A Survey on Software Coupling Relations and Tools

Enrico Fregnan

University of Zurich, Switzerland

Tobias Baum

Leibniz Universität Hannover, Germany

Fabio Palomba and Alberto Bacchelli

University of Zurich, Switzerland

Abstract

Context: Coupling relations reflect the dependencies between software entities and can be used to assess the quality of a program. For this reason, a vast amount of them has been developed, together with tools to compute their related metrics. However, this makes the coupling measures suitable for a given application challenging to find.

Goals: The first objective of this work is to provide a classification of the different kinds of coupling relations, together with the metrics to measure them. The second consists in presenting an overview of the tools proposed until now by the software engineering academic community to extract these metrics.

Method: This work constitutes a systematic literature review in software engineering. To retrieve the referenced publications, publicly available scientific research databases were used. These sources were queried using keywords inherent to software coupling. We included publications from the period 2002 to 2017 and highly cited earlier publications. A snowballing technique was used to

*Corresponding author

Email addresses: fregnan@ifi.uzh.ch (Enrico Fregnan), tobias.baum@inf.uni-hannover.de (Tobias Baum), palomba@ifi.uzh.ch, bacchelli@ifi.uzh.ch (Fabio Palomba and Alberto Bacchelli)
retrieve further related material.

**Results:** Four groups of coupling relations were found: structural, dynamic, semantic and logical. A fifth set of coupling relations includes approaches too recent to be considered an independent group and measures developed for specific environments. The investigation also retrieved tools that extract the metrics belonging to each coupling group.

**Conclusion:** This study shows the directions followed by the research on software coupling: e.g., developing metrics for specific environments. Concerning the metric tools, three trends have emerged in recent years: use of visualization techniques, extensibility and scalability. Finally, some coupling metrics applications were presented (e.g., code smell detection), indicating possible future research directions.

**Keywords:** Software Engineering; Coupling relations; Software metrics

**2010 MSC:** 00-01, 99-00

1. **Introduction**

Software development is a complex task that requires careful planning and a high amount of time and energy [1]. Furthermore, maintainability [2, 3] and reliability [4] are important qualities that software should possess. To assess these properties, software complexity measures (coupling and cohesion) were introduced [5, 6]. As defined by Robbes et al. [7], coupling measures the amount of dependency between entities in a software. Over the years, different coupling measures have been proposed. Starting from structural metrics developed for procedural languages [5], new approaches were introduced to measure different relations in object-oriented environments [8]. Nonetheless, the central importance of these metrics for software engineering encouraged researchers to give birth to even more coupling measures in the attempt to evaluate further connections between software entities [9]. Excluding the already existing structural coupling, three new groups of coupling relations were created: dynamic, se-
mantic, and logical coupling. Dynamic coupling analyzes the run-time relations among different software entities \[10\]. Semantic coupling exploits the semantic relations among different elements in the source code using information retrieval techniques \[11\]. Finally, logical coupling approaches work by finding relations among system parts that are frequently changed together \[12\].

Due to the flourishing of this research field, a vast amount of original coupling measures have been proposed. However, all these different approaches can make it difficult for a software engineer to find the proper coupling relations to test the quality of the software on which he or she is working on. Some coupling relations can be applied only to particular groups of programming languages such as the object-oriented ones. Other metrics reveal themselves useful in specific situations: for example, evolutionary coupling is particularly helpful to highlight software changes. For these reasons, this work aims to provide a taxonomy of the coupling relations proposed so far, categorizing them in different groups and highlighting the commonalities and differences among them. Special attention has been given to the various trends that emerged in this field so far, highlighting the motivations behind the construction of new coupling relations. We argue that this study constitutes a good overview of software coupling relations and a starting point for further research in this field. Furthermore, we compare different tools developed by researchers to extract these relations in terms of output and required input information. The goal of this second part of our literature review is complementary to the first one. We argue that a researcher, having identified the coupling metrics of interest, may also be interested in which tools he/she may use to extract them. To the best of our knowledge, this constitutes a new contribution to the existing literature.

Other systematic reviews on coupling relations have been done by Kirbas et al. \[12\] and Nicolaescu et al. \[13\]. However, they have a different aim: Kirbas et al. limit their review to the field of logical coupling, while Nicolaescu et al. organize it in chronological order. The work of Kirbas et al. uses a measurement theory perspective to analyze logical coupling measures. This approach is reflected in the research questions identified in the study: questions like "Do
existing studies use a sound empirical relation system?” or “Do existing studies define measurement methods and procedures?” show the authors’ focus on evaluating how well logical coupling is currently captured by the different measures proposed. However, such an approach is not easily applicable to the broader perspective of our review. We consider logical coupling measures only as a subgroup of all the proposed coupling ones. Our review’s goal is to give a general classification of all the possible coupling measures introduced in the software engineering field and not to analyze in details the properties of a specific subgroup. For this reason, the focus on logical coupling is limited to an overview of the different measures introduced to assess it.

Nicolaescu et al. propose an analysis of the main trends of coupling metrics for object-oriented systems, considering both the fundamental research done in the field and new directions that have been explored in recent years. Although their work constitutes an extraordinary attempt to present an overview of this complex research area, they report the different proposed coupling metrics in chronological order instead of dividing them into groups. In fact, Nicolaescu et al. analyzed 26 research papers dividing them in three time periods: fundamental works (1990-1999), advanced approaches (2000-2010) and recent directions (2011-2015). On the contrary, in our work the main focus is to give a conceptual subdivision of the different coupling relations. In fact, our main concern is not the period in which the considered metrics have been proposed (although, if possible, we keep a chronological order for clarity), but the different rationales behind them, which gave birth to their classification.

2. Research questions

Coupling relations have fundamental importance in software development since they are useful in activities such as, among others, maintenance and program comprehension [7]. For this reason, researchers explored links between software entities in the attempt to capture different characteristics of software to ensure its quality [9]. Nonetheless, a systematic classification of these dif-
different techniques is still missing. Therefore, in this work we first answer the following research question:

**RQ1** Which different coupling relations have been proposed by the software engineering research community?

The goal of **RQ1** is to produce a taxonomy of the existing coupling relations, describing their core points together with the novelty that they introduce. Furthermore, the differences between them will be presented. Then, we will investigate different metric tools based on the relations found in **RQ1** with their outcome and input information. For this reason, our next research question is the following:

**RQ2** Which tools to extract coupling metrics have been developed by the software engineering research community?

**RQ2** aims to retrieve the tools that the software engineering research community has developed to extract different coupling relations. We will classify them based on the taxonomy produced by answering **RQ1**. Moreover, their different inputs and outputs will be highlighted, together with their limitations: e.g., the programming languages to which they are restricted.

3. Research strategy

In our investigation, we followed the guidelines given by Kitchenham [14]. Figure 1 shows the steps of our research strategy. To address **RQ1**, we conducted an initial query to evaluate the goodness of our approach. Based on the papers retrieved, in particular, the work by Bavota et al. [9], we refined our query including terms specific for each coupling group. Moreover, we checked for alternative spellings and synonyms. The terms identified were:

- Structural coupling
- Dynamic coupling
Figure 1: Methodology steps

- (Semantic or Conceptual) coupling
- (Logical or evolutionary or change) coupling

Finally, to investigate our first research question (RQ1) we combined using boolean operators all terms identified to create the following search string:

\[(\text{Software AND coupling}) \ OR \ (\text{coupling AND object-oriented}) \ OR \ \\
(\text{software AND coupling AND ((logical OR evolutionary OR change)} \ OR \ (\text{semantic OR conceptual}) OR \ dynamic OR structural)\]

We included the word “software” to reduce the number of results from research fields other than software engineering. However, this was not necessary when we used terms proper to the computer science area such as “object-oriented”. The same procedure was applied to develop a search string for our second research question (RQ2).

\[(\text{Coupling AND tool}) \ AND \ (\text{metrics OR (logical OR evolutionary OR change}) \ OR \ (\text{semantic OR conceptual}) OR \ dynamic OR structural)\]

These two research strings were used to investigate the following resources:
Table 1 shows the size of the papers set retrieved at each step of our investigation. *ScienceDirect* and *SpringerLink* returned a number of results too vast for an accurate analysis (more than 200′000 results). To restrict this set, we filtered the journals to the ones on computer science and software/software engineering and then we applied our queries to each of them (complete list available in Appendix A). In *SpringerLink*, we excluded the “preview-only” content.

<table>
<thead>
<tr>
<th>Source</th>
<th>Total results retrieved</th>
<th>Initial selection</th>
<th>Final selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEEE Xplore Digital Library</td>
<td>16391 + 1639</td>
<td>69</td>
<td>65</td>
</tr>
<tr>
<td>ACM Digital Library</td>
<td>2838 + 1120</td>
<td>+13</td>
<td>+13</td>
</tr>
<tr>
<td>ScienceDirect</td>
<td>2887 + 2014</td>
<td>+9</td>
<td>+8</td>
</tr>
<tr>
<td>SpringerLink Digital Library</td>
<td>3733 + 2967</td>
<td>+2</td>
<td>+2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>93</strong></td>
<td></td>
<td><strong>88</strong></td>
</tr>
</tbody>
</table>

The retrieved papers were evaluated based on a set of exclusion and inclusion criteria. The exclusion criteria were:

- Articles that do not focus on software coupling relations and/or tools.
- Articles that were not written in English.
- Articles whose full text is not available.

