Philipps-Universität Marburg
Fachbereich Mathematik und Informatik
AG Softwaretechnik

Analysis of Conflicts and Dependencies between User Stories using Graph Transformation

Masterarbeit
zur Erlangung des akademischen Grades
Master of Science

vorgelegt von
Amir Rabieyan Nejad
Matrikel-Nr: 3350269

29. April 2024

Selbstständigkeitserklärung

Hiermit versichere ich, Amir Rabieyan Nejad, dass ich die vorliegende Arbeit mit dem Titel *Analysis of Conflicts and Dependencies between User Stories using Graph Transformation* selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Die Masterarbeit wurde in der jetzigen oder ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen anderen Prüfungszwecken gedient.

Marburg, den 29. April 2024

Abstract

User stories, the basic building blocks of software development, serve as precise and testable descriptions of a software's functionality. Within the dynamic framework of *agile development*, these user stories are usually written informally in simple text and managed in the product backlog, which serves as a repository for prioritising and tracking development tasks.

Behaviour Driven Development (BDD), a special approach in the field of agile software development, places a strong emphasis on the iterative implementation of user stories. The sequence of user stories in BDD is a central aspect of the methodology. The correct sequence not only has an impact on the efficiency of development, but also on the overall success of the project. By effectively prioritising and sequencing user stories, development teams can deliver incremental value to users, respond to changing requirements and ensure that the most critical features are addressed first.

User stories often have dependencies on each other, leading to potential conflicts if one user story requires the deletion of a component that is essential for the successful execution of another user story, or if one user story introduces an element that runs counter to the realisation of another user story and thus prevents it. In addition, changing an existing requirement or adding a new requirement to the existing *product backlog* in agile software development can also cause conflicts due to changes in the needs and concerns of *system stakeholders*. This can result in a wide range of inconsistencies due to conflicting requirements, as requirements are raised by multiple stakeholders involved in product development to achieve different functions.

Normally, to minimise the occurrence of conflicts, teams should systematically identify and document dependencies between user stories within their backlog. Agile methods such as *Scrum* encourage cross-functional collaboration and daily stand-up meetings as mechanisms to address and mitigate dependencies and conflicts in a timely manner. However, this approach can be time and resource consuming. In cases where the backlog is very extensive, recognising existing conflicts between user stories can become a complex undertaking.

Model-based software engineering is a suitable method for coping with the ever-increasing complexity of software development processes. *Graphs* and *graph transformations* have proven to be useful for visualising such models and their changes.

Since there is no process in agile software development to systematically identify and manage conflicts between requirements created in natural language, this thesis aim to presents a well-structured workflow that uses a collection of techniques and tools from different domains to accelerate the automatic detection of conflicts and dependencies in natural language user stories.

Zusammenfassung

User Stories, die Grundbausteine der Softwareentwicklung, dienen als präzise und testbare Beschreibungen der Funktionalität einer Software. Innerhalb des dynamischen

Rahmens der *agilen Entwicklung* werden diese User Stories in der Regel informell in einfachem Text verfasst und im *Product Backlog* verwaltet, das als Repository für die Priorisierung und Verfolgung von Entwicklungsaufgaben dient.

Behaviour Driven Development (BDD), ein spezieller Ansatz im Bereich der agilen Softwareentwicklung, legt einen starken Schwerpunkt auf die iterative Umsetzung von User Stories. Die Reihenfolge der User Stories im BDD ist ein zentraler Aspekt der Methodik. Die richtige Reihenfolge hat nicht nur Auswirkungen auf die Effizienz der Entwicklung, sondern auch auf den Gesamterfolg des Projekts. Durch eine effektive Priorisierung und Abfolge der User Stories können die Entwicklungsteams den Benutzern einen inkrementellen Mehrwert bieten, auf sich ändernde Anforderungen reagieren und sicherstellen, dass die kritischsten Funktionen zuerst behandelt werden.

User Stories sind oft voneinander abhängig, was zu potenziellen Konflikten führen könnte, wenn eine User Story die Löschung einer Komponente erfordert, die für die erfolgreiche Ausführung einer anderen User Story unerlässlich ist, oder wenn eine User Story ein Element einführt, das der Realisierung einer anderen User Story zuwiderläuft und diese somit verhindert. Darüber hinaus kann auch die Änderung einer bestehenden Anforderung oder das Hinzufügen einer neuen Anforderung zum bestehenden Product Backlog in der agilen Softwareentwicklung zu Konflikten führen, da sich die Bedürfnisse und Anliegen der Systemstakeholder ändern. Dies kann zu einer Vielzahl von Inkonsistenzen aufgrund widersprüchlicher Anforderungen führen, da Anforderungen von mehreren an der Produktentwicklung beteiligten Stakeholdern gestellt werden, um unterschiedliche Funktionen zu erreichen.

Um das Auftreten von Konflikten zu minimieren, sollten Teams in der Regel systematisch Abhängigkeiten zwischen User Stories innerhalb ihres Backlogs identifizieren und dokumentieren. Agile Methoden wie *Scrum* fördern die funktionsübergreifende Zusammenarbeit und tägliche Stand-up-Meetings als Mechanismen, um Abhängigkeiten und Konflikte zeitnah anzugehen und zu entschärfen. Dieser Ansatz kann jedoch zeit- und ressourcenaufwendig sein. In Fällen, in denen das Backlog sehr umfangreich ist, kann das Erkennen bestehender Konflikte zwischen User Stories zu einem komplexen Unterfangen werden.

Modellbasiertes Software-Engineering ist eine geeignete Methode, um die ständig steigende Komplexität von Softwareentwicklungsprozessen zu bewältigen. *Graphen* und *Graphtransformationen* haben sich als nützlich erwiesen, um solche Modelle und deren Änderungen zu visualisieren.

Da es in der agilen Softwareentwicklung keinen Prozess zur systematischen Identifizierung und Verwaltung von Konflikten zwischen in natürlicher Sprache erstellten Anforderungen gibt, soll in dieser Arbeit ein gut strukturierter Arbeitsablauf vorgestellt werden, der eine Sammlung von Techniken und Werkzeugen aus verschiedenen Bereichen verwendet, um die automatische Erkennung von Konflikten und Abhängigkeiten in natürlichsprachlichen User Stories zu beschleunigen.

Contents

1 Introduction

User stories (USs) are fundamentally informal units in the field of software development, which should be as precise and testable as possible descriptions of the functionality of a software in the dynamic framework of *agile development.* The iterative implementation of these user stories, especially in the context of *behaviour driven development* (BDD), plays a decisive role in the efficiency and success of a project[30]. As development teams navigate through the *product backlog* to prioritise and sequence the user stories, the inherent interdependencies between them become clear, presenting challenges to efficient implementation and potential conflicts.

Product backlog in agile software development acts as a repository for development tasks, reflecting the evolving needs and concerns of *system stakeholders*[35]. Additionally, the involvement of a *product owner* (PO) in agile development has broken down traditional barriers between development teams and end-users, fostering an environment where the product backlog becomes a central artifact guiding the development process[35]. However, there is no formal language for expressing stories or modelling backlogs from a practical point of view. Managing dependencies between user stories is increasingly important, and although agile methodologies such as *scrum* advocate collaboration and daily stand-up meetings to resolve conflicts, the process can be resource intensive, especially with extensive backlogs.

In between automated support for extracting domain models from requirements artifacts such as USs play a central role in effectively supporting the detection of dependencies and conflicts between user stories. Domain models are a simple way to understand the relationship between artifacts and the whole system. For example, Mosser et al. propose a model engineering method (and the associated tooling) to exploit a graph-based meta-modelling and compositional approach. The objective is to shorten the feedback loop between developers and POs while supporting agile development's iterative and incremental nature[30].

Natural language processing (NLP) is a *computational* method for the automated analysis and representation of human language [8]. NLP techniques offer potential advantages to improve the quality of USs and can be used to parse, extract, or analyze user story's data. It has been widely used to help in the software engineering domain (*e.g.,* managing software requirements [3], extraction of actors and actions in requirement document [1]. Furthermore, the incorporation of computational lexicon resources like *VerbNet*[1] aids in semantic analysis, capturing linguistic and semantic data for a comprehensive understanding.

To systematically identify conflicts and dependencies between USs, the *critical pair analysis*(CPA) extension of Henshin is used to determine whether USs influence each other through model-driven transformation rules. CPA for graph rewriting [15] has been adapted to rule-based model transformation, *e.g.* to find conflicting functional requirements for software systems [16], or to analyses potential causal dependencies between model refactorings [27], which helps to make informed decisions about the most

---

[1]https://verbs.colorado.edu/verbnet

appropriate refactoring to apply next.

This thesis attempts to introduce a well-structured workflow. The overall goal is to develop a framework that accelerates the automatic detection of conflicts and dependencies in USs expressed in natural language using graph transformation.

For this reason, we use the model presented by Mooser et al. 2022 and present an extended model, which is visualised in Figure 1 and illustrates the subsequent phases:

- Initially, Mooser et al. used the only publicly available residue dataset as the dataset[10]. They optimise the backlog and transfer it to the *Conditional Random Fields* (CRF) tool. The CRF tool is used to generate a graph-based model that represents the refined and annotated backlog to recognise *Entities*, *Actions*, *Personas* and *Benefits* of USs[30].

- In the next step, we want to use annotated blacklog datasets generated by CRF tool and utilise the actions generated by CRF that correspond to the verbs in each user story. These actions are then processed through a *computational lexical resource VerbNet*, to categorize them into four distinct categories namely "create", "delete", "forbid" and "preserve".

- Following that, the *Henshin* tool, specifically the transformation module through Henshin's *application programming interface* (API), is employed to create a class model and formulate transformation rules aligned with each user story.

- We then use a Henshin extension known as *Critical Pair Analysis* (CPA) to automatically analyse conflicts and dependencies between user stories programmatically via their API.

- With the outcomes of the CPA analysis, the requirements engineer can visualize conflicts and dependencies between user stories in a text-based format. This presentation is designed to be lightweight and facilitates comprehension for both the development team and the client.

This thesis is structured as follows: Related work is listed in section 2. In Section 3, we present a product backlog containing a set of annotated user stories to evaluate them with the QUS framework and AQUSA. In Section 4, we implement a method to extract all verbs from the product backlog to categorise them into four categories, namely "create", "delete", "forbid" and "preserve". In Section 5, we implement a method to convert annotated user stories into transformation rules and take action on each verb category. In Section 6, we implement a method to automatically analyse conflicts and dependencies between user stories using Henshin's CPA programming interface and conclude with Section 7. [h]

2 Related Work

This section discusses relevant work such the role of USs and backlogs in agile development (Section 2.1) and the most frequently used pattern of USs is presented.
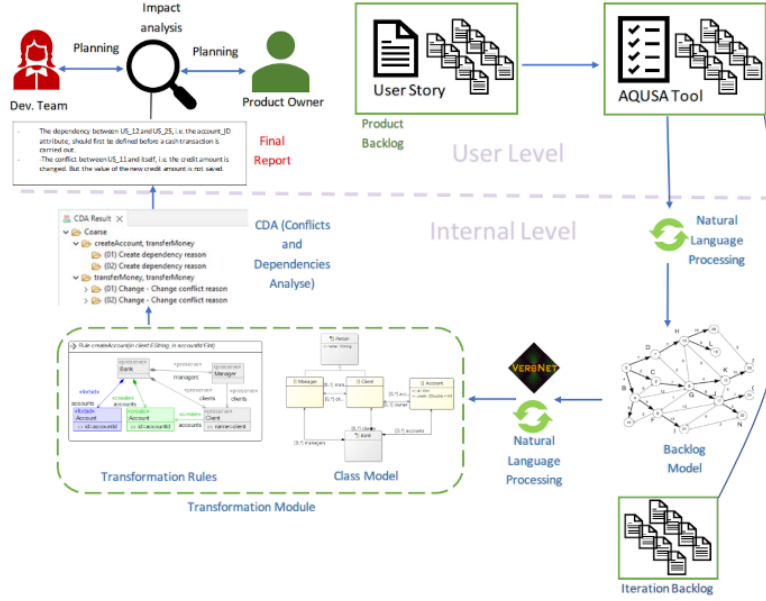
Figure 1: Extending backlog model represented by Mosser et al., 2022[30]

In Section 2.2, the focus is on the *quality user story* (QUS) framework and its tool "AQUSA" to answer the question of the existence of criteria and an automatic way to manage and identify conflicts and dependencies between user stories. Following this investigation, we draw a conclusion as to whether the existing criteria are able to optimise the USs in terms of their conflicts and dependencies.

2.1 Role of User Stories and Backlogs in Agile Development

The agile software development paradigm broke the wall that classically existed between the development team and end-users. Thanks to the involvement of a *product owner* (PO) who acts as a proxy to end-users for the team, the product backlog [35] became a first-class citizen during the product development.

Furthermore, thanks to a set of user stories expressing features to be implemented in the product in order to deliver value to end-users, the development teams were empowered to think in terms of added value when planning their subsequent developments. The product is then developed iteration by iteration.

