavaScript*, first ed. O'Reilly Media, Inc..
- Rauschmayer, A., 2015. *Exploring ES6. Upgrade to the next version of javascript*. Rollup, 2020. Rollup module bundler. <https://rollupjs.org>, Accessed: 2020.
- Rostami, S., Eshkevari, L., Mazinanian, D., Tsantalis, N., 2016. Detecting function constructors in javascript. In: 2016 IEEE Int. Conf. on Software Maintenance and Evolution (ICSME), pp. 488–492.
- Shore, J., 2004. Fail fast. *IEEE Softw.* 21, 21–25. <http://dx.doi.org/10.1109/MS.2004.1331296>.
- Silva, L.H., Valente, M.T., Bergel, A., 2017. Statically identifying class dependencies in legacy javascript systems: First results. In: 2017 IEEE 24th Int. Conf. on Software Analysis, Evolution and Reengineering (SANER).
- Silva, L.H., Valente, M.T., Bergel, A., 2017. Refactoring legacy JavaScript code to use classes: The good, the bad and the ugly. In: 16th Int. Conf. on Software Reuse (ICSR), pp. 1–16.
- Silva, L.H., Valente, M.T., Bergel, A., Anquetil, N., Etien, A., 2017c. Identifying classes in legacy javascript code. *J. Softw.: Evol. Process* 1, 1–37.
- Soares, E., Ribeiro, M., Amaral, G., Gheyi, R., Fernandes, L., Garcia, A., Fonseca, B., Santos, A., 2020. Refactoring test smells: A perspective from open-source developers. In: *Proceedings of the 5th Brazilian Symposium on Systematic and Automated Software Testing*, pp. 50–59.
- Sousa, L., Oizumi, W., Garcia, A., Oliveira, A., Cedrim, D., Lucena, C., 2020. When are smells indicators of architectural refactoring opportunities: A study of 50 software projects. In: *Proceedings of the 28th International Conference on Program Comprehension*, pp. 354–365.
- UMD, 2020. Universal module definition (UMD) API. <https://github.com/umdjs/umd>, Accessed: 2020.
- Vazquez, H., Bergel, A., Vidal, S., Pace, J., Marcos, C., 2018. Slimming javascript applications: an approach for removing unused functions from javascript libraries. *Inf. Softw. Technol.*
- Webpack, 2019. Webpack module bundler. Tree shaking. <https://webpack.js.org/guides/tree-shaking/>, Accessed: Nov. 2019.
- Wirfs-Brock, A., Eich, B., 2020. Javascript: The first 20 years. *Proc. ACM Program. Lang.* 4, <http://dx.doi.org/10.1145/3386327>.
- Wohlin, C., Runeson, P., Hö, M., Ohlsson, M.C., Regnell, B., Wesslén, A., 2000. *Experimentation in Software Engineering: An Introduction*. Kluwer Academic Publishers, Norwell, MA, USA.
- Ying, M., Miller, J., 2013. Refactoring legacy ajax applications to improve the efficiency of the data exchange component. *J. Syst. Softw.* 86, 72–88.
- Zakas, N., 2012. *Maintainable JavaScript*. O'Reilly Media, Inc..

Katerina Paltoglou received her BS degree in Informatics from the Department of Informatics & Telecommunications, National and Kapodistrian University of Athens, in 2014 and her M.Sc. in Information Systems from the Athens University of Economics and Business in 2017. Her research interests include software analysis and refactoring.

Vassilis E. Zafeiris is a Researcher and Teaching Staff in the Department of Informatics at the Athens University of Economics and Business. He received his Dipl. in Electrical Engineering from the National Technical University of Athens (NTUA) in 2001, his M.Sc. in Information Systems and Ph.D. in Informatics from the Athens University of Economics and Business in 2003 and 2011 respectively. His research interests include software refactoring, mining of software repositories and testing of mobile applications.

Nikos Diamantidis received his BS degree and Ph.D. in Informatics from the Athens University of Economics and Business in 1993 and 1998 respectively. His research interests include machine learning and medical diagnosis, software engineering.

Emmanuel A. Giakoumakis is a professor of Software Engineering in the Informatics Department of the Athens University of Economics and Business. He received his Dipl. in Electrical Engineering and Ph.D. in Computer Engineering from the National Technical University of Athens (NTUA) in 1981 and 1988 respectively. His research interests include software engineering, biomedical applications, regulation and competition in electronic services.