---

1IEEE, [http://ieeexplore.ieee.org/Xplore/home.jsp](http://ieeexplore.ieee.org/Xplore/home.jsp)
2ACM, [https://dl.acm.org/](https://dl.acm.org/)
3SpringerLink, [https://link.springer.com](https://link.springer.com)
4ScienceDirect, [http://www.sciencedirect.com](http://www.sciencedirect.com)
5We reported the number of results obtained with the first query and the number of results obtained with the second one. We did not compute their sum since the two queries presented overlapping results.
• Not peer-reviewed articles, e.g., Ph.D. or M.Sc. thesis

To assess the quality of the retrieved research material, in other words, if the papers identified by our queries contained information useful to answer our research questions (respecting the first criterion), the following three-step procedure was applied. In the first step, the papers’ titles and abstracts were carefully read to exclude the ones clearly irrelevant to the focus of our research. The second step consisted in skimming the whole text of the material left after the first selection to assure that it contained information related to coupling relations, measures and/or tools to extract them. Finally, the third step was an accurate reading through the whole text to ensure that this information was effectively helpful to address the two research questions: for the tools, we checked that their input, output and limitations were described.

At the same time, we applied the following inclusion criteria:

• Year of publication: only papers published between 2002 and 2017 were accepted;

OR

• Number of citations: only papers referenced by more than 100 other publications were accepted;

The two criteria do not have to be valid simultaneously: they are connected with the logical operator OR. Therefore, a paper is selected if it meets at least one of the two criteria. The first criterion was selected to include recent publications on the topic. We included in our work only papers published in the last 15 years at the moment on which this research is conducted: from January 2002 to December 2017. However, we argue that this criterion may have excluded fundamental papers on coupling. Although they have been published before 2002, their contribution could be fundamental to obtain insights on trends and characteristics of more recent metrics. For this reason, we introduced the second criterion to augment the first one. We selected 100 citations as threshold because we were interested in the analysis of solid and well-established resources on the topic of
interest. The number of citations has been attested using Google Scholar to have a verification system independent from the single data source. However, we recognize the potential limits of this approach: a paper with a high number of citations is not necessarily an important paper on the subject. To mitigate this problem, we complemented this research strategy with a snowballing technique [15]. We applied forward and backward snowballing on all the papers included in our final selection set until saturation was reached. The purpose of using the snowballing technique was to compensate for fundamental material that may have been left out by our previously mentioned search queries.

<table>
<thead>
<tr>
<th>Final set</th>
<th>Snowballing</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>88</td>
<td>48</td>
<td>136</td>
</tr>
</tbody>
</table>

As a further check on the goodness of the retrieved material and its correct use, we reached authors of other similar works on coupling to ask them to check if their papers were cited correctly and suggest us any other published work relevant for our review. 17 authors were reached and 6 replied. As result, we discovered one paper previously not included and we added it to the set of primary studies.

4. RQ1: Coupling relations

In this section different existing coupling relations and techniques will be presented, based on the result obtained investigating our first research question.

As stated by Briand et al. “Coupling refers to the degree on interdependence among the components of a software system” [16]. A component can be a module of the system or a smaller entity such as a class or an object. Moreover, coupling can indicate a relation between two components but also a property of an entity compared to all the other related entities in the system: e.g., CCBC (Conceptual Coupling Between Classes) and CoCC (Conceptual Coupling of a Class) [17]. Bavota et al. identified four different measures of coupling [9]:

\[
\text{CCBC} = \text{CoCC} = \text{CoCM} = \text{CoCA}
\]
Structural coupling exploits the static relations in the source code: it focuses on finding patterns such as called methods, relations among classes (inheritance and friendship) and accessed variables. Dynamic coupling also reflects calls between classes and methods but it does that at run-time, instead of looking at the static code. Semantic coupling relies on Information Retrieval techniques to find relations in the code lexicon, while logical coupling intends to assess the entities that are frequently changed together, and therefore share a link, using historical information. Finally, other approaches try to combine these groups of relations in a complementary way or present coupling measures for domain-specific programming languages. Figure 2 shows an overview of the coupling relation taxonomy that we have constructed in our review. The goal of this first part of our work is to present the evolution of the coupling relations and metrics proposed, while keeping intact the categories presented by Bavota et al. [9].

4.1. Structural coupling

Structural coupling relations exploit static connections among software entities (modules, objects or classes). Measures to assess them have been initially developed for procedural languages, but, later, extended to the object-oriented paradigm. Furthermore, some structural coupling relations have been proposed specifically for object-oriented languages. In general, it is possible to divide them into two broad groups: procedural programming coupling measures and object-oriented coupling measures [18].

Myers divided the coupling for procedural programming languages in 6 different levels, reported in Table 3 ordered from the worst to the best in terms...
Figure 2: Coupling relation taxonomy
of consequences on the maintainability and quality of the software [19]. These coupling levels have been extended by Offutt et al. to include global and bidirectional coupling measures, previously left uncovered [5]. Although it has been originally introduced for procedural languages, this subdivision remains valid also for object-oriented ones.

In 1981, Henry and Kafura proposed an information-flow technique to construct different measures for a software system [21]. Their idea constitutes an interesting approach to compute coupling relations. In particular, the information flow metrics can recognize content coupling and common coupling. The authors argue that content coupling is equivalent to their direct local flows. Common coupling is equivalent to the global flow measure. Henry and Kafura developed two metrics fan-in and fan-out. Fan-in counts the number of local flows to the considered procedure together with the number of data structures read by the procedure. Fan-out measures the quantity of local flows from a selected procedure plus the number of data structures on which the procedure writes. At a later time, Allen et al. proposed to measure coupling metrics using the links and information obtainable building the system graph of a software [22, 23]. The strength of this approach is that it can be applied to many software design abstractions and to all kinds of programming languages (procedural and object-oriented).

The flourishing of the object-oriented paradigm brought the researchers to propose metrics and relations to cover its new aspects. Coupling relations for object-oriented systems have been investigated in the work done by Eder et al. in 1994. The authors identified three groups of coupling relations [24]: interaction, component and inheritance coupling. Their classification is shown in Table 4.

A fundamental structural coupling metric for object-oriented software is CBO (Coupling Between Object) [25], which belongs to the interaction coupling subgroup [24]. CBO reflects the degree to which an object acts upon another, excluding the parent-child relation. It constitutes one of the core coupling metrics and it has been further refined and applied in different other domains. Chidamber and Kemerer also defined RFC (Response For a Class), a
Table 3: Structural coupling relations for procedural languages

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content coupling</td>
<td>Refers to two modules of which one directly accesses the contents of the other: e.g., module A modifies a statement or branches to a local label of module B.</td>
</tr>
<tr>
<td>Common coupling</td>
<td>Happens when two modules have access to the same global data: for example, both modules can read and write the same global record. Schach points out that if the access to the data is read-only, then this can not be considered common coupling [20].</td>
</tr>
<tr>
<td>Control coupling</td>
<td>Happens when a module explicitly controls the logic of another. However, this does not happen if the first module passes only data.</td>
</tr>
<tr>
<td>External coupling</td>
<td>Happens when two modules exchange information using an external element such as a file.</td>
</tr>
<tr>
<td>Stamp coupling</td>
<td>Exists between two modules if one of them passes a data structure as an argument to the second one, but the called module does not operate on all the components of the data structure.</td>
</tr>
<tr>
<td>Data coupling</td>
<td>Exists among two modules if the arguments that they pass to each other are all homogeneous data items: simple arguments or data structures in which all elements are used by the calling module [20].</td>
</tr>
<tr>
<td>Level</td>
<td>Description</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Interaction coupling</td>
<td>Defined as the invocations among different methods and their sharing of variables</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Component coupling</td>
<td>A class $A$ is related to another class $B$ if and only if $A$ is referenced in $B$: this happens when $A$ is the domain of an instance or local variable, a method’s parameter or a parameter of a method called inside a method of $B$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Inheritance coupling</td>
<td>Relates two classes if one of them is a subclass of the other</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
coupling measure related to CBO \cite{26, 27} that measures the total communication potential. These two metrics were further analyzed by Briand et al. \cite{28} and formalized to remove possible sources of ambiguities. Moreover, Briand et al. introduced CBO’ to include the ancestor classes in the metric computation, previously left excluded \cite{25}. Other important metrics that they considered were Message Passing Coupling (MPC) and Data Abstraction Coupling (DAC), originally defined by Li and Henry \cite{29}. Briand et al. further refined DAC in DAC’, a metric that counts the number of classes used as types of attributes. Finally, other important structural coupling measures are effenter and afferent coupling (C_e and C_a \cite{30}), Coupling Factor (COF) \cite{31, 32} and Information-flow-based coupling (ICP) \cite{33}. Li defined two new coupling metrics to complement Chidamber and Kemerer’s metrics suite: CTA (Coupling Through Abstract data type) and CTM (Coupling Through Message passing) \cite{34}. Similarly to the DAC’ \cite{28} metric, CTA measures the relation between two classes created when one of them uses the other in its data declaration. CTM (Coupling Through Message passing) measures the number of messages sent by a considered class to the other classes in the system, excluding the ones sent to objects used locally by the methods of the class. MPC, RFC and CBO were also modified to be applied to program slices \cite{35}, creating the new metrics SMPC, SRFC and SCBO. The method-level metrics proposed by Briand et al. have been adapted to a finer granularity by English et al. to distinguish the different types of constructs with which they might be related \cite{36}. A specific focus has been given to the friendship relation.