Sedano et al. posited that a "product backlog is an informal model of the work to be done" [35]. A backlog implements a shared mental model among the practitioners working on a given product, acting as a boundary artefact between stakeholders. This model is voluntarily kept informal to support rapid prototyping and brainstorming sessions. Classically, backlogs are stored in project management systems, such as Jira[2] . These tools store user stories as tickets, where stakeholders write text as natural lan-

---

[2]https://www.atlassian.com/en/software/jira

10

guage. Meta-data (*e.g.*, architecture components, severity, quality attribute) can also be attached to the stories. However, there is no formal language to express stories or model backlogs from a state of practice point of view.

A *user story* (US) is a brief, semi-structured sentence and informal description of some aspect of a software system that illustrates requirements from the user's perspective [33]. Large, vague stories are called epics. While user stories vary widely between organizations, most observed stories included a motivation and acceptance criteria. The brief motivation statement followed the pattern: As a *<user>* I want to *<action>* so that *<value>*. This is sometimes called the *Connextra* template. The acceptance criteria followed the pattern: Given *<context>*, when *<condition>* then *<action>*. This is referred to as *Gherkin syntax*[41]. It consists of three aspects, namely aspects of *who*, *what* and *why*. The aspect of "who" refers to the system user or actor, "what" refers to the actor's desire, and "why" refers to the reason (optional in the user story) [33].

The US components consist of the following elements[39]:

- *role*: abstract behaviour of actors in the system context; the aspect of who represents.

- *goal*: a state or circumstance that is desired by stakeholders or actors

- *task*: specific things that need to be done to achieve goals.

- *Capability*: the ability of actors to achieve goals based on certain conditions and events.

User stories usually have dependencies, i.e. the order in which they are implemented plays a decisive role. This raises the question of how to recognise and identify the relationships between the user stories.

2.2 QUS Framework and AQUSA as a Tool

Lucassen et al. [26] present a quality user story (QUS) framework consisting of 13 quality criteria that US authors should strive for. The criteria analysed determine the intrinsic quality of USs in terms of *syntax*, *pragmatics* and *semantics*. Figure 2 illustrates the structure of the agile requirements verification framework. Table 2 also shows the QUS framework, which defines 13 criteria for the quality of user stories.

Based on QUS, Lucassen et al. present the automatic quality user story artisan (AQUSA) software tool for assessing and enhancing US quality automatically. Relying on NLP techniques, AQUSA detects quality defects and suggests possible remedies.

A US should follow a pre-defined, agreed template, chosen from the many templates available. In the conceptual model the skeleton of the template is called *format*, which the *role*, *means*, and optional *end(s)* are interspersed to form a US [40].

Because USs are a controlled language, the QUS framework's criteria are organize in *Lindland's* categories [23]:

- *Syntactic quality*, about the textual structure of a US without taking its meaning into account;

- *Semantic quality*, about the relationships and the meaning of (parts of) the US text;

- *Pragmatic quality*, takes into account not only syntax and semantics, but also the subjective interpretation of the US text by the audience.

By Lucassen et al. introduced quality criteria divided into two categories, namely *Individual*, which applies to single US, and *Set*, which applies to a bundle of USs. Individual criteria can be evaluated against an individual US:
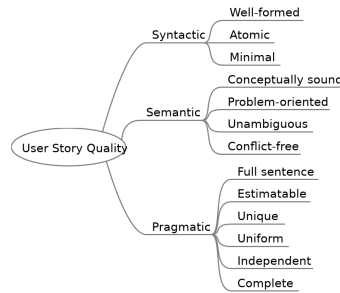

Figure 2: Agile Requirements Verification Framework [26]

well-formed

To be considered US, the core text of the request must contain a role and the expected functionality: the *means*. Looking at the US "I want to see an error message if I can't see recommendations after uploading an article". It is likely that the US author has forgotten to specify the role. The error can be fixed by adding the role: "As a member, I would like to see an error message if I cannot see recommendations after uploading an article".

Atomic

A US should only concern one characteristic. Although it is common in practice, combining multiple USs into a larger, general US affects the accuracy of effort estimation[24]. For example, the US "As a user, I can click on a specific location on the map and thereby perform a search for landmarks associated with that combination of latitude and longitude" consists of two separate requests: Clicking on a location and viewing the associated landmarks. This US should be split into two parts:

- "As a user, I am able to click on a specific location on the map";

- "As a user, I am able to see landmarks associated with the combination of latitude and longitude of a specific location".

Criteria	Description	Individual/Set
Syntactic		
Well-formed	A user story includes at least a role and a means	Individual
Atomic	A user story expresses a requirement for exactly one feature	Individual
Minimal	A user story contains nothing more than role, means, and ends	Individual
Semantic		
Conceptually sound	The means expresses a feature and the ends expresses a rationale	Individual
Problem-oriented	A user story only specifies the problem, not the solution to it	Individual
Unambiguous	A user story avoids terms or abstractions that lead to multiple interpretations	Individual
Conflict-free	A user story should not be inconsistent with any other user story	Set
Pragmatic		
Full sentence	A user story is a well-formed full sentence	Individual
Estimable	A story does not denote a coarse-grained requirement that is difficult to plan and prioritize	Individual
Unique	Every user story is unique, duplicates are avoided	Set
Uniform	All user stories in a specification employ the same template	Set
Independent	The user story is self-contained and has no inherent dependencies on other stories	Set
Complete	Implementing a set of user stories creates a feature-complete application, no steps are missing	Set

Table 2: Quality User Story framework that defines 13 criteria for user story quality [26]

Minimal

User stories should contain a role, a means and (ideally) some goals. Any additional information such as comments, descriptions of expected behaviour or notes on testing should be noted in additional notes. View the US "As a supervisor, I would like to see the registered hours for this week (split into products and activities). See: Mockup by Alice NOTE-first create the overview screen-then add validations": In addition to a role and resources, it includes a reference to an undefined mockup and an indication of how the implementation should be approached. The requirements engineer should move both to separate US attributes such as the description or comments and keep only the basic story text: "As a care professional, I would like to see this week's registered hours".

Conceptually sound

The middle and end parts of a US play a special role. The middle part should capture a specific feature, while the end part expresses the rationale for that feature. Let's look at the US "As a user, I want to open the interactive map so that I can see the location of the points of interest": The purpose is actually a dependency on another (hidden) feature that is required for the purpose to be fulfilled, which requires the presence of a database of points of interest that is not mentioned in any of the other stories. An

important additional function that is misrepresented as an end, but should be a means in a separate US, for example:

- "As a user, I would like to open the interactive map";

- "As a user, I would like to see the location of points of interest on the interactive map".

problem orientated

According to the principle of problem specification for requirements engineering proposed by Zave and Jackson, a US should only specify the problem. If absolutely necessary, implementation notes can be included as comments or descriptions. Apart from the violation of the minimum quality criteria, this US contains "As a care professional, I would like to save a refund - Save button top right (never greyed out)" Implementation details (a solution) within the US text. The text could be rewritten as follows "As a carer I would like to save a reimbursement".

Unambiguous

Ambiguity is inherent in natural language requirements, but the requirements engineer writing USs must avoid it as much as possible. A US should not only be internally unambiguous, but should also be unambiguous in relation to all other USs. The taxonomy of ambiguity types [6] provides a comprehensive overview of the types of ambiguity that can occur in a systematic requirements specification.

In this US "As a user, I am able to edit the content I have added to a person's profile page", "content" is a superclass that refers to audio, video, and text media uploaded to the profile page, as specified in three other, separate user stories in the real US record. The requester should explicitly mention which media is editable; for example, the story can be modified as follows: "As a user, I am able to edit video, photo and audio content that I have added to a person's profile page".

Full sentence

A US should read like a complete sentence, without typos or grammatical errors. The US "server configuration", for example, is not formulated as a complete sentence (and does not correspond to the syntactic quality). By reformulating the feature as a complete sentence US, it automatically specifies what exactly needs to be configured. For example, US "Server configuration" can be converted to "As an administrator, I would like to configure the sudo-ers of the server".

Estimatable

The larger and more complex the US becomes, the more difficult it is to accurately estimate the required effort. Therefore, any US should not become so large that it

becomes impossible to estimate and plan with reasonable certainty[3].For example, the US "As a carer, I would like to see my route list for the next/future days so that I can prepare (e.g. I can see when I should start the journey)" a route list so that carers can prepare.

This may be an unordered list of places to visit during the working day. However, it is equally likely that the function includes an algorithmic arrangement of routes to minimise the distance travelled and/or display the route on a map. These many functionalities make an accurate estimate difficult and make it necessary to split the US into several user stories, for example:

- "As a carer, I would like to see my route list for the next/future days so that I can prepare";

- "As a manager, I would like to upload a route list for carers".

The following quality criteria refer to a collection of USs. These quality criteria are relevant for assessing the quality of the entire project specification, as they focus on the entirety of the project specification and not on the individual review of each individual story:

Unique and conflict-free

The concept of unique USs, which emphasises the avoidance of semantic similarity or duplication within a project. For example, consider EP_a: "As a visitor, I can see a list of messages so I can stay up to date" and US_a: "As a visitor, I can see a list of messages so I can stay up to date". This situation can be improved by offering more specific messages, such as:

- US_{a1} "As a visitor, I am able to see the latest news;"

- US_{a2} "As a visitor I am able to see sports news"

It is also important to avoid conflicts between user stories to ensure their quality. A requirements conflict occurs when two or more requirements cause an inconsistency[31] [34]. For example, consider the story US_b: "As a User, I am able to edit any landmark" contradicts the requirement that a user can edit any landmark (US_c: "As a User, I am able to delete only the landmarks that I added"), if we assume that edit is a general term that includes delete. US_b refers to any landmark, while US_c refers only to those that the user has added. One possible way to fix this is to change US_b: "As a user, I can edit the landmarks I have added". [26]

To recognise these types of relationships, each US part must be compared with the parts of the other USs using a combination of similarity measures that are either syntactic (e.g. Levenshtein distance) or semantic (e.g. using an ontology to determine synonyms). If the similarity exceeds a certain threshold, a human analyst must analyse the USs for possible conflicts and/or duplicates.

---

[3]http://xp123.com/articles/invest-in-good-stories-and-smart-tasks/

Definition 2.1. *A US μ is a 4-tuple $\mu = (r, m, E, f)$, where r is the role, m is the mean, $E = (e_1, e_2, ...)$ is a set of ends and f is the format. A means m is a 5-tuple $m(s, av, do, io, adj)$, where s is a subject, av an action verb, do a direct object, io an indirect object and adj an adjective (io and adj can be zero, see figure 3). The set of user stories in a project is denoted by $U = (\mu_1, \mu_2, ...)$.*
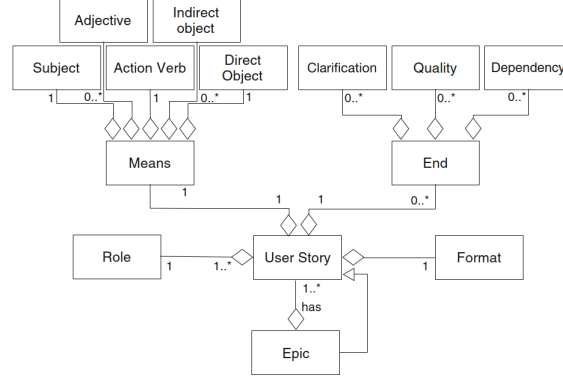

Figure 3: Conceptual model of the user stories [26]

uniform

Uniformity refers to the consistency of a US format where the majority of USs are within the same set. To assess uniformity, the Requirements Engineer determines the most common format, which is usually determined in collaboration with the team. For example, the US "As administrator, I receive an email notification when a new user is registered" is presented as a non-uniform US and can be rewritten as follows to improve uniformity: "As an administrator, I would like to receive an email notification when a new user is registered".

Independent

USs should be able to be planned and implemented in any order and should not overlap conceptually.

It is recommended that all dependencies are made explicit and visible, as complete independence is not always achievable. In addition, some interdependencies may not be possible to resolve, and you may want to consider making these interdependencies visible in a practical way, for example by adding notes to story cards or by linking to them in the issue tracker. Two examples of dependency are given:

- *causality*: Sometimes a US (l_1) needs to be completed before another (l_2) may start. which states that l_1 is causally dependent on l_2 if certain conditions are satisfied.

- *superclasses*: USs can contain an object (*e.g.*, "content" in US "As a user, I can edit the content I've added to a person's profile page") that references several other objects in different histories. This means that the object in l_1 serves as a parent or superclass for the other objects.

Complete

The implementation of a series of USs should result in a complete application. Whilst it is not necessary for the USs to cover 100% of the application's functionality from the outset, it is important not to overlook any essential USs as this can create a significant functionality gap that hinders progress. Take for example the US "As a user, I can edit the content I have added to a person's profile page", which requires the presence of another story describing the creation of content. This scenario can be extended to USs with action verbs that refer to non-existent direct objects, such as reading, updating or deleting an item, which requires its prior creation. To address these dependencies with respect to the direct object of the agent, Lucassen et al. introduce a conceptual relation.

The Automatic Quality User Story Artisan (AQUSA)

The QUS framework provides guidelines for improving the quality of USs. To support the framework, Lucassen et al. propose the AQUSA tool, which exposes defects and deviations from good US practice [26]. AQUSA primarily targets easily describable and algorithmically determinable defects in the clerical part of requirements engineering, focusing on syntactic and some pragmatic criteria, while omitting semantic criteria that require a deep understanding of requirements' content [26]. AQUSA consists of five main architectural components (Figure 4): linguistic parser, US base, analyzer, enhancer, and report generator.
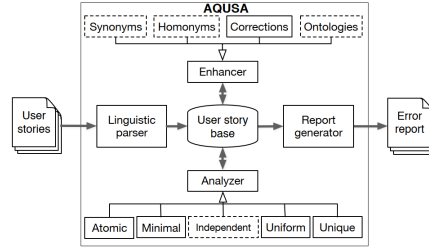
Figure 4: Functional view on architecture of AQUSA. Dashed components are not fully implemented yet [26]

The first step for every US is validating that it is well-formed. This takes place in the linguistic parser, which separates the US in its role, means and end(s) parts. The US base captures the parsed US as an object according to the conceptual model, which acts as central storage. Next, the analyzer runs tailormade method to verify specific syntactic and pragmatic quality criteria—where possible enhancers enrich the US base,

improving the recall and precision of the analyzers. Finally, AQUSA captures the results in a comprehensive report [26].

In the case of story analysis, AQUSA v1 conducts multiple analyses, beginning with the *StoryChunker* and subsequently executing the Unique-, Minimal-, WellFormed-, Uniform-, and *AtomicAnalyzer* modules. If any of these modules detect a violation of quality criteria, they engage the *DefectGenerator* to record the defect in the associated database tables related to the story. Additionally, users have the option to utilize the AQUSA-GUI to access a project list or view a report of defects associated with a set of stories.

Linguistic Parser: Well-Formed

One of the essential aspects of verifying whether a string of text is a US is splitting it into *role*, *means*, and *end(s)*. This initial step is performed by the linguistic parser, implemented as the StoryChunker component. It identifies common indicators in the US, such as "As a", "I want to", "I am able to", and "so that". The linguistic parser then categorizes words within each chunk using the Stanford NLP POS Tagger and validates the following rules for each chunk:

- Role: Checks if the last word is a noun representing an actor and if the words preceding the noun match a known role format (*e.g.*, "as a").

- Means: Verifies if the first word is "I" and if a known means format like "want to" is present. It also ensures the remaining text contains at least one verb and one noun (*e.g.*, "update event").

- End: Checks for the presence of an end and if it starts with a recognized end format (*e.g.*, "so that").

The linguistic parser validates whether a US adheres to the conceptual model. When it cannot detect a known means format, it retains the full US and eliminates the role and end sections. If the remaining text contains both a verb and a noun, it's tagged as a "potential means," and further analysis is conducted. Additionally, the parser checks for a comma after the role section.

User Story Base and Enhancer

Linguistically parsed USs are transformed into objects containing role, means, and ends components, aligning with the first level of decomposition in the conceptual model. These parsed USs are stored in the US base for further processing. AQUSA enriches these USs by adding potential synonyms, homonyms, and relevant semantic information sourced from an ontology to the pertinent words within each chunk. Additionally, AQUSA includes a corrections' subpart, ensuring precise defect correction where possible.

Analyzer: Explicit Dependencies

AQUSA enforces that USs with explicit dependencies on other USs should include navigable links to those dependencies. It checks for numbers within USs and verifies whether these numbers are enclosed within links. For instance, if a US reads, "As a care professional, I want to edit the planned task I selected—see 908," AQUSA suggests changing the isolated number to "See PID-908," where PID represents the project identifier. When integrated with an issue tracker like Jira or Pivotal Tracker, this change would automatically generate a link to the dependency, such as "see PID-908 (http://company.issuetracker.org/PID-908." It's worth noting that this explicit dependency analyzer has not been implemented in AQUSA v1 to ensure its universal applicability across various issue trackers.

Analyzer: Atomic

AQUSA examines USs to ensure that the means section focuses on a single feature. To do this, it parses the means section for occurrences of the conjunctions "and, &, +, or". If AQUSA detects double feature requests in a US, it includes them in its report and suggests splitting the US into multiple ones. For example, a US like "As a User, I'm able to click a particular location from the map and thereby perform a search of landmarks associated with that latitude-longitude combination" would prompt a suggestion to split it into two USs: (1) "As a User, I want to click a location from the map" and (2) "As a User, I want to search landmarks associated with the lat-long combination of a location."

AQUSA v1 verifies the role and means chunks for the presence of the conjunctions "and, &, +, or". If any of these conjunctions are found, AQUSA checks whether the text on both sides of the conjunction conforms to the QUS criteria for valid roles or means. Only if these criteria are met, AQUSA records the text following the conjunction as an atomicity violation.

Analyzer: Minimal

AQUSA assesses the minimality of USs by examining the role and means of sections extracted during chunking and *well-formedness* verification. If AQUSA successfully extracts these sections, it checks for any additional text following specific punctuation marks such as dots, hyphens, semicolons, or other separators. For instance, in the US "As a care professional I want to see the registered hours of this week (split into products and activities). See: Mockup from Alice NOTE: First create the overview screen—Then add validations," AQUSA would flag all text following the first dot (".") as non-minimal. Additionally, any text enclosed within parentheses is also marked as non-minimal. AQUSA v1 employs two separate minimality checks using regular expressions. The first check searches for occurrences of special punctuation marks like ", -, ?, ., *." and marks any text following them as a minimality violation. The second check identifies text enclosed in brackets such as "(), [], {}, <>" and records it as a minimality violation.

Analyzer: Uniform

AQUSA, in addition to its chunking process, identifies and extracts the format parts of USs and calculates their occurrences across all USs in a set. The most frequently occurring format is designated as the standard US format. Any US that deviates from this standard format is marked as non-compliant and included in the error report. For example, if the standard format is "I want to," AQUSA will flag a US like "As a User, I am able to delete a landmark" as non-compliant because it does not follow the standard. After the linguistic parser processes all USs in a set, AQUSA v1 initially identifies the most common US format by counting the occurrences of indicator phrases and selecting the most frequent one. Later, the uniformity analyzer calculates the edit distance between the format of each individual US chunk and the most common format for that chunk. If the edit distance exceeds a threshold of 3, AQUSA v1 records the entire story as a uniformity violation. This threshold ensures that minor differences, like "I am" versus "I'm," do not trigger uniformity violations, while more significant differences in phrasing, such as "want" versus "can," "need," or "able," do.

Analyzer: Unique

AQUSA has the capability to utilize various similarity measures, leveraging the WordNet lexical database to detect semantic similarity. For each verb and object found in the means or end of a US, AQUSA performs a WordNet::Similarity calculation with the corresponding verbs or objects from all other USs. These individual calculations are combined to produce a similarity degree for two USs. If this degree exceeds 90%, AQUSA flags the USs as potential duplicates.

AQUSA-GUI: report generator

After AQUSA detects a violation in the linguistic parser or one of the analyzers, it promptly creates a defect record in the database, including details such as the defect type, a highlight of where the defect is located within the US, and its severity. AQUSA utilizes this data to generate a comprehensive report for the user. The report begins with a dashboard that provides a quick overview of the US set's quality. It displays the total number of issues, categorized into defects and warnings, along with the count of perfect stories. Below the dashboard, all USs containing issues are listed, accompanied by their respective warnings and errors. An example is illustrated in figure 5.

Conclusion

In the QUS framework, a conflict is defined as a requirements conflict that occurs when two or more requirements cause an inconsistency. As far as the inconsistency is concerned, it is not clear to which form it actually belongs. Is it a content inconsistency or an inconsistency in the sense that two USs are very close to each other and describe the matter slightly differently, which leads to an inconsistency that we can understand as a similarity?
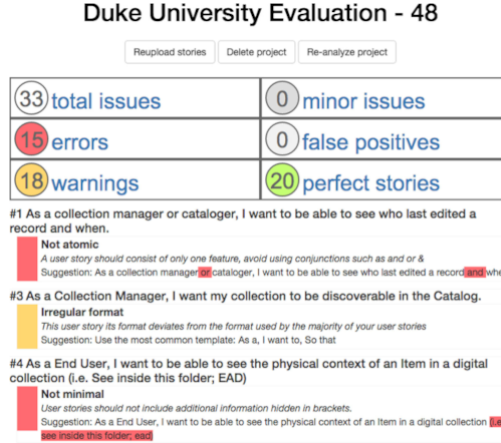
Figure 5: Example report of a defect and warning for a story in AQUSA [26]

If we look at the analysis of the similarity of texts, *e.g.* with transformation tools or similar, we realize that our analysis cannot effectively recognise strong overlaps between USs and will not act precisely. We try to capture the content of the USs through the annotation just through the action, so we can do very lightweight conflict analysis similarity of texts. Very lightweight because there are many inaccuracies, as is often the case with the USs. But otherwise it is as if the analysis of two USs are similar but not the same, and somehow contradict each other.

One is the conflict that can arise in the final described system versus the inconsistencies during the specification of USs. The content-related conflict in the quality criterion is not even mentioned. They described how the USs were written down and do not refer to content conflicts.

Similarly, with dependencies, there are explicit dependencies to define in the USs by saying this US is based on that and giving it an ID. In another case, we can analyse implicit dependencies by determining, *e.g.*, something has to be created at this position first so that it can be used in another position and so creating and using it happens in two different USs. We call this implicit dependency because it is not explicitly stated in USs, but if we analyse what is there semantically, we can find out that there is a dependency.

In addition, we could use the US-ID to obtain information about the implicit dependencies between USs and compare these with the explicit dependencies to ensure that explicit and implicit dependencies are consistent with each other?

Both conflicts and dependencies are interesting and both terms and both analyses we would sharpen and make clear what can be analysed at all. However, this only partially fits in with the quality framework here from AQUSA-tool.

Example 2.1. *Considering two USs: US_1: "As a user, I am able to edit any landmark" and US_2: "As a user, I am able to delete only the landmarks that I added". First, we try to minimize US_1 and divided it into three USs namely:*

- US_1a: "As a user, I am able to add any landmark."

- US_1b: "As a user, I am able to modify any landmark."

- US_1c: "As a user, I am able to delete any landmark."

US_1c means that two users are allowed to delete the same landmark, which would lead to a conflict. This conflict can be avoided if US_1c is replaced with US_2, as two users are then no longer allowed to delete the same landmark. Furthermore, this situation cause an inconsistency between US_1c and US_2, e.g. if US_2 deletes the landmark that was added by a particular user, US_1c can no longer find this landmark and vice versa.

Moreover, analyzing dependencies through AQUSA v1 is a component of the forthcoming effort by Lucassen et al. Which means, the automatic analysis of dependencies and conflicts between USs is an area that requires future attention and development. In this context, we would like to contribute by addressing this unmet need.

3 Preliminaries

In the 3.1 Section, we introduce *conditional random fields* (CRF) for the extraction of domain models from agile product backlogs, which play a central role in effectively supporting the identification of dependencies and conflicts between user stories. Furthermore, we conduct a conclusion.

Next, we dive into the Section *nlp* and introduce *natural language processing* (NLP) and *VerbNet* as a computational lexicon resource as well as a conclusion.

Finally, in Section 3.3, some basic definitions of *graphs* and *graph transformation rules* are laid down for better understanding. We then look at the graph transformation tool *Henshin* and its extension *critical pair analysis* (CPA), which plays a central role in our methodology. We also draw a conclusion and explain why we have chosen these techniques and what their central idea is.

3.1 Extracting Domain Models from Textual Requirements

Automated support for extracting domain models from requirements artifacts such as USs play a central role in effectively supporting the detection of dependencies and conflicts between user stories. Domain models are a simple way to understand the relationship between artifacts and the whole system. In this subsection, we present a graph-based extraction modelling approach called CRF and conclude our review.

A Modelling Backlog as Composable Graphs

Mosser et al. propose a model engineering method (and the associated tooling) to exploit a graph-based meta-modelling and compositional approach. The objective is to shorten the feedback loop between developers and POs while supporting agile development's iterative and incremental nature.

The tool can extract what is called a conceptual model of a backlog in an ontology-like way. The conceptual models are then used to measure USs quality by detecting ambiguities or defects in a given story [30]. From a modelling point of view, Mosser et al. represents the concepts involved in the definition of a backlog in a metamodel, as depicted in figure 6. Without surprise, the key concept is the notion of story, which brings a benefit to a *Persona* thanks to an Action performed on an *Entity*. A Story is associated to a readiness *Status*, and might optionally contribute to one or more *QualityProperty* (*e.g.*, security, performance).
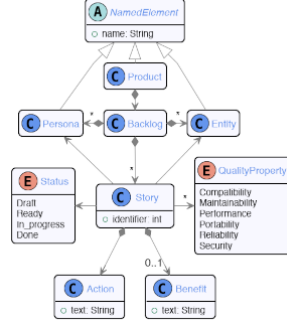

Figure 6: Backlog conceptual metamodel [30]

Consider, for example, the following story, extracted from the reference dataset [11]: "As a user, I want to click on the address so that it takes me to a new tab with Google Maps.". *This story brings to the user (Persona) the benefit of reaching a new Google Maps tab (Benefit) by clicking (Action) on the displayed address (Entity).*

As Entities and Personas implement the *jargon* to be used while specifying features in the backlog, they are defined at the *Backlog level*. On the contrary, Actions belong to the associated stories and are not shared with other stories. Finally, a *Product* is defined as the *Backlog* used to specify its features.