To take into account indirect coupling relations and the strength of coupling between two classes, Li developed a new metric \cite{37}. Indirect coupling has also been considered by Yang et al. \cite{38, 39} (creating also a tool, Indirect Coupling Detector) and later by Almugrin et al. \cite{40}.

New measures have been introduced to allow an evaluation of the level of object-orientation of a program to estimate the possibility that an object-oriented fault happens. For this purpose, Tang et al. proposed new coupling metrics \cite{41}: IC (Inheritance Coupling) and CBM (Coupling Between Meth-
ods). Gui and Scott focused instead on metrics specific for component reusability [12], defining measures for the direct coupling among two classes (CoupD), the transitive coupling among two classes (CoupT) and the total coupling of a software system (WTCoup).

An interesting approach is the one proposed by Aloysius and Arockiam, where a comprehensive coupling metric, CWCBO (Cognitive Weighted Coupling Between Objects), is defined to give an overall measure of different degrees of coupling [43]. This metric considers different kinds of coupling measures such as data coupling, control coupling, global coupling and interface coupling and applies to them a weighting factor. Using a comparative study, the authors supported their claim that CWCBO is a better indicator than CBO to measure the complexity of a class since it takes into consideration different coupling levels.

4.2. Dynamic coupling

Dynamic coupling rules were introduced to address problems left not completely answered by previous static coupling measures: e.g., dynamic binding and polymorphism [10]. In fact, these metrics lose precision when dealing with intensive use of inheritance and dynamic binding. Furthermore, they aim to evaluate software quality looking not only at the source code complexity, but also at its operational environment [44]. Further research confirmed that these metrics provide additional information to the structural metrics [45]. Dynamic coupling approaches can be further divided according to the coupling direction, import or export coupling, and their mapping level, object or class-level oriented [46]. Coupling direction captures the difference between a sending entity and a receiving entity [46] [10]. Considering the messages exchanged between entities, the distinction is:

- **Import coupling**: focus on the messages sent from an entity
- **Export coupling**: focus on the messages received by an entity
The mapping level reflects the applicability domain of dynamic coupling rules: object-level or class-level coupling. Moreover, Arisholm et al. proposed three different approaches to evaluate the strength of a coupling relation: the number of messages, the number of distinct method invocations and the number of distinct classes. The first one refers to the quantity of distinct messages exchanged between two entities. The other two represent the amount of methods called and classes used, respectively, by a method in an object. The classification of these metrics, as given by Arisholm [46], is summarized in Table 5.

<table>
<thead>
<tr>
<th>Direction</th>
<th>Mapping</th>
<th>Strength</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Import Coupling</td>
<td>Object-level</td>
<td>Dynamic messages</td>
<td>IC_OD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distinct Methods</td>
<td>IC_OM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distinct Classes</td>
<td>IC_OC</td>
</tr>
<tr>
<td></td>
<td>Class-level</td>
<td>Dynamic messages</td>
<td>IC_CD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distinct Methods</td>
<td>IC_CM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distinct Classes</td>
<td>IC_CC</td>
</tr>
<tr>
<td>Export Coupling</td>
<td>Object-level</td>
<td>Dynamic messages</td>
<td>EC_OD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distinct Methods</td>
<td>EC_OM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distinct Classes</td>
<td>EC_OC</td>
</tr>
<tr>
<td></td>
<td>Class-level</td>
<td>Dynamic messages</td>
<td>EC_CD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distinct Methods</td>
<td>EC_CM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distinct Classes</td>
<td>EC_CC</td>
</tr>
</tbody>
</table>

The direction of coupling was taken into account also by Mitchell and Power; their idea was to expand the previously defined CBO metric [44]. The authors presented two new coupling relations, both divided to account for internal and external coupling. The first one, Run-time coupling between objects ($R$) (external or internal) counts the number of accesses made by/to a class divided by the total number of accesses. The second one, run-time import (or export) degree of coupling ($RD$), gives the strength of the coupling relation computed...
as the number of classes that access (or are accessed by) a selected class. In further research, [Mitchell and Power] in their attempt to verify if CBO can be used efficaciously as a dynamic metric, defined two new coupling measures [47]: Run-time Coupling Between Objects (RCBO) counts the amount of classes that a specific class accesses at run-time. The second one, the Number of object class clusters (Noc), counts the number of clusters obtained considering a class in the system and studying the distribution of unique accesses per object.

Work at object-level was also done by [Yacoub et al. 48]. The authors proposed two dynamic coupling metrics that operate on the export and import side, respectively. The former one, Export Object Coupling (EOC) measures the percentage of messages sent from an object to the other, compared with the total amount of messages exchanged in the complete execution. Import Object Coupling (IOC) works in the opposite way, reflecting the number of messages that an object receives from another. From EOC, [Yacoub et al.] defined OQFS (Object reQuest For Service) as the sum of the EOC among a selected object and all the other objects in the design. IOC was instead developed into OPFS (Object resPonse For Service), defined as the sum of IOC between the given object and all the other objects in the application during the execution of a specific scenario. These values affect the maintainability, understandability, reusability and the errors distribution in the code. [Zaidman and Demeyer] refined OQFS to work at class-level, introducing CQFS (Class reQuest For Service) [49]. CQFS counts all the methods that a given class calls during the application execution. Every method is counted once: calling the same method more than one time does not increase the count.

[Hassoun et al.] propose a general relation, called DCM (Dynamic Coupling Metric) to formalize the idea of dynamic coupling [50]. Their metric works at object-level and it can be used to analyze the coupling of a selected object or a system during the program execution.

Dynamic coupling metrics require analysis conducted at run-time, but the impact of the metrics is higher if they are computed at early stages of the development. To address this issue, pseudo-dynamic coupling metrics were pro-
posed: static metrics that consider the expected usage profile (derived from UML graphs during the design phase) [51]. Referring to Chidamber and Kemerer’s CBO, the pseudo-dynamic CBO metric was defined as the value of the static CBO multiplied by the value of the operational profile. It presents a strong correlation with the dynamic CBO metric. A similar static approach to compute dynamic metrics has been proposed by Liu and Milanova [52]. A different approach calculates dynamic metrics from the system use case maps and the interactions between different scenarios [53].

An interesting evolution of the metrics defined by Arisholm [46] has been introduced by Abualese et al. to evaluate the importance of a class in the understanding process that a developer has to face when approaching code written by a different person [54].

4.3. Semantic coupling

Classes can not only be structurally related to each other but also conceptually. Based on this idea, semantic coupling uses information from comments and identifiers to identify relations among software entities [11]. The technique proposed by Poshyvanyk and Marcus relies on Latent Semantic Indexing (LSI) [55]: a machine learning model developed to analyze relations between words and documents. To investigate coupling aspects left unaddressed by the previous metrics, the authors created four progressive coupling relations, each of them based on the previous one: CCM (Conceptual Coupling Between Methods), CCMC (Conceptual Coupling Between a Method and a Class), CCBC (Conceptual Coupling Between two Classes), also called CSBC (Conceptual Similarity Between two Classes) and CoCC (Conceptual Coupling of a Class) [11, 17]. Poshyvanyk and Marcus also considered the idea to exclude weak coupling links in the computation of the metrics defining a new metric called CSMC\textsubscript{m}. From it, they also recomputed CSBC and CoCC accordingly, producing the two new metrics CSBC\textsubscript{m} and CoCC\textsubscript{m}. Újházi et al. have further improved this approach with a new metric called CCBO (Conceptual Coupling Between Object classes), which does not merely take the maximum but identifies a threshold to distin-
guish between a strong and weak semantic coupling [56].

All the aforementioned semantic coupling measures use LSI to create an initial semantic corpus for the analysis. Gethers and Poshivyanyak proposed a coupling approach based on a different technique: Relational Topic Model (RTM), a model that can find connections between documents based on the context [57]. The authors introduced a measure called Relational Topic-based Coupling (RTC). This metric individuates new aspects of coupling between classes compared to the metrics based on LSI, such as CoCC. Furthermore, a fundamental benefit of this model is that it does not need any previous knowledge about the links between classes.

Revelle et al. extended semantic coupling relations to work at feature level, aiming to identify which parts of a system are linked to a specific function [59, 58]. In fact, a feature represents the implementation of a functionality described in the requirements. Since features can be represented by structured information (source code and related artifacts) and unstructured information (textual information), two different metrics were proposed: SFC (Structural Feature Coupling) and TFC (Textual Feature Coupling). Furthermore, the authors introduced HFC (Hybrid Feature Coupling) to consider together the complementary information captured by SFC and TFC.