Mosser et al. propose in the context of backlog management a system which represented in figure 7 is proposed for utilization. Building upon the efficiency of NLP approaches. Mosser et al. suggest employing an NLP-based extractor to create a backlog model. This model will subsequently assist teams in the planning phase by aiding in the selection of stories for implementation during the upcoming iteration [30].

Composable Backlogs

In order to support team customization (*e.g.*, a given team might want to enrich the backlog metamodel with additional information existing in their product management system) Mossser et al. chose open-world(ontological) representation by modelling backlog as graphs [30]. The graph is equipped with constraints (*e.g.*, a story always refers to a persona and an entity) to ensure that the minimal structure captured in the previously defined metamodel is guaranteed.
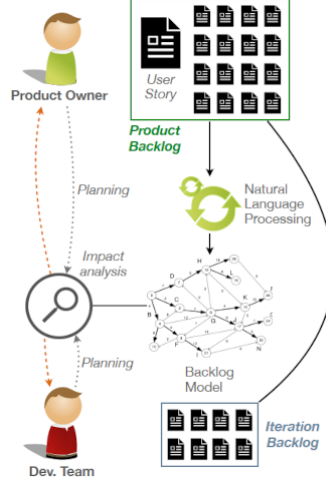
Figure 7: Providing early feedback at the backlog level [30]

Definition 3.1 (**Story**). *A Story$s \in S$ is defined as a tuple (P, A, E, K), where $P = \{p_1, ..., p_i\}$ is the set of involved personas, $A = \{a_1, ..., a_i\}$ the set of performed actions, and $E = \{e_1, ..., e_k\}$ the set of targeted entities. Additional knowledge (e.g., benefit, architectural properties, status) can be declared as key-value pairs in $K = \{(k_1, v_1), ..., (k_l, v_l)\}$. The associated semantics is that the declared actions bind personas to entities. Considering that story independence is a pillar of agile methods (as, by definition, stories are independent inside a backlog), there is no equivalence class defined over $S : \forall(s, s') \in S^2, s \neq s' \Rightarrow s \not\equiv s'$.*

Definition 3.2 (**Backlog**). *A backlog $b \in B$ is represented as an attributed typed graph $b = (V, E, A)$, with V a set of typed vertices, E a set of undirected edges linking existing vertices, and A a set of key-value attributes. Vertices are typed according to the model element they represent $v \in V, type(v) \in \{Persona, Entity, Story\})$. Edges are typed according to the kind of model elements they are binding. Like backlogs, vertices and edges can contain attributes, represented as (key, value) pairs. The empty backlog is denoted as $\emptyset = (\emptyset, \emptyset, \emptyset)$.*

Example 3.1. *Backlog excerpt: Content Management System for Cornell University — CulRepo [11].*

- *s_1: As a faculty member, I want to access a collection within the repository.*

Associated model:

- *$s_1 = (\{facultymember\}, \{access\}, \{repository, collection\}, \emptyset) \in S$*

A backlog containing a single story s_1: ("As a faculty member, I want to access a collection within the repository").

$b_1 = (V_1, E_1, \emptyset) \in B$
$V_1 = \{Persona(faculty\ member, \emptyset)\ ,$
$\qquad Stroy\,(s_1, \{(action, access)\})$
$\qquad Entity\,(repository, \emptyset),$
$\qquad Entity(collection, \emptyset)\}$
$E_1 = \{has\_for\_persona(s_1, faculty\ member),$
$\qquad has\_for\_entity\,(s_1, repository)$
$\qquad has\_for\_entity(s_1, collection)\}$

Conditional Random Fields (CRF)

CRFs [21] are a particular class of *Markov Random Fields*, a statistical modelling approach supporting the definition of discriminative models. They are classically used in pattern recognition tasks (labelling or parsing) when context is important identify such patterns [4].

To apply CRF Mosser et al. transform a given story into a sequence of tuples. Each tuple contains minimally three elements: *(i)* the original word from the story, *(ii)* its syntactical role in the story, and *(iii)* its semantical role in the story. The syntactical role in the sentence is classically known as *Part-of-Speech* (POS), describing the grammatical role of the word in the sentence. The semantical role plays a dual role here. For training the model, the tags will be extracted from the annotated dataset and used as target. When used as a predictor after training, these are the data Mosser et al. will ask the model for infer.

The main limitations of CRF are that *(i)* it works at the word level (model elements can spread across several words), and *(ii)* it is not designed to identify relations between entities [4]. To address the first limitation, Mosser et al. use a glueing heuristic. Words that are consecutively associated with the same label are considered as being the same model element, *e.g.*, the subsequence ["UI", "designer"] from the previous example is considered as one single model element of type *Persona*.

Mooser et al. applied this heuristic to everything but verbs, as classically, two verbs following each other represent different actions. They used again heuristic approach to address the second limitation. Mooser et al. bound every *Persona* to every primary *Action* (as *trigger* relations), and every primary *Actions* to every primary *Entity* (as *target* relations) [4].

Example 3.2. *Consider the following example:*

$S = [' As',' a',' UI',' designer',',',\ .\ .\ .]$
$POS(S) = [ADP, DET, NOUN, NOUN, PUNCT, . . .]$
$Label\,(S) = [\emptyset, \emptyset, PERSONA, PERSONA, \emptyset, . . .]$

S represents a given US (Table 6). POS (S) represent the Part-of-speech analysis of S. The story starts with an adposition (ADP), followed by determiner (DET), followed by a noun, followed by another noun, ... Then, Lables (S) represents what we interest

in: the first two words are not interesting, but the 3rd and 4^{td} words represent a Persona. A complete version of the example is provided in Table 6.
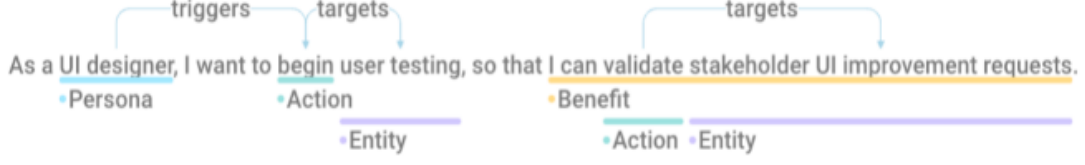

Figure 8: Example of annotated user using Doccano Annotation UI [4]

Word	As	a	UI	designer	,	I	want	to	begin	user	testing	,
POS	ADP	DET	NOUN	NOUN	PUNCT	PRON	VERB	PART	VERB	NOUN	NOUN	PUNCT
Label	-	-	PER	PER	-	-	-	-	P-ACT	P-ENT	P-ENT	-

Word	so	that	I	can	validate	stakeholder	UI	improvement	requests	.
POS	SCONJ	SCONJ	PRON	AUX	VERB	NOUN	NOUN	NOUN	NOUN	PUNCT
Label	-	-	-	-	S-ACT	S-ENT	S-ENT	S-ENT	S-ENT	-

POS tags are the Universal POS tags
Labels: PER (Persona), P-ACT (Primary Action), P-ENT (Primary Entity), S-ACT (Secondary Action), S-ENT (Secondary Entity)

Table 6: Minimal Feature Set, associating part-of-speech (POS) and semantic labels to each word in a given story [4]

Conclusion

Conditional Random Fields (CRF) approach is graph-based and promises a high degree of precision and recall, which is particularly important in the context of domain concept extraction. CRF can cover both syntactic and semantic aspects, especially when complemented by a suitable conceptual metamodel, making them suitable for definition as a type graph in Henshin. The annotations generated by CRF can then be used for transformation into a rule-based graph transformation system, which improves support for DevOps practices.

3.2   NLP and VerbNet as a Computational Lexical Resource

Natural language processing (NLP) is a computational method for the automated analysis and representation of human language [8]. NLP techniques offer potential advantages to improve the quality of USs and can be used to parse, extract, or analyze US's data. It has been widely used to help in the software engineering domain (*e.g.*, managing software requirements [3], extraction of actors and actions in requirement document [1].

NLP techniques are usually used for text preprocessing (*e.g.*, tokenization, *Part-of-Speech* (POS) tagging, and dependency parsing). Several NLP approaches can be used (*e.g.*, syntactic representation of text and computational models based on semantic features). Syntactic methods focus on word-level approaches, while the semantic focus on multiword expressions [8].

A computational lexicon resource is a systematically organized repository of words or terms, complete with linguistic and semantic data. These lexicons play a pivotal role in facilitating NLP systems focused on semantic analysis by offering comprehensive insights into language elements, encompassing word forms, part-of-speech (POS) categories, phonetic details, syntactic properties, semantic attributes, and frequency statistics. Lexical classes, defined in terms of shared meaning components and similar (morpho-) syntactic behaviour of words, have attracted considerable interest in NLP [8]. These classes are useful for their ability to capture generalizations about a range of (cross-) linguistic properties. NLP systems can benefit from lexical classes in a number of ways. As the classes can capture higher level abstractions (*e.g.* syntactic or semantic features) they can be used as a principled means to abstract away from individual words when required. Their predictive power can help compensate for lack of sufficient data fully exemplifying the behaviour of relevant words [19].

Upon the completion of the transformation of USs utilizing a Conditional Random Fields (CRF) approach, wherein entities, actions (both primary and secondary), persona, and their relational attributes (specifically, triggers, targets and contains) are meticulously annotated and structured as a graph-based representation, a preliminary imperative emerges. This imperative entails the determination of a representative semantic interpretation for the ascertained actions. This determination, in turn, serves as a prerequisite for the generation of corresponding transformation rules, namely, "create", "delete", "preserve" and "forbidden" rules.

The attainment of this representative semantic interpretation hinges upon the application of a suite of foundational lexical resource techniques of a conceptual nature. These techniques assume a pivotal role in furnishing the essential cognitive infrastructure, facilitating a comprehensive grasp of the semantic roles, syntactic characteristics, and the systematic categorization to "creation", "deletion", "preservation" and "forbiddance" of linguistic elements embedded within the construct of the US.

VerbNet

VerbNet (VN) is a hierarchical domain-independent, broad-coverage verb lexicon with mappings to several widely-used verb resources, including WordNet [28], Xtag [32], and FrameNet [5]. It includes syntactic and semantic information for classes of English verbs derived from Levin's classification, which is considerably more detailed than that included in the original classification.

Each verb class in VN is completely described by a set of members, thematic roles for the predicate-argument structure of these members, selectional restrictions on the arguments, and frames consisting of a syntactic description and semantic predicates with a temporal function, in a manner similar to the event decomposition of Moens and

Steedman [29]. The original Levin classes have been refined, and new subclasses added to achieve syntactic and semantic coherence among members.

Syntactic Frames

Semantic restrictions, such as constraints related to animacy, humanity, or organizational affiliation, are employed to limit the types of thematic roles allowed within these classes. Furthermore, each syntactic frame may have constraints regarding which prepositions can be used.

Additionally, there may be additional constraints placed on thematic roles to indicate the likely syntactic nature of the constituent associated with that role. Levin classes are primarily characterized by NP (noun phrase) and PP (prepositional phrase) complements.

Some classes also involve sentential complementation, albeit limited to distinguishing between finite and non-finite clauses. This distinction is exemplified in VN, particularly in the frames for the class Tell-37.2, as shown in Examples (1) and (2), to illustrate how the differentiation between finite and non-finite complement clauses is implemented.

1. Sentential Complement (finite):
   "Susan told Helen that the room was too hot."
   Agent V Recipient Topic [+sentential – infinitival]

2. Sentential Complement (nonfinite):
   "Susan told Helen to avoid the crowd."
   Agent V Recipient Topic [+infinitival – wh_inf]

Semantic Predicates

Each VN frame also contains explicit semantic information, expressed as a conjunction of Boolean semantic predicates such as "motion", "contact", or "cause". Each of these predicates is associated with an event variable E that allows predicates to specify when in the event the predicates are true $start(E)$ for preparatory stage, $during(E)$ for the culmination stage, and $end(e)$ for the consequent stage).

Relations between verbs (or classes) such as antonymy and entailment present in WordNet and relations between verbs (and verb classes) such as the ones found in FrameNet can be predicted by semantic predicates. Aspect in VN is captured by the event variable argument present in the predicates.

The VerbNet Hierarchy

VerbNet represents a hierarchical structure of verb behaviour, with groups of verb classes sharing similar semantics and syntax. Verb classes are numbered based on common semantics and syntax, and classes with the same top-level number (e.g., 9-109) have corresponding semantic relationships.

For instance, classes related to actions like "putting", such as "put-9.1", "put_spatial-9.2", "funnel-9.3", all belong to class number 9 and relate to moving an entity to a location. Classes sharing a top class can be further divided into subclasses, as seen with "wipe" verbs categorized into "wipe_manner" (10.4.1) and "wipe_inst" (10.4.2) specifying the manner and instrument of "wipe" verbs in the "Verbs of Removing" group of classes (class number 10).

The classification encompasses class numbers 1-57, derived from Levin's classification [22], and class numbers 58-109, developed later by Korhonen and Briscoe [20]. The later classes are more specific, often having a one-to-one relationship between verb type and verb class. This hierarchical structure helps categorize and organize verbs based on their semantic and syntactic properties.

Verb Class Hierarchy Contents

Each individual verb class within VerbNet is hierarchical. These classes can include one or more "subclasses" or "child" classes, as well as "sister" classes. All verb classes have a top-level classification, but some provide further specification of the behaviours of their verb members by having one or more subclasses.

Subclasses are identified by a dash followed by a number after the class information. For example, the top class might be "spray-9.7", and a subclass would be denoted as "spray 9.7-1". This hierarchy allows for a more detailed and structured organization of verb behaviour within VerbNet.

- Top Class: The highest class in the hierarchy; all features in the top class are shared by every verb in the class. The top class of the hierarchy consists of syntactic constructions and semantic role labels that are shared by all verbs in this class.

- Parent Class: Dominates a subclass; all features are shared with subordinate classes.

- Subclasses: VerbNet subclasses inherit features from the top class but specify further syntactic and semantic commonalities among their verb members. These can include additional syntactic constructions, further selectional restrictions on semantic role labels, or new semantic role labels.

- Child Class: Is dominated by a parent class; inherits features from this parent class, but also adds information in the form of additional syntactic frames, thematic roles, or restrictions.

- Sister Class: A subclass directly dominated by a parent class. This parent class also, directly dominates another subclass, so the two subclasses are sisters to one another. Sister classes do not share features.

Figure 9 illustrate an example of class hierarchy from spray-9.7 class. Verb Classes are numbered according to shared semantics and syntax, and classes which share a top-level number (9-109) have corresponding semantic relationships.

Figure 9: Class hierarchy for spray-9.7 class [17]

For instance, verb classes related to putting, such as put-9.1, put_spatial-9.2, funnel-9.3, etc. are all assigned to the class number 9 and related to moving an entity to a location.

Classes that share a top class can also be divided into subclasses, such as wipe verbs in wipe_manner (10.4.1) and wipe_inst (10.4.2) which specify the manner and instrument of wipe verbs in the "Verbs of Removing" group of classes (class number 10).

An example of top-class numbers and their corresponding types is given in Table 8. Class numbers 1-57 are drawn directly from Levin's (1993) classification. Class numbers 58-109 were developed later in the work of Korhonen & Briscoe (2004). Notably, the verb types of the later classes are less general, as indicated by the fact that most of these classes have a one-to-one relationship between verb type and verb class.

Class Number	Verb Type	Verb Class
10	Verbs of Removing	banish-10.2
		cheat-10.6.1
		clear-10.3
		debone-10.8
		fire-10.10
		mine-10.9
		pit-10.7
		remove-10.1
		resign-10.11
		steal-10.5
		wipe_manner-10.4.1
26	Verbs of Creation and Transformation	adjust-26.9
		build-26.1
		convert-26.6.2
		create-26.4
		grow-26.2.1
		knead-26.5
		performance-26.7
		rehearse-26.8
		turn-26.6.1
13	Verbs of Change of Possession	berry-13.7
		contribute-13.2
		equip-13.4.2
		exchange-13.6.1
		fulfilling-13.4.1
		future_having-13.3
		get-13.5.1
		give-13.1
		hire-13.5.3
		obtain-13.5.2

Table 8: An example of top-class numbers and their corresponding verb-type[**?**]

Conclusion

The methodological categorisation of verbs into hierarchical classes in VerbNet provides a structured and all-encompassing framework for understanding verb behaviour. This alignment with our project goals is of utmost importance as it supports our task of formulating transformation rules based on semantic interpretations of the actions (verbs)

recognised by CRF and described in USs. VerbNet's organisational structure fits seamlessly with the requirements of our project and ensures the necessary precision and granularity that is essential for our semantic analysis and rule generation.

3.3 Analysis by Graph Transformation Tool

In a software development process, the class architecture is getting changed over the development, *e.g.* due to a change of requirements, which results in a change of the class diagram. During runtime of a software, an object diagram can also be modified trough creating or deleting of new objects.

Many structures, that can be represented as graph, are able to change or mutate. This suggests the introduction of a method to modify graphs through the creation or deletion of nodes and edges. This graph modification can be performed by the so-called graph transformations. There are many approaches to model graph transformations *e.g.* the double pushout approach or the single pushout approach, which are both concepts based on pushouts from category theory in the category Graphs.

In this section, some basic definitions about graph transformation and transformation rules are first established for better understanding. Afterwards, we dive into graph transformation tool Henshin [2], which play a pivotal role in our methodology.

Graphs and Typed Graphs

A graph consists of nodes and edges, with each edge connecting precisely two nodes and having the option to be directed or undirected. When an edge is directed, it designates a distinct start node (source) and an end node (target). For the purpose of this discussion, we will focus on directed graphs.

Definition 3.3 (**graph**). *A graph $G = (V, E, s, t)$ contains V, a set of nodes, E, a set of edges, $s : E \to V$, a source function, where $s(e)$ is the start node of $e \in E$ and a target function $t : E \to V$, where $t(e)$ is the end node of a edge $e \in E$.*

Definition 3.4 (**Transformation Rule**). *A transformation rule denotes which nodes and edges of a graph have to be deleted and which nodes and edges have to be created. In the double-pushout approach a transformation rule $p = L \xleftarrow{l} K \xrightarrow{r} R$ consists of three graphs L, K, R, two graph morphisms $l : K \to L$ and $r : K \to R$, where K contains all elements, that remain in the graph, $L \setminus l(K)$ contains the elements that are removed and $R \setminus r(K)$ contains the elements, that are created.*

Definition 3.5 (**Graph Transformation**). *In the context of graph transformations, when we have two graphs G and H, along with a transformation rule p, we can apply this rule to graph G at match m. This application, denoted as $G \xRightarrow{p,m} H$, results in graph H. The match, represented as $m : L \hookrightarrow G$, is an injective graph morphism, and L contains all the nodes and edges of p that remain intact and are not deleted during the transformation.*

As outlined in section 3.1, where we elucidated our utilization of conditional random fields (CRF) as a graph-based metamodeling and compositional approach for annotating USs, each US is meticulously structured and annotated in the form of a graph. Subsequently, we will apply transformation rules to these CRF-generated graphs to modelling graph transformation.

To accomplish this objective, we have harnessed the capabilities of established lexical resources such as VerbNet. This utilization of VerbNet enables us to categorize actions (which are verbs) extracted from the US into four distinct categories, namely: "create", "delete", "preserve", "forbid". These categories serve as essential components of transformation rules that articulate precise changes within the graph-based representation.

Henshin: A Tools for In-Place EMF Model Transformations

The Eclipse Modelling Framework (EMF) provides modelling and code generation facilities for Java applications based on structured data models. Henshin is a language and associated tool set for in-place transformations of EMF models.

The Henshin transformation language uses pattern-based rules on the lowest level, which can be structured into nested transformation units with well-defined operational semantics. So-called amalgamation units are a special type of transformation units that provide a forall-operator for pattern replacement. For all of these concepts, Henshin offers a visual syntax, sophisticated editing functionalities, execution and analysis tools. The Henshin transformation language has its roots in attributed graph transformations, which offer a formal foundation for validation of EMF model transformations [2].

Graph Types

Graph transformation-based approaches, essentially define model transformations using rules consisting of a pre-condition graph, called the left-hand side (LHS), and a post-condition graph, called the right-hand side (RHS) of the rule. Informally, the execution of a model transformation requires that a matching of objects in the model (host graph) to the nodes and edges in the LHS is found, and these matched objects are changed in such a way that the nodes and edges of the RHS match these objects [37].

The performance of graph transformation-based model transformations is mainly determined by the efficiency of the match finding of the LHS. Consequently, model transformation languages offer different options to add constraints to the LHS of model transformations to improve the performance of the matching [37]. To be efficient, graph transformation tools usually employ heuristics such as search plans to provide good performance (e.g. [38]).

Structure and Application of Rules

The Henshin transformation language is defined by means of a metamodel. The Henshin metamodel is closely aligned to the underlying formal model of double pushout (DPO) graph transformations [37]. Thus, rules consist of a left-hand side and a right-hand side

graph as instances of the *Graph* class. Rules further contain node mappings between the LHS and the RHS which are omitted here for better readability. Graphs consist of a set of Nodes and a set of Edges. Nodes can additionally contain a set of Attributes. These three kinds of model elements are typed by their corresponding concepts in the Ecore metamodel of EMF.

Application Conditions

To conveniently determine where a specified rule should be applied, application conditions can be defined. An important subset of application conditions is negative application conditions (NACs) which specify the non-existence of model patterns in certain contexts. In the Henshin transformation model, graphs can be annotated with application conditions using a *Formula*. This formula is either a logical expression or an application condition, which is an extension of the original graph structure by additional nodes and edges. A rule can be applied to a host graph only if all application conditions are fulfilled [2].

Attribute and Parameters

Nodes may also include a set of Attributes. Rules inherit from Units and can thus include various Parameters. A common use of parameters is to transmit an Attribute value (such as a name) from a node to be matched in the rule. To restrict the application of a rule, the metamodel encompasses concepts for representing nested graph conditions [14] as well as attribute conditions.

State Space Exploration

Henshin support in-place model transformation, Arendt et al. have developed a state space generation tool, which allow to simulating all possible executions of a transformation for a given input model, and to apply model checking, similar to the GROOVE [18] tool. Henshin can generate finite as well as large state space exploration. Regarding generation and analysis of large state space, the tool's ability to utilize parallel algorithms, taking advantage of modern multi-core processors, which enables the handling of state spaces with millions of states.

Analyzing Conflicts and Dependencies

The elements that comprise the system under construction interact with each other, establishing dependencies among them [18]. In Figure 10, *element A* requires *element B*, generating a dependency between them. Such dependencies are naturally inherited by the USs (US_i cannot be implemented until US_j is implemented).

Therefore, the natural dependencies between USs should be accepted as inevitable. In fact, only a fifth of the requirements can be considered with no dependencies [9]. The existence of dependencies between USs makes it necessary to have some implemented before others [9],[13],[25]. If the order of USs implementation does not consider these

dependencies, it may have a large number of preventable refactorings, increasing the total cost of the project needlessly. Identifying beforehand the dependencies increases the ability to effectively deal with changes.

Hence, light systematic mechanisms are needed to help identify dependencies between USs [12]. The critical pair analysis (CPA) for graph rewriting [15] has been adapted
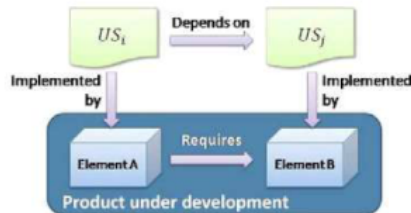

Figure 10: Inherited dependencies by user stories[12]

to rule-based model transformation, *e.g.* to find conflicting functional requirements for software systems [16], or to analyses potential causal dependencies between model refactorings [27], which helps to make informed decisions about the most appropriate refactoring to apply next. The CPA reports two different forms of potential causal dependencies, called conflicts and dependencies [7]. The application of a rule r_1 is in conflict with the application of a rule r_2 if

– r_1 deletes a model element used by the application of r_2 (**delete/use**), or

– r_1 produces a model element that r_2 forbids (**produce/forbid**), or

– r_1 changes an attribute value used by r_2 (**change/use**)[4].

Different between Model Checking and Conflict and Dependency Analysis

In this subsection, we shall delineate two distinct analytical methodologies, specifically model checking and Conflict and Dependency Analysis. Their respective purposes are delineated in Table 10, which serves to elucidate their appropriateness for modelling USs.

Example 3.3. *Now, we exemplify conflicts and dependencies within Henshin using two US namely US_1:"As an administrator, I can add a new person to the database" and US_2: "As a visitor, I can view a person's profile"). Figure 11 delineates the class model LDAP (Lightweight Directory Access Protocol). In Figure 12, the defined rules in Henshin, specifically the rule view_profile linked to US_2, and add_person_profile corresponding to US_1, are depicted. The representation uses black to signify object preservation and green for new objects. In addition, Figure 13 (CDA result) shows the dependencies between US_2 and US_1 as "Create dependency", which highlights that* Profile *node must first be*

---

[4]Dependencies between rule applications can be characterized analogously.

Aspect	Model Checking for User Stories	Conflict and Dependency Analysis for User Stories
Purpose	Verify user story properties and system behavior	Understand dependencies and interactions between user stories
Method	Large state spaces exploration	Rule-based model transformation
Automated vs. Manual	Automated	Automated
Scope	Ensuring user stories meet specified requirements and system behavior	Understanding how user stories relate to each other, managing dependencies
Use Cases	Ensuring user story correctness and system behavior	Agile development, impact analysis, and managing user story dependencies
Result	Verification of user story properties (e.g., acceptance criteria)	Identification of user story dependencies, potential conflicts, and their impact on the development process

Table 10: Comparative analysis between model checking and conflict and dependency methods

created by US_1 in order to be used in US_2. Last but not least, a special instance graph is not required in Henshin.