Semantic coupling has been combined with evolutionary coupling [60] or domain-based relations [61]. Domain-based coupling individuates relations among domain variables, functions and User Interface Components (UIC) [62, 63] and has been applied to fields such as code clone detection with promising results [64]. Gethers et al. defined CSE (Conceptual Similarity between Entities) and CSED (Conceptual Similarity between two UICs) to perform impact analysis in hybrid software systems [61]. Moreover, semantic and domain-based coupling relations have been checked to assure their orthogonality. This analysis confirmed that these relations capture different aspects of the analyzed system and, therefore, they can be efficaciously combined. Based on CBE, Kagdi et al. defined CSEMC and CSEBC [65]. Furthermore, semantic coupling has been combined with structural coupling to create a metric that takes into account
both aspects at the same time \cite{66}. The authors defined four coupling metrics, each of them based on the previous one (in a fashion similar to the one used by Poshyvanyk et al. \cite{17}): MPC (Method Pair Coupling), HCMC (Hybrid Coupling between Method and a Class), HCCC (Hybrid Coupling between two classes) and SSCM (Coupling of a class in an object-oriented system). Moreover, they positively performed an evaluation to confirm that these metrics identify aspects not covered by structural and semantic coupling relations alone.

4.4. Logical coupling

Logical coupling (sometimes also called evolutionary or change coupling) works by finding similar change patterns in the release history: it aims to investigate “the sequential dependencies such as if module A is changed in one release, module B is changed in the next release” \cite{67}. This approach has been further developed to be applied at class level in the research conducted by Gall et al. \cite{68}, with the aim to identify classes that share a common change behavior. The authors proposed a distinction between internal and external links: internal coupling happens between classes in the same module, while external coupling involves classes contained in different modules. Their approach works using data extracted from the CVS (Concurrent Versions System) release history. Further research focused on a finer-grained analysis of system evolution, compared to the description obtained using CVS. Robbes et al. \cite{7} argued that this method is imprecise because it employs the commits as basic analysis blocks. For this reason, they defined coupling metrics to work using information collected during software development through a tool that saves all the changes made to a system in development together with the exact time at which they were made. Alali et al. proposed to further extend these metrics analyzing the contribution of age and pattern distance measures \cite{69}. Age is defined as the period of time between the appearance of the specific evolutionary coupling relation and its disappearance. Pattern distance represents the tree distance between two files in a program. Another interesting approach is the one proposed by D’Ambros et al. \cite{70}. They introduced two weighted change coupling measures EWSOC
(Exponentially Weighted Sum of Coupling) and LWSOC (Linearly Weighted Sum of Coupling). They both emphasize recent changes over past ones, but the latter penalizes them less than the former.

4.5. Recent or isolated trends

Outside this classification, other coupling relations have been proposed. They are novel techniques, still too recent to be considered a proper subgroup of metrics, or relations developed for specific domains.

A first novel relation is interaction coupling. Interaction coupling aims to group artifacts that are likely to implement the same task. Zou et al. \cite{71} worked on the task interaction history, defining the strength of the relation between two entities as the number of times they are accessed together. Although interaction coupling may be considered similar to logical coupling, the former requires information from the task interaction histories and involves not only artifacts that are changed together, but also entities that are viewed in the same portion of time. Interaction coupling and logical coupling have been combined by Bantelay et al. to predict future interactions \cite{72}.

The usefulness of general coupling relations led the researchers to tailor them to domain-specific applications: e.g., knowledge-based systems. Kramer and Kaindl proposed the Degree of Coupling of Frame (DCpF) metric to measure the number of slots in a frame connected to other slots in different frames by a rule \cite{73}. Coupling measures have also been developed for Web Ontology Language to evaluate the complexity of the system \cite{74}. Table \ref{table6} summarizes the metrics proposed for this application. Furthermore, coupling metrics have been modified to be applied to Agent-oriented software development. Jordan and Collier proposed a reformulation of the CBO metric as coupling between abstractions, defining Coupling Between Elements (CBE) \cite{75}: two elements are coupled if any direct dependencies exist between any of their parts. If an element accesses or modifies the implementation details of another one, this leads to a dependency.
<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEC (Number of external classes)</td>
<td>Number of other classes outside the selected ontology</td>
<td>[74]</td>
</tr>
<tr>
<td>REC (References to External Classes)</td>
<td>Counts the number of references to external classes considering a selected ontology</td>
<td>[74]</td>
</tr>
<tr>
<td>RI (Referenced Includes)</td>
<td>Computes the number of includes used in an ontology</td>
<td>[74]</td>
</tr>
<tr>
<td>CBE-in (Coupling Between Entities)</td>
<td>Considers the class to be in the property domain</td>
<td>[76, 77]</td>
</tr>
<tr>
<td>CBE-out (Coupling Between Entities)</td>
<td>Considers the class to be in the property range</td>
<td>[76, 77]</td>
</tr>
<tr>
<td>SC (Self-Coupling)</td>
<td>Captures the properties with a class contemporary in the property range and domain</td>
<td>[76, 77]</td>
</tr>
</tbody>
</table>
Coupling relations have also been adapted for Aspect-Oriented (AO) software, where the basic entities are aspects and classes, to assess its reusability and maintainability [78]. Sant’Anna et al. [78] and Ceccato and Tonella [79] (further improved by Shen and Zhao [80] with the addition of other seven metrics) proposed an initial set of metrics to measure aspect-oriented coupling relations. An example is CBC (Coupling Between Components) [78], a general measure of coupling that accounts for different relations between classes and aspects in AO programs. However, these metrics have been criticized by the software engineering community for not taking into account finer dimensions of class-aspect coupling and their lack of empirical validation. Moreover, their adoption was disregarded by the software developers [81]. For this reason, Burrows et al. aimed to assess the quality of these metrics and, if necessary, how they might be improved. The authors defined a new AOP coupling measure called Base-Aspect Coupling (BAC) that quantifies the strength of the link between the base and the aspect code. In the same year, Bernardi and Lucca proposed a further set of coupling metrics based on their classification of aspects’ interactions [82]. They proposed a metric for coupling due to interactions altering the static structure (CLSS), to interactions altering the control flow (CLCF) and to interactions altering the state of an object (CLSO). Finally, they considered all these interactions together in the metric CLA (Coupling Level of an Aspect).

New metrics were proposed by Bennett and Mitropoulos to address the problem of aspect interference [83]: an aspect that causes unexpected changes to the flow of a class or a method. The authors argued that previously proposed AO coupling metrics did not cover all the interaction necessary to describe potential aspect interference. To solve this issue, two new metrics were introduced: IP (Interference Potential) and ICP (Interference Causality Potential). Furthermore, these two metrics have been combined in a new one, TIP (Total Interference Potential) [84]. At the same time, attempts have been made to create a comprehensive framework, independent from the language considered, to define AO coupling measures [85] [86] [87].

Moreover, specific coupling metrics have also been developed for Service-
Oriented Architectures (SOA) [88, 89, 90, 91, 92, 93, 94, 95]. A SOA is an architectural model to combine different services in one platform. It can be formed by a combination of technologies, products, APIs and various other components and is not related to a particular programming language [96]. Among the metrics proposed, we report ASSD (Average Service State Dependency), ASPD (Average Service Persistent Dependency), ARSD (Average Required Service Dependency) [88], SOCI (Service Operational Coupling Index), ISCI (Inter-Service Coupling Index) [89] and ASOU (Average Service Operation Coupling) [93]. ASSD and ASPD compute the average of the services’ states and persistent state dependencies, respectively. A persistent state dependency happens between services that share a state, which all of them can use and update. Finally, ARSD measures the average number of services to which each service in the system provides its functionalities. SOCI (Service Operational Coupling Index) and ISCI (Inter-Service Coupling Index) measure the dependence of a service on other services and on messages, respectively. The former was adapted by the object-oriented metric RFC, the latter from the CBO metric. ISCI can be considered as the opposite of ARSD [88]. To measure the dependency based on messages, a new metric was formulated: SMCI (Service Message Coupling Index). Even if it may seem to have a dynamic nature, it is computed statically from the information model of the domain. Finally, ASOU computes the coupling of a service as the sum of its synchronous and asynchronous invocations divided by the total number of services in the domain. Karhikeyan and Geetha identified five types of dependencies that influence the coupling of a Service-Oriented system: direct, indirect, state, IO and delayed message dependency [94]. They developed a metric for each of them and proposed a fuzzy model to evaluate the overall coupling of a system. Pautasso and Wilde proposed a framework to evaluate coupling in Service-Oriented design, with a specific focus on Web technologies [97]. They defined 12 coupling directions (or facets) to assess different design aspects, giving an instrument for comparing different Web services options.

The discussed coupling metrics for SOA are all static. Based on the promising results obtained by dynamic coupling in object-oriented systems, Quynh and
Thang introduced a set of dynamic metrics for Service-Oriented systems [91]: CBS (Coupling Between Services), which has been derived from CBO, IMS (Instability Metric for Service), DC2S (Degree of Coupling between 2 Services) and DCSS (Degree of Coupling within a given Set of Services).

Semantic coupling relations have also been further developed to deal with Service-Oriented Architecture (SOA). New metrics needed to be created since the ones proposed by Poshyvanyk et al. [17] could not be applied in this domain: comments and identifier names are not accessible for services and, furthermore, the required concepts can also be obtained using business level artifacts [98]. For these reasons, Kazemi et al. developed three coupling metrics [98]: CCO (Conceptual Coupling between Operations), CDSO (Conceptual Dependency of a Service to an Operation) and CCS (Conceptual Coupling of a Service).

An interesting application of coupling measures is to assess the information security of object-oriented designs [99]. To this purpose, a new metric CCC (Critical Classes Coupling) has been defined. CCC computes the degree of interconnection among classes and classified attributes in a given software design. Moreover, it is based on design graphs (e.g., UML). However, to extract security information these graphs need to be annotated using tools such as UMLsec [100] or SPARK’s annotations [101].

Finally, coupling measures have been defined for Computational Science and Engineering (CSE) applications [102] and real-time application design [103]. In the context of real-time application design, Ahmed and Shoaib defined a set of metrics (e.g., MEF, Message Exchange Factor) to evaluate the system in its early development phases [103]. Kamble et al. investigated coupling in Computational Science and Engineering (CSE) software [102] to perform software integration. They claimed that this domain is different from others due to the complex algorithms and functions involved.
5. RQ2: Developed tools

Different tools have been proposed to extract the measures discussed in Section 4. Based on the previous classification, the aim is to identify how they work, the information that they require and their advantages and disadvantages. Table 7 shows a summary of the tools that we have considered divided based on the kind of metrics that they extract (structural, dynamic, semantic, logical coupling or relations belonging to smaller groups). For each tool its input and output are reported, together with its limitations: mainly its restriction to a particular programming language (or set of languages).

Some tools are stand-alone kits that simply extract a set of metrics: e.g., CCMETRICS [105]. However, during our analysis, two trends emerged clearly: the use of visualization techniques to improve users’ understanding and the focus on extensibility. Moreover, in tools developed for dynamic coupling metrics, due to the significant amount of data that needs to be analyzed, researchers focused on scalability to improve the tools’ performance.

5.1. Extensibility

A problem of metric tools proposed by the software engineering community is that the majority of them can not be extended to support new metrics or languages [133]. For this reason, some metrics tools focused on extensibility with the specific intent to support future metrics. Examples can be found in QScope [106], which provides an explicit mechanism to include new metrics and a framework to develop and test them, OOMeter [112] and AMT [107]. This last tool takes a further step towards extensibility, being expandable not only with new metrics but with new languages too. To achieve an implementation independent from the programming language analyzed it takes as input a representation of the source code using XML. However, this representation should be created from the source code using a parser. For this reason, at the time of the publication of their research, Kayarvizhy and Kanmani’s tool AMT
<table>
<thead>
<tr>
<th>Coupling</th>
<th>Tool</th>
<th>Input</th>
<th>Output</th>
<th>Limitations</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structural coupling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>dJava</td>
<td>Java files or JAR files</td>
<td>CBO (Coupling Between Objects)</td>
<td>Restricted to Java applications</td>
<td>[104]</td>
</tr>
<tr>
<td></td>
<td>CVMETRICS</td>
<td>Source code</td>
<td>DAC (Data Abstraction Coupling)</td>
<td>Restricted to object-oriented languages</td>
<td>[98]</td>
</tr>
<tr>
<td></td>
<td>Qscope</td>
<td>XML database of the program representation</td>
<td>CBO and RFC</td>
<td>and their graphical visualization</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AMT</td>
<td>Source code</td>
<td>CBO, CBO′, RFC, MPC, DAC and DAC′</td>
<td>Restricted to Java and C#</td>
<td>[107]</td>
</tr>
<tr>
<td></td>
<td>WebMetrics</td>
<td>Source code</td>
<td>CBO and RFC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Fan-in and Fan-out</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DependencyViewers</td>
<td>Java files or JAR archives</td>
<td>C′ (afferent coupling) and C′′ (efferent coupling)</td>
<td>Limited to Java applications</td>
<td>[111]</td>
</tr>
<tr>
<td></td>
<td>OOMeter</td>
<td>Source code (Java or C#) or UML diagrams (in XML)</td>
<td>CBO (Coupling Between Objects)</td>
<td>Only supports UML in XML format</td>
<td>[123]</td>
</tr>
<tr>
<td></td>
<td>CLUSTERCHANGES</td>
<td>CodeFlow changes</td>
<td>Clusters of diff-regions (visualized a tree graph)</td>
<td>Restricted to CFlow</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SOPA</td>
<td>UMLsec or SPARK graphs (generated with the tool)</td>
<td>CCC</td>
<td>Specific for assessing security (using UMLsec or SPARK’s annotations)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OSN tool</td>
<td>UML diagrams (in XMI format)</td>
<td>DAC, DAC′</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ABA tool</td>
<td>Java source code</td>
<td>SOC, DIT</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>JACT</td>
<td>Java source files</td>
<td>MPC</td>
<td>Restricted to Java</td>
<td>[118]</td>
</tr>
<tr>
<td></td>
<td>JCAT</td>
<td>Java source files</td>
<td>GCC, ICC, ICC′, DCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>JACTiva</td>
<td>Java source files</td>
<td>CBO, CTI</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Indirect Coupling</td>
<td>Eclipse IDE</td>
<td>nav-def indirect coupling</td>
<td>Eclipse plugin</td>
<td></td>
</tr>
<tr>
<td>Dynamic coupling</td>
<td>JDissect</td>
<td>Running Java program</td>
<td>Dynamic coupling measures</td>
<td>Restricted to Java</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SSS tool</td>
<td>Running Java program</td>
<td>Total Dynamic Messages (TDM)</td>
<td>Restricted to Java applications</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Distinct ClassCouple (DCC)</td>
<td>Restricted to Java applications</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Distinct MethodCouple (DMC)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DMC</td>
<td>2nd files</td>
<td>EC_CU, EC_CM and EC_CCD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>IC_CU, IC_CM and IC_CCD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semantic coupling</td>
<td>IC3MI</td>
<td>Source code</td>
<td>CIC, CIC′</td>
<td>Restricted to C/C++ programs</td>
<td>[125]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CIC′′ and CIC′′′</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FLAT</td>
<td>Source code</td>
<td>TFC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Feature-Method Mapping</td>
<td>Eclipse plugin</td>
<td></td>
</tr>
<tr>
<td>Logical coupling</td>
<td>ROSE</td>
<td>CVS data</td>
<td>Locations for further changes</td>
<td>Warnings about probable missing changes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exhibition Radar</td>
<td>CVS data</td>
<td>Graphical visualization of coupling between modulos and files</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>YHDC tool</td>
<td>CVS data</td>
<td>List of files changed together with the selected one</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other couplings</td>
<td>OWL-VisMod</td>
<td>Ontologies code</td>
<td>Graphical visualization of CBE-in and CBE-out relations</td>
<td>Restricted to OWL language</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AJATO</td>
<td>Source code and Concern maps (XML) and Buks (XML)</td>
<td>CBC (Coupling Between Components)</td>
<td>Restricted to Aspect-Orientated applications</td>
<td>[103]</td>
</tr>
<tr>
<td></td>
<td>AlfMetrix</td>
<td>AspectJ files and Java files</td>
<td>CAM, CAA, CAF, and CAF′ (CBE-in and CBE-out relations)</td>
<td>Restricted to Aspect-Oriented software</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CT tool</td>
<td>AspectJ source code</td>
<td>CAF, CAF′, CCM′, CFA′</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SSP tool</td>
<td>UML diagrams</td>
<td>SOCC, RSCT and SMCT</td>
<td>Restricted to MAS</td>
<td></td>
</tr>
</tbody>
</table>
only supports Java and C#. A similar approach has also been implemented in WebMetrics \cite{webmetrics}. The tool implements an architecture that includes an intermediate level of abstraction between the code and the metrics computation: the code is processed to extract a list of relations, which are analyzed in a second step to compute the desired metrics. This allows an easier implementation of new measures in the tool since the developer does not need to know how the parser operates, but only the generated intermediate relations.

5.2. Visualization techniques

Applying visualization techniques to metric tools constitutes another important trend in research. The goal is not only to extract a set of software metrics but to support and improve the users’ understanding. A step in this direction has been made by DependencyViewer \cite{dependencyviewer} and OOMeter \cite{oometer} in 2005. Both of them can show the metrics extracted using simple graphs: e.g., DependencyViewer reports the metrics computed for a package as a column graph.

In the field of logical coupling metrics, Evolution Radar \cite{evolutionradar} (2006) and the tool proposed by Hanakawa \cite{hanakawa} (2007) implemented a visualization technique. D’Ambros et al. \cite{dambros} argue that visualization techniques give immediateness to the user. Evolution Radar shows as output the coupling links existing between a selected module and the other system’s modules. The visualization interface uses the distance between the center (where the selected module is located) as a measure of the strength of the coupling relation: the closer a module is to the center, the stronger is the link. Furthermore, due to the interactivity of this approach, it is possible to see more information related to the selected entity such as the author, timestamp, comments, lines added and removed and its source code and the logical coupling among entities over time. Hanakawa’s visualization tool presents two maps: a module coupling map and a logical coupling one. Both of them can be shown at the same time.

\footnote{When the tool’s name is not explicitly stated in the referenced research, we will call it with the initial letters of the authors’ surnames}
JCTIViz [119] (2008) uses a polymetric view to display software metrics. Each class or interface is represented with a node: the dimensions of the node (height, width and depth) represents the value of a metric. In particular, the node depth represents the CBO value. A different approach considers the creation of tools as plug-ins for existing IDEs. EPOSpix [134] exploits this idea showing related classes in Eclipse with a pixel map. eROSE [135] is an Eclipse plug-in that computes logical coupling to suggest related changes to the developer.