Example 3.4. *Looking at the USs mentioned in the example 2.1, US_1: "As a user, I am able to edit any landmark." and US_2:"As a user, I am able to delete only the landmarks that I added.". First, we minimize US_1 and divided it into three USs as follows:*

- *US_1a:"As a user, I am able to add any landmark."*

- *US_1b:"As a user, I am able to modify any landmark."*

- *US_1c:"As a user, I am able to delete any landmark."*

Figure 14 shows the class model "Map" while figure 15 shows the defined rules in Henshin, with each rule corresponding to a user story.

In this example, we assume that a user only performs one action at a time. If US_1c is translated into a rule and then CDA (Conflict and Dependency Analysis) is applied (Figure 16), Henshin would find a "Delete-Delete" conflict between two "Actions (verbs)" in US_1c and US_2 where two users are allowed to delete the same landmark. This conflict can be avoided if US_1c is replaced with US_2, as two users are then no longer allowed to delete the same landmark.

Figure 16 shows the conflicts and dependencies between two rules. For Instance, there is a "Create-Delete" conflict between US_1b and US_2. If the specific landmark are deleted, US_1b cannot modify that landmark anymore.
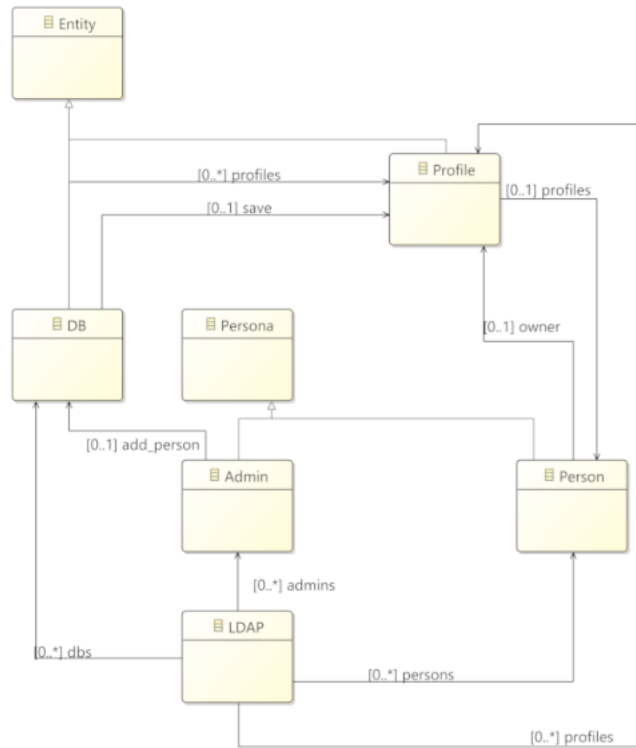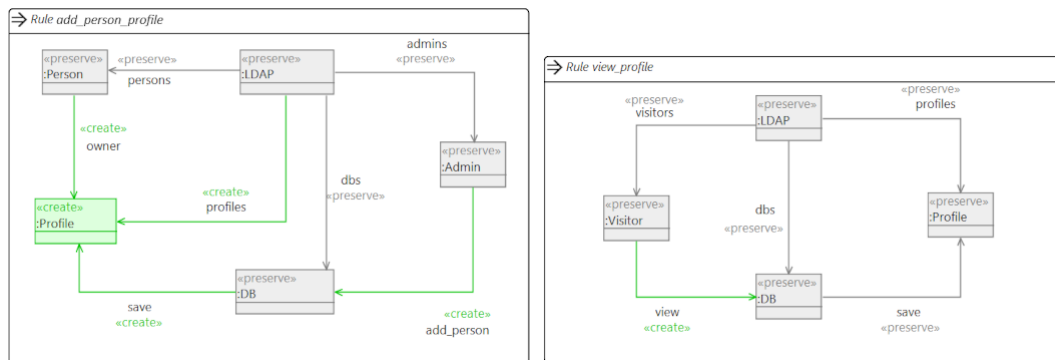
Figure 11: Henshin Class Model LDAP

Figure 12: Illustrated rules in Henshin: view_profile rule related to US_2 and add_person related to US_1
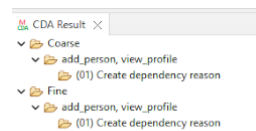

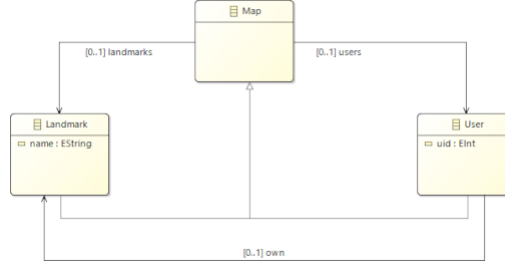Figure 13: Henshin CDA result visualises dependencies between user stories

Figure 14: Henshin Class Model Map

CPA Tool

The provided CPA extension of Henshin can be used in two different ways: Its application programming interface (API) can be used to integrate the CPA into other tools and a user interface (UI) is provided supporting domain experts in developing rules by using the CPA interactively [27].

After invoking the analysis, the rule set and the kind of critical pairs to be analyzed have to be specified. Furthermore, options can be customized to stop the calculation after finding a first critical pair, to ignore critical pairs of the same rules, etc. The resulting list of critical pairs is shown and ordered along rule pairs.

Conclusion

It is notable that Henshin's critical pair analysis (CPA) remains agnostic to the specifics of instance graphs. Consequently, model checking in Henshin merely requires the set of rules. In the Henshin approach, instance graphs are not explicitly selected; instead, only rule pairs are analysed, considering conflicts and dependencies.

Since Henshin enables the specification of constraints and conditions within rules, which can be useful for enforcing and verifying US requirements to ensure that constraints are met.

Moreover, a compelling factor in favor of Henshin is the inherent support of a versatile *application programming interface* (API) by the CPA extension, facilitating seamless integration of CPA functionality into various tools, including Java-based platforms.

4   Analysing Redundancy

USs are a fundamental component of agile software development. They provide concise and clear descriptions of the software functionality from the end user's perspective. These stories guide the development process and ensure that the software meets the actual needs of the users.

However, as the project backlog grows, they often become overloaded with redundant USs that can obscure project priorities, waste resources and complicate maintenance.
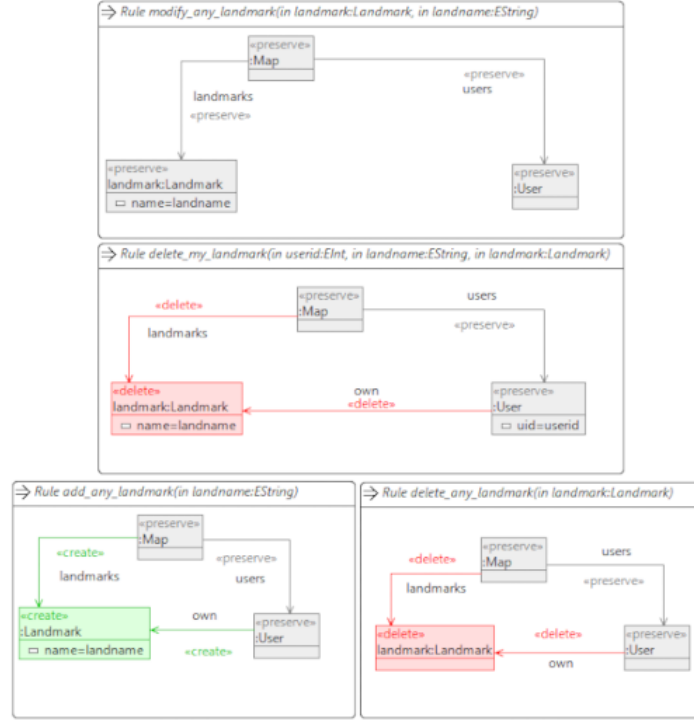
Figure 15: Illustrated rules in Henshin: add_any_landmark rule related to US_1a, modify_any_landmark rule related to US_1b, delete_any_landmark rule related to US_1c and delete_my_landmark rule related to US_2

Analysing the redundancy within these USs is key to maintaining clarity and efficiency in the development process.

The main goal of this analysis is to streamline the software development workflow by identifying redundancies in the user USs within a project's backlog.

By identifying and consolidating overlapping functionalities, companies can avoid redundant work, streamline development work and optimise testing processes. This strategic approach not only shortens project timelines but also improves the clarity and coherence of project documentation, making it easier for team members to navigate and manage project requirements.

In section 4.1 we present the requirements and functional needs that are used as input for the design phase to fulfil the requirements. In section 4.2 we explain the design decisions of the workflow shown in Figure ?? and explain how the architecture is structured.

In section 4.3 follows the implementation to show how the components and their classes look in relation to figure ??. In Section 4.4 we show what and woh far we tested it. Finally, we apply our approach to 19 datasets of backlogs annotated with the CRF
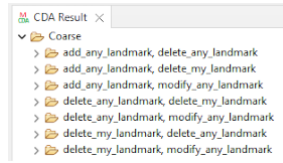
Figure 16: Henshin CPA result visualises conflicts and dependencies between user stories

tool [5]. In Section 4.5 we evaluate the results.

4.1 Requirements

In order to accomplish the analysis of redundancy in USs we try to address following requirements:

- "As a user, I want to perform syntactic analysis on user stories within a specified project backlog, so that I can identify and address redundancies effectively."

- "As a User, I want a report of user story pairs that contain identical syntactic clauses in both the action and benefit parts, so that I can modify them as needed."

- "As a User, I want to apply a filter to the redundancy report to exclusively display US-pairs, which have at least one clauses as *Targets* with the terms "Action" (as a verb) and "Entity" (as a noun), so that I can efficiently reduce the number of potential redundant pairs."

- As a user, I want to mark found redundancy clauses as Triggers with a hash symbol (#) and show those that have a redundancy in "Persona" (as a noun) and "Action" (as a verb), so that I can better see if the persona in is also recognised as a redundancy.

- As a user, I want to mark founded redundancy clauses as *Contains* with a hash symbol (#) and show those that contain two "Entity"(as a noun), so that I can better see whether the contained entity is also recognised as a redundancy.

- As a user, I would like to have a redundancy report that shows founded US texts in US-pairs and adds a hash symbol (#) at the beginning and end of the founded words as a marker, so that I can better see which words are redundant in US-pairs.

- As a user, I want to see how many redundancy clauses were founded in main and benefit parts of each US-pair, so that I can aggregate each founded redundancy US-pair for further statistical purposes on that basis.

- As a user, I want a table at the top of the redundancy report that lists the US-pairs and the number of redundancy clauses contained in each pair, so that I can quickly see all the US-pairs founded and the number of redundancy clauses.

---

[5]https://github.com/ace-design/nlp-stories

- As a user, I want to know whether the redundancy clauses belong to the main or benefit part of the US, so that I can use it for further statistical purposes.

- As a user, I want to know if a pair of US is partially or fully redundant, so that I can delete or modify them.

4.2 Design

This section describes the operational flow and architectural considerations that underpin the framework.

Design Overview

To address the redundancy detection requirements specified in 4.1, our system initiates with backlogs annotated with Conditional Random Fields (CRF) as the primary input. These annotated USs are used to generate graph transformation rules for each US in the backlog using Henshin tool.

Subsequently, these rules serve as inputs for the Henshin Conflict and Dependency Analysis tool (CDA) [27], which automates the identification of potentially redundant US pairs. The output from the CDA tool is then utilized to create a report that compiles information on these potentially redundant pairs. This report, in turn, becomes the input for the evaluation process, which outputs statistical data concerning redundancy among USs.

Figure 17 illustrates how each step in this sequence is interconnected, with the output of one step feeding directly into the next. This diagram effectively demonstrates the toolchain and process workflow, highlighting how each tool transforms artefacts and contributes to the overall objective of redundancy detection.
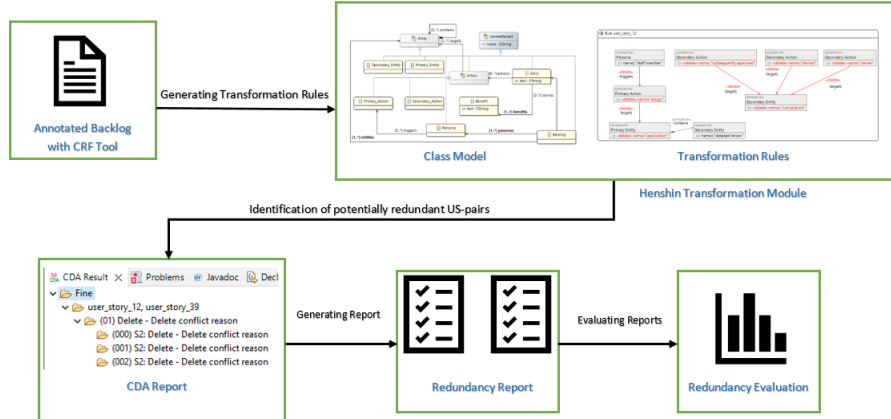

Figure 17: Step-by-step visualisation of the tool chain and its inputs and outputs

US labelling in JSON files As the annotated USs in the JSON files do not contain identifiers, a customised Python script, *nummerise_us.py*, is used to assign an unique identifier to each US, which is stored in a JSON object named "US_Nr" [6]. This addition improves the system's ability to distinguish and process individual USs within the analytical pipeline.