Visualization techniques have been applied also by Garcia et al. to coupling relations among OWL ontologies [76, 129]. Their tool, OWL-VisMod, requires as input the ontology’s code and it shows the coupling CBE-in/out relations among the classes. Classes are displayed using a radial layout, where the selected class occupies the center. On the right and left side are displayed the classes coupled by a CBE-out or a CBE-in relation, respectively. An edge link couples classes and its color indicates the direction of the coupling relation.

5.3. Scalability and Dynamic coupling

Different ways exist to collect dynamic metrics: using run-time information or relying on simulating the execution behavior of a system using interaction diagrams, such as UML or Real-time Object Oriented Modeling (ROOM) language [136].

A first tool to find dynamic relations, proposed by Arisholm, is JDissect [46]. The tool works in two phases: in the first one, it gathers information from a running program, while in the second step the collected data are analyzed. However, its first limitation consists in its restriction to Java applications, due to its connection with the JVM (Java Virtual Machine). In fact, this tool uses the JVM interfaces to collect dynamic information. For what concerns the input required, JDissect needs a running Java program to extract the dynamic coupling relations in it. Another tool to extract dynamic coupling metrics is DynaMetrics, proposed by Singh and Singh [122]. It can compute both dynamic and static metrics, analyzing the data collected at run-time (specifically, event log files).
Extracting a significant amount of data from the execution of a program may require a vast amount of time and resources. To mitigate this problem, in 2015 Sarvari et al. [120] proposed to parallelize this process using Hadoop MapReduce. Hadoop MapReduce needs the XML file of the program to be executed. For this reason, the authors utilized JP2 [137]: an open source tool that creates CTT XML files from a running Java program. Furthermore, Hadoop MapReduce can be used both locally and on the cloud: a cloud-based approach further helps in dealing with large quantities of data. For this reason, in 2017 Dogra et al. [121] proposed DMA (Dynamic Metric Analysis), a tool based on Platform as a Service (PaaS). Like Sarvari et al.’s tool, DMA relies on JP2 but adapts it to be streamed to the cloud. In this way, it allows the user to have a real-time analysis of the coupling metrics during the program execution. Sarvari et al.’s tool returns three dynamic coupling measures: TDM (Total Dynamic Messages), DCC (Distinct Class Coupling) and DMC (Distinct Method Couples). The authors introduced this nomenclature for the first time in the software engineering research field. However, these metrics are the same as the ones developed earlier by Arisholm et al. [10]: TDM corresponds to IC_CD (or EC_CD, depending on the considered direction of the relation), DCC is equal to IC_CC (or EC_CC) and DMC is the same as IC_CM (or EC_CM).

Another approach is to collect dynamic coupling data from the UML diagrams of the program [46]. On the one hand, since these diagrams are usually done in the early design phase, the main advantage of this approach resides in the possibility of using dynamic relations to take design decisions. On the other hand, the coupling measures collected tend to be underestimated due to the impossibility to distinguish the different messages in the set of possible messages in the system using UML. Unfortunately, our systematic review did not find any examples of tools that implemented this approach to extract dynamic metrics. Tools such as OOMeter [112] and the tool proposed by Girgis et al. [115] extract coupling metrics from design diagrams, but they are restricted to struc-
### Table 8: Coupling Metrics Summary

<table>
<thead>
<tr>
<th>Coupling group</th>
<th>Metric</th>
<th>Metric Ref</th>
<th>Tool(s)</th>
<th>Tool ref</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structural coupling</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>fan-in, fan-out</td>
<td></td>
<td>WebMetrics</td>
<td>[21]</td>
</tr>
<tr>
<td></td>
<td>CBO</td>
<td></td>
<td>ckjm, QScope, AMT, WebMetrics</td>
<td>[8]</td>
</tr>
<tr>
<td></td>
<td>RFC</td>
<td></td>
<td>ckjm, QScope, AMT, Web Metrics</td>
<td>[8]</td>
</tr>
<tr>
<td></td>
<td>CBO′</td>
<td></td>
<td>AMT</td>
<td>[28]</td>
</tr>
<tr>
<td></td>
<td>MPC, DAC</td>
<td></td>
<td>CCMETRICS, AMT</td>
<td>[29]</td>
</tr>
<tr>
<td></td>
<td>DAC′</td>
<td></td>
<td>AMT</td>
<td>[28]</td>
</tr>
<tr>
<td></td>
<td>SSIPC, SRFC, SCBO</td>
<td></td>
<td>no tool</td>
<td>[35]</td>
</tr>
<tr>
<td></td>
<td>C_a</td>
<td></td>
<td>DependencyViewer, DynaMetrics</td>
<td>[30]</td>
</tr>
<tr>
<td></td>
<td>C_b</td>
<td></td>
<td>DependencyViewer, DynaMetrics</td>
<td>[30]</td>
</tr>
<tr>
<td></td>
<td>COF</td>
<td></td>
<td>no tool</td>
<td>[32]</td>
</tr>
<tr>
<td></td>
<td>ICP</td>
<td></td>
<td>no tool</td>
<td>[33]</td>
</tr>
<tr>
<td></td>
<td>NAS</td>
<td></td>
<td>no tool</td>
<td>[34]</td>
</tr>
<tr>
<td></td>
<td>CTA, CTM</td>
<td></td>
<td>no tool</td>
<td>[35]</td>
</tr>
<tr>
<td></td>
<td>IC, CBM</td>
<td></td>
<td>no tool</td>
<td>[36]</td>
</tr>
<tr>
<td></td>
<td>CWCBO</td>
<td></td>
<td>no tool</td>
<td>[37]</td>
</tr>
<tr>
<td></td>
<td>CCC</td>
<td></td>
<td>SCPA</td>
<td>[38]</td>
</tr>
<tr>
<td></td>
<td>RMC, RIC</td>
<td></td>
<td>no tool</td>
<td>[39]</td>
</tr>
<tr>
<td></td>
<td>MPEIC, MPIC, AFM</td>
<td></td>
<td>no tool</td>
<td>[40]</td>
</tr>
<tr>
<td></td>
<td>PLC</td>
<td></td>
<td>no tool</td>
<td>[41]</td>
</tr>
<tr>
<td></td>
<td>CRC, CNIC, MC, CC, AMC</td>
<td></td>
<td>no tool</td>
<td>[42]</td>
</tr>
<tr>
<td></td>
<td>CoupD, CoupT, WTCoup</td>
<td></td>
<td>no tool</td>
<td>[43]</td>
</tr>
<tr>
<td><strong>Dynamic coupling</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IC and IC′</td>
<td></td>
<td>DMA, JDissect</td>
<td>[44]</td>
</tr>
<tr>
<td></td>
<td>R, RD</td>
<td></td>
<td>no tool</td>
<td>[45]</td>
</tr>
<tr>
<td></td>
<td>RCBO</td>
<td></td>
<td>DynaMetrics</td>
<td>[46]</td>
</tr>
<tr>
<td></td>
<td>Noc</td>
<td></td>
<td>no tool</td>
<td>[47]</td>
</tr>
<tr>
<td></td>
<td>EOC, IOC, OQFS, OPFS</td>
<td></td>
<td>DynaMetrics</td>
<td>[48]</td>
</tr>
<tr>
<td></td>
<td>CQFS</td>
<td></td>
<td>DynaMetrics</td>
<td>[49]</td>
</tr>
<tr>
<td></td>
<td>DCM</td>
<td></td>
<td>no tool</td>
<td>[50]</td>
</tr>
<tr>
<td></td>
<td>TBD, DCC, DMC</td>
<td></td>
<td>SSS tool</td>
<td>[51]</td>
</tr>
<tr>
<td></td>
<td>EUC, EREC</td>
<td></td>
<td>no tool</td>
<td>[52]</td>
</tr>
<tr>
<td></td>
<td>ICV</td>
<td></td>
<td>no tool</td>
<td>[53]</td>
</tr>
<tr>
<td><strong>Semantic coupling</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CCM, CC5MC, CS5MC</td>
<td></td>
<td>no tool</td>
<td>[54]</td>
</tr>
<tr>
<td></td>
<td>CSBC (CCBC), C5CC</td>
<td></td>
<td>BCC5M</td>
<td>[55]</td>
</tr>
<tr>
<td></td>
<td>CSBCn (CCBCn), C5CCn</td>
<td></td>
<td>BCC5M</td>
<td>[56]</td>
</tr>
<tr>
<td></td>
<td>CCBO</td>
<td></td>
<td>no tool</td>
<td>[57]</td>
</tr>
<tr>
<td></td>
<td>FTC</td>
<td></td>
<td>no tool</td>
<td>[58]</td>
</tr>
<tr>
<td></td>
<td>SFC, TPC</td>
<td></td>
<td>FLAT</td>
<td>[59]</td>
</tr>
<tr>
<td></td>
<td>SFC, HFC</td>
<td></td>
<td>no tool</td>
<td>[60]</td>
</tr>
<tr>
<td></td>
<td>CSE, CS5ED</td>
<td></td>
<td>no tool</td>
<td>[61]</td>
</tr>
<tr>
<td></td>
<td>CS5EM, CS5EB</td>
<td></td>
<td>no tool</td>
<td>[62]</td>
</tr>
<tr>
<td><strong>Logical coupling</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LC</td>
<td></td>
<td>ROSE, Evolution Radar, YINNC tool and Hamakawa’s tool</td>
<td>[63]</td>
</tr>
<tr>
<td></td>
<td>CC, TC</td>
<td></td>
<td>no tool</td>
<td>[64]</td>
</tr>
<tr>
<td></td>
<td>IC</td>
<td></td>
<td>no tool</td>
<td>[65]</td>
</tr>
<tr>
<td></td>
<td>NOCC, SOC, EWSOC, LWSOC</td>
<td></td>
<td>no tool</td>
<td>[66]</td>
</tr>
<tr>
<td><strong>Semantic + structural coupling</strong></td>
<td>MPC, HC5MC, BCC5C, SSC5M</td>
<td></td>
<td>no tool</td>
<td>[67]</td>
</tr>
</tbody>
</table>
Table 9: Other Coupling Metrics Summary

<table>
<thead>
<tr>
<th>Applicability field</th>
<th>Metric</th>
<th>Tool(s)</th>
<th>Metric Ref</th>
<th>Tool ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge-based Systems</td>
<td>DCpF</td>
<td>no tool</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Web Ontology Language</td>
<td>NEC, REC, RI</td>
<td>no tool</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CBE-in, CBE-out</td>
<td>OWL-VisMod</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SC, DC, CBE-in, CBE-out</td>
<td>no tool</td>
<td></td>
<td></td>
</tr>
<tr>
<td>aspect-oriented Software</td>
<td>CBE</td>
<td>AIKINE</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CFA, CBD, CME, CBE</td>
<td>no tool</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CAM, CAA, CAF, RPM, BRF</td>
<td>AIKINE</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BAC</td>
<td>no tool</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IP, ICP</td>
<td>no tool</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TIP</td>
<td>no tool</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CoAT, CoPT, CoAR, CoOI, CoI, CoHA</td>
<td>no tool</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CLAS, CLCF, CLIO, CLA</td>
<td>no tool</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agent-Oriented Paradigm</td>
<td>CBE</td>
<td>no tool</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Service-Oriented Architecture</td>
<td>ASO, ASD, ASED</td>
<td>no tool</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SOCI, SSI, SSCI</td>
<td>SSI tool</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WSCIE, WEISCE, WESCIC, WEISCI, WESDOC, WESRC, WEISRC, WESURF, WESURF</td>
<td>no tool</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ASO1</td>
<td>no tool</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CBS, IMS, DCIS, DCNS</td>
<td>no tool</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CVO, CDVO, CCS</td>
<td>no tool</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DDI, DSI, DSID, DSID, DSID, DSI, DILogD</td>
<td>no tool</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remote-component-based Systems</td>
<td>CCSP</td>
<td>no tool</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real-time application design</td>
<td>MEF</td>
<td>no tool</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. Discussion

A vast quantity of coupling metrics and relations has been proposed for different paradigms and applications, starting with procedural languages and, later, object-oriented ones. Due to their importance in assessing the software quality and analyzing programs’ features, coupling relations were proposed to investigate aspects left uncovered by previous research and to be applied to specific application domains.

Our research showed how CBO (Coupling Between Objects), proposed by Chidamber and Kemerer as part of their metrics suite, became a fundamental coupling metric used as base for further metrics and refinements by other researchers: examples can be found in CBO’ [28], CWCBO [43] and CBE (Coupling Between Elements) [75]. Moreover, our investigation revealed that sometimes the researchers encountered difficulties in retrieving previously proposed metrics. Analyzing the material collected in our review, we noticed inco-

---

7 at different granularity levels
sistencies in the metrics names: e.g., afferent and efferent coupling have been proposed as $C_a$ and $C_e$ by Martin [30], but later referred by Singh and Singh [122] as AFC and EC. This is only a formal issue, but different nomenclatures for referring to the same metric may undermine the coherence of the research corpus in this field. The problem of formally defining the metrics and validate them led to the creation of many frameworks: e.g., the one defined by Tempero and Ralph [146].

Software metrics should undergo a theoretical and empirical validation when they are introduced. Our investigation revealed that coupling metrics are evaluated referring to the properties defined by Kitchenham et al. [147], Weyuker [148] and Briand et al. [28]. Metrics such as CTA and CTM [34], CWCBO [43] and CCBC [11] have been validated using this process. However, we noticed that a vast number of metrics have been proposed without undergoing a theoretical evaluation: for instance, based on properties like Representation condition [147]. Many studies performed only an empirical evaluation. Using a set of test cases, the goal of the studies was to assess that the newly proposed metric achieves better performance than a previous one as an indicator for a specific application: e.g., fault prediction. Moreover, a common trend is to apply correlation analysis techniques (Spearman correlation or Principal Component Analysis) to verify the orthogonality of a new metric compared to previously presented ones. More emphasis has been given by the researchers on this second aspect of the evaluation. The theoretical evaluation does not seem to be considered as fundamental as the empirical one since the latter contributes to highlight the novelty of the metric. Thus, we suggest novel metrics to employ both a theoretical and empirical validation/correlation analysis.

Table 8 shows a summary of the metrics retrieved in our systematic literature review. They are grouped based on the category of coupling relations to which they belong. Furthermore, they are associated with the tools that can be used to extract them (if any). While for the structural, dynamic and semantic coupling relations a set of metrics has been defined, for the logical coupling relations no strict metric definitions seem to exist. In the table, we referred to the
classification given by Robbes et al. [7], but their definitions allow different interpretations of the same metric. Further efforts should be devoted to provide a consistent formal definition of logical coupling metrics. Table 9 contains an overview of the metrics belonging to the Other coupling approaches group. They are grouped based on the field of applicability (e.g., Aspect-Oriented software). As in the previous table, the tools that can be used to extract them (if retrieved in our systematic review) are reported.

Our analysis of the coupling metric tools proposed by the researchers revealed two interesting trends: the progressive use of visualization techniques as a means to show the information to the user and the focus on making easily extendible tools. Visualization techniques, used in tools such as Evolution Radar [125] or OWL-VisMod [129], help the user to have a better understanding of the considered software properties. Usually, this approach allows changing the considered entity interactively. D’Ambros et al. stated that the idea of recurring to visualization is based on the following motives: “it provides effective ways to break down the complexity of information” and “it has been shown to be a successful means to study the evolution of software systems” [125]. As the second trend, easily expandable tools want to overcome the problem of having metric tools that work only on a specific programming language (or groups of languages). Researchers proposed modular designs in which new metrics can be implemented without the need to understand the whole tool implementation. Examples can be found in tools such as AMT [107] and WebMetrics [108].

Tahir and MacDonell stated that dynamic metrics could be collected using a run-time analysis or executable modules and interaction diagrams (UML or ROOM) [136]. Although both of these approaches have been analyzed in the literature, in our review we did not find any tool that implemented a methodology based on interaction graphs. This could be caused by the lack of precision that dynamic coupling metrics computed during the design phase are likely to have, which may have discouraged further research attempts in this direction. However, it is also necessary to highlight that this may be caused by the limited scope of our review, as given by our procedure and especially the choice to
restrict the analysis to academic-developed tools. Also the semantic coupling area suffers from a lack of tools to extract its correlated metrics: IRC²M [17] and FLAT³ [58] are the only ones retrieved in our systematic review. This can be explained by the fact that semantic coupling relations have been investigated only by a restricted group of researchers.

Coupling relations can be used to cluster related code changes, helping developers in the process of reviewing and modifying their software. Logical coupling is particularly suited for this task, due to its intrinsic nature: logical coupling relations were introduced to find similar change patterns in the code release history [67]. An example can be found in ROSE [123], which gives suggestions to the user regarding which portions of code are likely to have to be changed with the current one. However, also structural or semantic coupling relations can be effectively used with this intent. CLUSTERCHANGES [113] uses data coupling to cluster code diff-regions that influence each other and, therefore, should be inspected together when modifying one of them. On the contrary, we argue that dynamic coupling metrics are unsuitable for this task since they reflect run-time relations among software elements which can not be easily collected when dealing with code changes. An interesting way to approach the problem of grouping related code changes is given by the evolutionary coupling relations proposed by Zou et al. [71]. Information on which entities have been accessed together during the development phase may constitute a sound basis on which grouping together portions of code: in fact, these are likely to implement the same functionality.

7. Coupling Relations: A Research Roadmap

While the research community heavily investigated ways to measure coupling relations, we believe that future research directions should and will be devoted to the application of such coupling metrics as well as the definition of effective combinations of metrics that would allow a better estimation of the actual coupling of software classes. This section aims at reporting a (non-exhaustive)
roadmap for further research in the field.

**Applications.** There are plenty of opportunities to use coupling metrics to support other software maintenance and evolution tasks. For instance, their use in the context of code review may represent an effective method to improve the way developers detect defective source code. Specifically, change-based code review constitutes an important trend in software development and improving the existing techniques may lead to a significant contribution to software engineering [149, 150, 151]. Coupling relations may be applied to analyze the code contained in different changes and, consequently, cluster similar changes together. Baum et al. proposed an ordering theory for code changes based on the relations that they share with each other [149]. In particular, they conducted a survey among developers to evaluate which relations were considered important. Among all of them, they mentioned the similarity relation. We argue that logical and semantic coupling relations may be applied as practical implementation of this relation. Still in the context of code review, coupling metrics might be exploited in conjunction with just-in-time defect prediction [152]: we envision the introduction of coupling-related information on top of the recommendations provided by defect prediction models, so that developers might be informed on the classes having relations with a defective file and possibly assess the risk of defect propagation over these classes.