Creation of an Ecore meta model The creation of an Ecore meta-model is essential for rule generation in Henshin. The constructed meta-model, as shown in Figure 19, is based on the conceptual meta-model outlined in Figure 6 and ensures consistency with the JSON-structured data. This correspondence between the meta-model and the JSON representation underpins the rule generation and transformation processes of the system.

Rule creation process The transition from identified USs to implementable transformation rules within Henshin is facilitated by the package *org.eclipse.emf.henshin.model. compact*. The class *RuleCreator* within the package *org.henshin.backlog.code.rule* is important in this process as it uses different classes to instantiate transformation rules, nodes, edges and attributes. These components are annotated with actions such as *Delete*, *Create* or *Preserve*, which are necessary for the subsequent application of CDA to recognise redundant US-pairs. This methodological approach improves the system's ability to detect redundancies and thus optimises the efficiency of the residue analysis process.

Conflict and Dependency Analysis (CDA) After the creation of transformation rules and the creation of the corresponding Henshin file by the *RuleCreator* class, the system is able to perform a conflict and dependency analysis (CDA) to identify potential redundancy pairs.

Extraction of text reports The creation of a text report is a crucial step in consolidating the results of the conflict and dependency analysis (CDA), which aims to highlight essential information such as the identification of potentially redundant US-pairs, the enumeration of redundancy clauses, the categorisation of these clauses within the main or benefit parts of the USs and a tabulation of potentially redundant pairs alongside the total number of redundancy clauses.

Evaluating the reports After creating the reports, we are now able to evaluate the level of redundancy between US-pairs, aiming to determine the level of redundancy in user stories (USs) based on JSON reports.

The evaluation process involves a detailed examination of elements such as triggers, targets, and contains within the USs, comparing them to identify exact matches or significant overlaps that could indicate redundancy.

---

[6]https://github.com/amirrabieyannejad/USs_Annotation/tree/main/Script/numberise_us

Software Architecture

In this section, we present the basic structures of our workflow and the discipline of creating such structures. Each structure comprises software elements, relations among them, and properties of both.

- Annotated USs with CRF Tool: Mosser et al. integrated a dedicated Natural Language Processing (NLP) technique—Conditional Random Fields (CRF). CRFs represent a specialized class of Markov Random Fields, tailored for discriminative modelling in pattern recognition tasks where context plays a pivotal role[4]. As artefact we receive a graph-based model with JSON format, which represents the refined and annotated dataset for the recognition of *entities*, *actions*, *persons* and *benefits* of USs [30].

- Eclipse as IDE[7]: Eclipse is an integrated development environment (IDE) used in computer programming. It contains a base work workspace and an extensible plug-in system for customizing the environment. We chose this IDE because it offers the Henshin tool specifically for model-based development.

- Eclipse Modelling Project[8]: It focuses on the evolution and promotion of model-based development technologies within the Eclipse community by providing a unified set of modelling frameworks, tooling, and standards implementations.

- Eclipse Modelling Framework (EMF)[9]: The EMF project is a modeling framework and code generation facility for building tools and other applications based on a structured data model. From a model specification described in XMI, EMF provides tools and runtime support to produce a set of Java classes for the model, along with a set of adapter classes that enable viewing and command-based editing of the model, and a basic editor.

- Henshin[10]: Henshin is an in-place model transformation language for the Eclipse Modelling Framework (EMF). It supports direct transformations of EMF model instances (endogenous transformations), as well as generating instances of a target language from given instances of a source language (exogenous transformations).

  Because Henshin enables the specification of restrictions and conditions within rules, we use it to enforce and verify US requirements to ensure that the restrictions are met.

- Henshin API: It provides the specification and execution of transformation modules, units and rules. This API is beneficial due to the dynamic creation and modification of transformation modules as part of an automated tool chain[11].

---

[7]https://eclipseide.org/
[8]https://eclipse.dev/modeling/
[9]https://eclipse.dev/modeling/emf/
[10]https://wiki.eclipse.org/Henshin
[11]https://wiki.eclipse.org/Henshin/Compact_API

- RuleCreator Class: This class developed within the *org.henshin.backlog.code.rule* package, serves as an integral component of our software architecture, leveraging the Henshin API to automate the process of transformation rule creation based on dynamic JSON input. This class facilitates the programmable generation of graph transformation rules. It operates by reading JSON files, which specify USs and their associated actions, entities, and relationships, then systematically constructs corresponding Henshin modules, rules, nodes, and their attributes.

  the RuleCreator class effectively bridges the gap between high-level requirements specified in JSON and the low-level execution capabilities of the Henshin, ensuring that USs translated into executable transformation rules that reflect the specified software behaviour. This translation process is crucial for maintaining consistency and traceability across the software development lifecycle, from requirements specification to implementation.

- Henshin's CDA feature[12]: Henshin's conflict and dependency analysis (CDA) feature enables the detection of potential conflicts and dependencies of a set of rules.

  After the rules and the corresponding henshin file have been created by the Rule-Creator class as artefact, we are now able to pass them to the conflict and dependency analysis (CDA) to find potential redundancy pairs.

  Since the analysis of conflicts and dependencies related to the *attribute* is not yet considered in the CDA API, we decided to use the user interface (UI) of the CDA extension of Henshin, which supports analysis of conflict and dependencies of rules through the interactive use of CDA.

- ReportExtractor: The class from the *org.henshin.backlog.code.report* package is used in our software architecture to extract and format reports from the CDA-generated *Minimal-Model.ECore* that contain all information about redundant US-pairs such as redundant elements with their name and type. It uses classes from the *org.eclipse.emf.ecore* package to handle EMF (Eclipse Modelling Framework) resources that support the management and manipulation of *Minimal-Model.ECore* data in a structured format.

  In operation, the ReportExtractor class reads minimal-model.ecore files containing detailed information about redundant US-pairs. It then processes this data to generate reports in both text and JSON formats, aiding in the systematic analysis of potential redundancies within the USs. This is facilitated through methods that dynamically read and interpret the JSON data, extracting key information such as actions, entities, and their interactions.

- Evaluation class: The class is part of the *org.henshin.backlog.code.evaluation* package, serves a critical function within our software architecture, focusing on the analysis and determination of redundancy between USs based on specified criteria. This class utilizes JSON processing to assess the overlap and redundancy

---

[12]https://wiki.eclipse.org/Henshin/Conflict_and_Dependency_Analysis

between elements of USs, such as triggers, targets, and contains, which are vital for identifying potential redundancies in the USs.

Key functionalities of this class include reading and interpreting JSON data, where it extracts detailed elements related to USs and evaluates them against redundancy criteria defined in its methods. The class method *evaluateRedundancyCriteria* is particularly central, as it processes the JSON objects representing USs to determine if they are fully or partially redundant based on the presence and matching of various JSON array elements like triggers, targets, and contains in the main and benefit parts of the USs.

Figure 18 shows the architectural composition, highlighting the integral components and their user interface and artefacts.
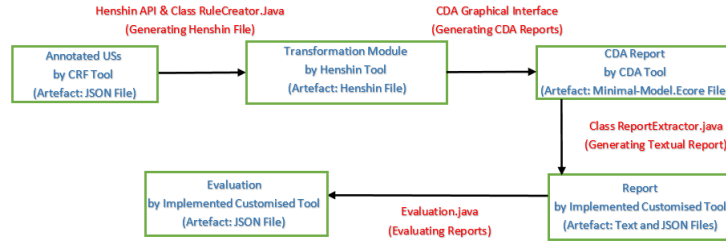

Figure 18: Design phases

Regarding redundancy, some definitions are clarified:

Definition 4.1 (Main and Benefit Parts in User Story). *A user story (US) is divided into two distinct parts that collectively describe what the user wants, why they want it, and how it will benefit them. These are:*

- *The main part, which is essential as it clearly and concisely summarises the persona, the intended functionality and the resources required to perform the action. This part usually has follow the format: "As a [persona], I want [what]".*

  Here, the persona helps contextualise the requirement by linking it to a user type, which promotes understanding and empathy. The intended functionality describes the action that the persona wants to perform or the function they need, providing a clear statement of the requirement.

  Specifying the resources required to perform the action helps with planning and resource allocation and ensures that the development team is aware of the tools, technologies and time required.

- *The benefit part, which formulates the potential benefit for the end user and typically begins with the phrase "so that". This part of the user story is essential for justifying the need for a feature by explaining the value or improvement it brings*

44

to the user's experience. It connects the functionality directly to user satisfaction, efficiency, or productivity gains, making it easier for the development team to prioritize features based on their impact.

It is worth noting that sometimes the benefit part may not exist in the structure of USs. In such cases, the main part can stand alone.

Definition 4.2 (**Redundancy**). *Redundancy refers to situations where some phrases of user stories (USs) duplicate or largely overlap with others. This is the case when only a part (main or benefit part) or entire USs are syntactically identical, meaning that some or all phrases and their orderings match perfectly between two USs.*

To comprehensively assess redundancy, it is important to consider not only the textual content, but also the functional and contextual relevance of each phrase within the USs. By analysing triggers, targets and contains relationships, we can uncover redundancies that may not be immediately apparent through a simple text comparison.

Definition 4.3 (**Clause**). *In this context, a clause is constructed as a pair of words, each pair being linked by a set of predefined relational constructs. These constructs, which are central to the contextual structuring of USs, are described below:*

- *Triggers: A relationship between a persona and an action that includes the invocation of the action by the persona within the narrative of a US.*

- *Targets: A relationship between an action and an entity means that the action influences the entity in question.*

- *Contains: relation between two entities, indicating that one entity contains another one, in which one entity merges into or is enclosed by another.*

Including the specified clauses in the analysis framework ensures that the words identified in potentially redundant US-pairs are linked by one of the relational constructs outlined: Triggers, Targets, or Contains.

This methodological approach supports the identification of redundancy not only on the basis of lexical similarity, but through the contextual relationships that determine the interactions within the US narratives.

Example 4.1. *When evaluating the textual similarities between two USs, consider the following pair:*
user_story_01: "As a Staff member, I want to Manage Ordinances, so that I can maintain accurate ordinance information in the System."
user_story_02: "As a Staff member, I want to Manage Affidavits, so that I can ensure compliance with the requirements prior to the hearing."

Upon comparing these two user stories, we identify shared phrases: "Staff member" and "Manage." These common elements indicate that both user stories involve the same persona ("Staff member") engaged in a similar type of action ("Manage"). However, the

focus of their management tasks differs significantly, with one managing "Ordinances" and the other "Affidavits."

This distinction in the resources being managed suggests that while there are textual overlaps and commonalities in persona and action, the user stories are not duplicates, as they address different aspects of the staff member's responsibilities and goals within the organization.

Definition 4.4 (**Full Redundancy**)**.** *A US-pair was categorised as "full redundancy" in the main/benefit part if all occurring clauses in the main/benefit part consisting of triggers (for the main part only), targets and contains were syntactically identical.*
This means that for a pair to fall into this category, every element including the wording, order and structure of the clauses relating to the triggers and targets must be identical. The assessment also extends to the "contains" elements, if these are present, to ensure that these also match perfectly without any deviation.

This level of redundancy in main part means that the redundant story provides no additional information and can be consolidated or eliminated without compromising the completeness or operational integrity of the system specifications.

On the other hand, having this level of redundancy in the benefit part means that USs achieve the same goal; they can be categorised in the same group. This means that if the end results or benefits described by the USs are identical, regardless of the different actions/entities or triggers that lead to these results, the USs effectively serve the same purpose within the system.

Example 4.2. *For example, the following US-pair is identified as "Full Redundancy" in the main part:*
user_story_01: "#g14# as a #publisher#, i want to #publish# a #dataset#, so that i can view just the dataset with a few people."
user_story_02: "#g14# as a #publisher#, i want to #publish# a #dataset#, so that i can share the dataset publicly with everyone."

Since triggers (persona: "Publisher" and action: "Publish") and targets (action: "Publish" and entity: "Dataset") are identical in the main parts of the USs, we therefore have "Full Redundancy" in the "Main" parts.

Example 4.3. *Following US-pair is identified as "Full Redundancy" in "Benefit" part:*
user_story_02: #g05# as a data publishing user, I want to be able to edit the model of data I have already imported, so that I can #fix# #bugs# or #make# #enhancements# in the #API# built for my #data#.
user_story_07: #g05# as a data publishing user, I want to be able to edit the data source of data I have already imported, so that i can #fix# #bugs# or #make# #enhancements# in the #API# built for my #data#.

As we can see, targets(["fix", "bugs"], ["make", "enhancements"]) and contains (["API","enhancements"],["data","API"]) in the benefit part are identical between USs, therefore, so we have "Full Redundancy" in the "Benefit" parts.

Definition 4.5 (**Partial Redundancy**). *Partial redundancy in USs occurs when only certain clauses, such as targets, have significant overlap but are not completely identical. This means that while there are shared aspects such as targets between USs, other clauses such as triggers or contains may not overlap, indicating an incomplete match. Such a scenario indicates that there is substantial, but not fully, a match between the USs.*

Example 4.4. *Following US-pair is identified as "Partial Redundancy" in main part:*
user_story_09: #g04# as a #user#, I want to be able to #view# a #map display# of the public recycling bins around my #area#.
user_story_10: #g04# as a #user#, I want to be able to #view# a #map display# of the special waste drop off sites around my #area#.
As we can see, there are some redundancy clauses such as Triggers (["user", "view"]), Targets (["view", "map display"]) and Contains(["map display", "area"]) between the USs. There are also clauses such as Contains (["public recycling bins"] vs. ["hazardous waste collection points"]) that are distinct elements justifying the maintaining of separate USs. We therefore assess this as "Partial redundancy" in the "Main part".