Another promising research field in Software Engineering is Code Smell detection [153, 154, 155]. Recent works started to exploit it by using machine learning techniques [156, 157] and to classify the severity of a code smell issue [158]. While some structural and logical coupling metrics have already been used as features of these models, there is still room for improvement: as shown by our survey, the role of many complementary coupling metrics can be explored to improve the (not always good [159]) performance of currently available code smell prediction models. Still in the same area, the application of conceptual coupling metrics have been explored by Palomba et al. [160, 161]. The authors also suggested that the exploration of a combination between structural and conceptual metrics may lead to promising results. This is something that is still
unknown and that might lead to new research directions on how to combine
the output of different metrics. At the same time, it remains unclear what is
the value of other coupling metrics in the context of code smell detection: for
example, to the best of our knowledge, implications of using dynamic coupling
metrics to detect code smells are still to be evaluated. This seems to be a natu-
ral fit for the identification of Message Chain instances [162], given its intrinsic
dynamic nature: in fact, it occurs when a long chain of method invocations is
required for the operations of a class [163].

Finally, coupling relations have found a major field of application in Change
Prediction, a research area dealing with identifying the classes that are more
prone to be modified in the future [164]. Most works rely on the use of struc-
tural coupling metrics (among others) as indicators of these classes [165]. A
recent study conducted by Elish and Al-Rahman Al-Khiaty evaluates a set of
evolution metrics for change prediction purposes [166]. The authors reported
that these metrics measure different dimensions than the classical Chidamber
and Kemerer’s metrics suite [25] and that their combination improved the ac-
curacy of their prediction model. Based on the promising results of their work,
we argue that the application of logical coupling or conceptual coupling metrics
to this context may be worthy. This metric may be combined with structural
or dynamic ones and tested to see if the performance of a model that takes into
account these different aspects increases: we expect so from the moment that
recent studies [164, 167] showed how an improved description of the change pre-
diction phenomenon, done through the addition of orthogonal information, can
dramatically increase the overall ability of prediction models in discriminating
the classes that are more likely to change in the future.

Coupling relations and metrics have been applied in many different contexts,
of which the ones cited above (e.g., code review, code smells detection or change
prediction) constitute just a small part. Depending on the application consid-
ered, combining two or more groups of coupling metrics may be worthwhile: the
existing techniques could increase their performance. An example may be found
in the research conducted by Palomba et al. to identify code smells with concep-
tual coupling metrics, where the authors argue that the possible combination of these metrics with others belonging to different groups (e.g., structural or conceptual) may lead to a further performance increase [160].

Combination. During our investigation, we noticed very few attempts to integrate previously proposed coupling metrics in an ensemble metric, i.e., a method able to combine the information coming from different sources. In our opinion, this represents an important research direction that might be worth to investigate to come up with more powerful solutions for measuring coupling relations. As an example, consider the application of machine learning approaches in this context: coupling metrics computed using different data sources (e.g., structural vs conceptual coupling) might be nicely adopted as features of a regressor able to estimate a combined form of coupling that may provide developers with a comprehensive view of the phenomenon, thus facilitating her ability to take informed decisions. At the same time, we envision a combination of those metrics through the use of search-based algorithms: a clear opportunity is represented by the possibility to apply them for refactoring purposes (e.g., to improve software re-modularisation by means of aggregate measures that optimize the locations of classes).

Furthermore, our work classifies new approaches or coupling metrics for specific domains in a generic group called “recent or isolated trends” (section 4.5). The knowledge on those metrics is still poor and the way they can effectively complement existing measures is still unknown. This represents an opportunity for future research, as researchers are called to investigate further how these emerging metrics can be combined with the set of metrics for which way more information is available.

8. Conclusion

This work presented a systematic review of the coupling relations and metrics proposed until now by the software engineering research community. In the first part of our research, we analyzed the trends that emerged over time in
the software coupling area. We developed a taxonomy, as complete as possible within the limitations of our approach, of these relations in the attempt to give a systematic classification of over thirty years of research in the field. Based on previous works, such as the one done by Bavota et al. [9], we divided the coupling relations into four main groups: structural, dynamic, semantic and logical. Furthermore, we included a fifth group of coupling metrics not listed with the previous ones, since they constitute new trends still in development or coupling metrics developed for a particular field of applicability (such as knowledge-based systems or aspect-oriented applications).

In the second part of our investigation, we presented the tools developed by the research community to extract (and sometimes even visualize) coupling relations. The tools retrieved have been summarized in Table 7 maintaining the structure used to answer our first research question: dividing the tools based on the coupling group of metrics that they extract. For each tool, we highlighted the input that it needs and the output that it produces, together with its possible limitations: e.g., a restriction to a particular programming language. Moreover, we analyzed three main trends noticed in the academic-proposed tools: application of visualization techniques, extensibility, and scalability (applied to dynamic coupling metrics). We proposed a discussion on our findings and a roadmap for future work. The complexity of this research field sometimes led to discrepancies among the introduced coupling metrics. As guidance for future work we highlighted interesting applications of the presented coupling relations and metrics (change clustering, code review, code smells detection and change prediction), reporting the groups of coupling metrics already applied in these fields together with the ones that are yet to be explored and that may constitute the starting point for future work. Furthermore, we discussed the possibility to combine existing coupling metrics to create ensemble metrics, able to combine information from different sources.
Acknowledgments

E. Fregnan, F. Palomba, and A. Bacchelli gratefully acknowledge the support of the Swiss National Science Foundation through the SNF Project No. PP00P2_170529.

List of primary studies

- N. Fenton, J. Bieman, Software Metrics: A rigorous and practical approach, CRC Press, 2014
- S. Chidamber, C. Kemerer, Towards a metrics suite for object oriented design, Conference Proceedings on Object-oriented programming systems, languages, and applications (OOPSLA ’91) (1991) 197–211.


- G. Myers, Reliable software through composite design, Mason and Lipscomb, 1975.


- Y. Liu, A. Milanova, Static Analysis for Dynamic Coupling Measures, in: Proceedings of the 2006 Conference of the Center for Advanced Studies on Collaborative Research, CASCON '06, IBM Corp., Riverton, NJ, USA,


- P. Rosner, S. Viswanathan, Visualization of Coupling and Programming to Interface for Object-Oriented Systems, in: 2008 12th International Conference
Information Visualisation, ISSN 1550-6037, 575–581, 1335


Testing, Verification and Validation Workshops, 566–571, 2011.


References


[15] C. Wohlin, Guidelines for Snowballing In Systematic Literature Studies and a Replication In Software Engineering, Proceedings of the 18th Inter-


[73] S. Kramer, H. Kaindl, Coupling and cohesion metrics for knowledge-based


23rd Asia-Pacific Software Engineering Conference (APSEC), ISSN 1530-

[147] B. Kitchenham, S. L. Pfleeger, N. Fenton, Towards a framework for soft-
ware measurement validation, IEEE Transactions on Software Engineering

URL https://doi.org/10.1109/32.6178.

[149] T. Baum, K. Schneider, A. Bacchelli, On The Optimal Order Of Read-
ing Source Code Changes for Review, IEEE International Conference on

[150] T. Baum, O. Liskin, K. Niklas, K. Schneider, A Faceted Classification
Scheme for Change-Based Industrial Code Review Processes, 2016 IEEE
International Conference on Software Quality, Reliability and Security

[151] L. Pascarella, D. Spadini, F. Palomba, M. Bruntink, A. Bacchelli, Informa-
on Human-Computer Interaction - CSCW 2018 .

[152] Y. Kamei, E. Shihab, B. Adams, A. E. Hassan, A. Mockus, A. Sinha,
N. Ubayashi, A large-scale empirical study of just-in-time quality assurance,

D. Poshyvanyk, When and why your code starts to smell bad (and whether
the smells go away), IEEE Transactions on Software Engineering 43 (11)


Appendix A. List of selected ScienceDirect Journals

- AASRI Procedia
- Advances in Engineering Software
- Astronomy and Computing
- Computer Fraud and Security
- Computer Languages
- Computer Languages, Systems and Structures
- Computer Methods and Programs in Biomedicine
- Computer Programs in Biomedicine
- Computer Standards and Interfaces
- Data Processing
- Digital Investigation
- Egyptian Journal of Basic and Applied Sciences
- Entertainment Computing
- Environmental Modelling and Software
- Environmental Software
- Euromicro Newsletter
- Future Computing and Informatics Journal
- Future Generation Computer Systems
- Information and Software Technology
- Information Systems
- Integration
• Intelligent Data Analysis

1920 • Journal of Computational Science
• Journal of Innovation in Digital Ecosystems
• The Journal of Logic and Algebraic Programming
• The Journal of Logic Programming
• Journal of Logical and Algebraic Methods in Programming

1925 • Journal of Parallel and Distributed Computing
• Journal of Systems Architecture
• Journal of Systems and Software
• Journal of Web Semantics
• Microprocessing and Microprogramming

1930 • Microprocessors
• Microprocessors and Microsystems
• Network Security
• Performance Evaluation
• Robotics

1935 • Science of Computer Programming
• SoftwareX