Example 4.5. *Following US-pair is identified as "Partial Redundancy" in benefit part:*
user_story_17: #g03# as a staff member, I want to manage approved proffers, so that I can #ensure# #compliance# with and satisfaction of the proffer in the future.
user_story_30: #g03# as a staff member, I want to manage affidavits, so that I can #ensure# #compliance# with the requirements prior to the hearing.
As we can see, there is a redundancy clause as targets (["ensure", "compliance"]) between the USs in the benefit part, but there is also a clause as contains (["proffer", "satisfaction"] vs ["requirements", "hearing"]), which leads us to evaluate this as "Partial Redundancy" in the "Benefit" part.

Structure of a US

The delineation and checking of redundancy clauses within USs requires a methodical approach, especially when distinguishing between the main and the benefit part of a US-pair. This distinction is crucial to ensure that redundancy identifications within one part are not mistakenly transferred to the other. Consequently, the analytical framework comprises three conditions, each of which specifies its own methodology for case processing:

- Presence of benefit in both USs of the redundancy-pair: if a benefit is identifiable in each US of the pair, a process of separation is used to split the main content from the benefit parts. After this separation, a targeted search for redundancy clauses is carried out only within the main part, whereby identified redundancies are annotated with a hash symbol (#). This process is repeated for the benefit parts to ensure a thorough check and marking of redundancies within each individual part.

47

- Exclusive presence of a benefit in one US of the redundancy-pair: In scenarios where only one US of the pair contains a benefit part, the analysis is limited to the main part of both USs. The aim remains the identification and annotation of redundancy clauses within this part. The lone benefit part remains in its original state and is excluded from the redundancy check.

- Absence of benefit parts in both USs of the pair: If neither of the two USs of the redundancy-pair contains a benefit part, the focus shifts completely to the main parts. The investigation is designed to highlight redundancy clauses within these parts, whereby the benefit parts are not taken into account due to their non-existence.

This structured and segmented approach ensures precise and efficient identification of redundancy clauses within the USs, optimising the clarity and effectiveness of textual report.

4.3 Implementation

In this section, we explain the objective and scope of the implementation, the system architecture, the functionality and the programming languages used.

The entire implementation is available in the GitHub repository [13].

Objective and Scope

The goal and scope of the work is divided into four phases. Firstly, converting the USs annotated by the CRF tool into graph transformation rules; secondly, using the CDA tool of the Henshin to automatically report redundancies between USs pairwise; thirdly, extracting important information from the CDA report into a text report; fifthly, evaluation of reports.

Methodology

This section explains and introduces tools that are required during the development process.

Following approach and tools are necessary in order to develop our workflow:

- Java as programming language[14]: Java is a widely used object-oriented programming language and software platform that is used to implement the Henshin and EMF APIs, which are critical to our approach to utilising them. Therefore, we use Java as our programming language.

- GitHub as version control[15]: GitHub is a developer platform that allows developers to create, store, manage and share their code. It uses Git software, providing the

---

[13]https://github.com/amirrabieyannejad/USs_Annotation/tree/main
[14]https://www.java.com/de/
[15]https://github.com/

distributed version control of Git plus access control, bug tracking, software feature requests, task management, continuous integration, and wikis for every project.

Implementation Phases

To provide a comprehensive overview of the implementation phases, this section explains each step of the process, from initial setup to final evaluation, using practical examples.

Step 1: Data Preparation

As primary input, we receive a graph-based model generated by the CRF tool, which represents the refined and annotated dataset for the recognition of *entities*, *actions*, *persons* and *benefits* of USs [30].

The datasets have the JSON format, the structure of which is very important in the Java classes *RuleCreator*, *ReportExtractor*, and *Evaluation*. Therefore, understanding the JSON format provided is needed for the further procedure.

Each JSON file for a backlog dataset contains a JSON-array in which each US entry is defined as a JSON-object. Listing 1 illustrates the format used for the US entry.

```
[
{
"Text": "...",
"Persona": [ "..." ],
"Action": {
        "Primary Action": [ "...", ... ],
        "Secondary Action": ["..."], ... },
"Entity": {
        "Primary Entity": [ "...", ... ],
        "Secondary Entity": [ "...", ...] },
"Benefit": "...",
"Triggers": [ [ "...", "..." ] ],
"Targets": [ [ "...", "..." ], ... ],
"Contains": [ [ "...", "..." ], ... ]
},
...
]
```

Listing 1: The JSON format of each US entry in JSON file

Mooser et al. have linked each *Persona* to each *Primary Action* as *Trigger* relationships, each *Primary Actions* to each *Primary Entity* as *Target* relationships and each *Primary/Secondary Entity* to each *Primary/Secondary Entity* implying a *Contains* relationship[4].

To interact with the entries in JSON file, we need to distinguish between the entries that are defined as JSON-objects, such as: Text, Action, Entity, Benefit, US_Nr and the entries that are defined as a JSON-array, such as: Persona, Primary/Secondary Action, Primary/Secondary Entity, Triggers, Targets, Contains.

Identifying USs in JSON-File

Annotated USs in each JSON file have no identifier. To distinguish USs, we use a Python script called *nummerise_us.py* [16], which receives JSON files as input and adds a JSON object named "US_Nr" with an identifier as value (e.g. user_story_01) to each US and returns the JSON files as output. Listing 2 illustrates the added JSON object "US_Nr" and its value in the JSON file.

```
[
{
"Text": "...",
"Persona": [ "..." ],
"Action": {
        "Primary Action": [ "...", ... ],
        "Secondary Action": ["..."], ... },
"Entity": {
        "Primary Entity": [ "...", ... ],
        "Secondary Entity": [ "...", ...] },
"Benefit": "...",
"Triggers": [ [ "...", "..." ] ],
"Targets": [ [ "...", "..." ], ... ],
"Contains": [ [ "...", "..." ], ... ]
},
"US_Nr" : "user_story_01"
,
...
]
```

Listing 2: The JSON format with the additional JSON object "US_Nr" and its value

Step 2: Creation of Rules

Step 2 of the implementation involves a central process in which the US data structured in JSON files is transformed into transformation rules using the Henshin API. This involves the creation of an Ecore meta-model that represents the structure of the data we are working with, followed by the generation of Henshin transformation rules using RuleCreator class.

Creating Ecore Meta-Model

To be able to create rules in Henshin, an Ecore (meta)-model should be available. Ecore is the core (meta)-model at the heart of the EMF (Eclipse Modelling Framework). It enables the formulation of other models by utilising its constructs.

---

[16]https://github.com/amirrabieyannejad/USs_Annotation/tree/main/Script/numberise_us

Accordingly, we create an Ecore meta-model as shown in Figure 19, which is inspired by the meta-model shown in Figure 6 and corresponds to the JSON-objects in the JSON-file as follows:

- *Persona* as a class in the meta-model corresponds to the JSON-object "Persona" in the JSON-file.

- *Entity* as an abstract class, from which *Primary/Secondary Entity* inherits as a class in the meta-model, corresponds to the JSON-object "Entity", which contains two JSON-arrays, namely "Secondary/Primary Entity" in the JSON-file.

- *Action* as an abstract class and *Primary/Secondary Action* as an inherited class in the meta-model correspond to the JSON-object "Action", which contains two JSON-arrays, namely "Secondary/Primary Action" in the JSON-file.

- *Benefit* as a class in the meta-model, which also has an attribute called "text" that corresponds to the JSON-object "Benefit" in the JSON-file.

- *Story* as a class in the meta-model that contains text from US, which also has an attribute called "text" that corresponds to the JSON-object "Text" in the JSON-file.

- Abstract class *NamedElement* has attribute *name*, which Primary/Secondary Action/Entity inherit from it, which corresponds to the value of *Primary/Secondary Action/Entity* in JSON-file.

- *Edge* with the name *triggers* between Persona and Primary Action in the meta-model, which corresponds to the JSON-array "Triggers", where each JSON-array in it contains a pair, the first element corresponding to the *Persona* and the second to the *Primary Action*.

- *Edge* named *targets* between Primary/Secondary Action and Primary/Secondary Entity in the meta-model, which corresponds to the JSON-array "Targets", where each JSON-array has a pair, the first element corresponding to "Primary/Secondary Action" and the second element corresponding to "Primary/Secondary Entity".

- *Edge* named *contains* between Primary/Secondary Entity and itself in the meta-model, which corresponds to the JSON-array "Contains", where each JSON-array in it has a pair where the first element corresponds to "Primary/Secondary Entity" and the second element corresponds to "Primary/Secondary Entity".
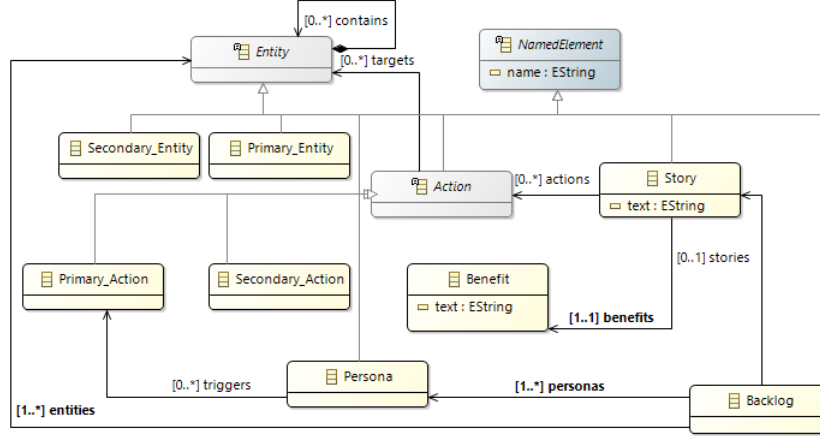
Figure 19: Ecore meta-model inspired by Mosser et al. [30]

Creating Rules

With the identified USs in the the JSON-file, we generate rules with the Henshin package *org.eclipse.emf.henshin.model.compact*, which is responsible for the creation of *transformation rules* and their *classes*, *attributes*, *edges* and annotates them with *<Delete>*, *<Create>*or*<Preserve>*, which are vital for the CDA tool to recognise the redundant pairs.

To generating rules we create a package named *org.henshin.backlog.code.rule* and specially the class *RuleCreator* which used following classes[17]:

- *org.eclipse.emf.henshin.model.compact.CModule*: CModule class can import elements from an Ecore file to use them in the transformation process responsible for linking the Ecore meta-model to the Henshin-file to be created.

- *org.eclipse.emf.henshin.model.compact.CRule*: Once we have a CModule, we can specify transformation rules with the CRule class and create them.

- *org.eclipse.emf.henshin.model.compact.CNode*: Now that we have a transformation rule, we want to fill this rule with nodes, edges and attributes. To create a node within a transformation rule, we need the CRule class. To create an edge we need to reference two nodes together. The default action when specifying a node, edge or an attributes is the *<preserve>*action. We can also specify a different action when we create a node or an edge, for example *<delete>*or*<create>*.

- org.henshin.backlog.code.rule.RuleCreator: We implement *RuleCreator* class that creates a rule with annotated nodes, edges and attributes based on a JSON-file as input and a Henshin-file containing all rules as output, where each rule and its members(nodes, attributes, edges) correspond to the individual US and their JSON-objects/arrays in the JSON-file.

---

[17]https://wiki.eclipse.org/Henshin/Compact_API

The most important design decision of this class is the way nodes, attributes and edges are annotated in order to be able to apply conflict and dependency analysis (CDA) for rules stored as a Henshin file.

We decided to annotate the "name" attribute of all Primary/Secondary Actions/Entities and their associated edges including "targets", "triggers" and "contains" as <delete>action.

The main goal is to increase the probability of identifying US-pairs characterised by matching names of *action* nodes and *entity* nodes in conjunction with an edge called *targets*. This congruence serves as a basic criterion for identifying potentially redundant US-pairs and simplifies the process of redundancy detection in the context of US analysis.

Methods of the RuleCreator Class

In this section, the methods of the RuleCreator class is described as follows:

1. assingCmodule: This method assign a CModule to a Ecore meta-model. It creates a new CModule object with the provided Henshin-file name, adds imports from the Ecore file, and returns the module.

2. processJsonFile: It takes parsed JSON array as input and processes their attributes, such as persona, actions, entities, text and their edges, such as targets, triggers. Corresponding elements are created as output in a the Henshin transformation module (CModule).

3. processRule: It takes the "US_Nr" JSON-object as input and creates a new CRule with the name of unique US identifier in the CModule.

4. processPersona: It receives as input the persona extracted from the JSON data and the associated CRule to create a new CNode representing the persona within the provided CRule and adds the attribute "name" with persona as value. Finally, the created CNode representing the persona is returned.

5. processText: It receives as input US text extracted from JSON data and the associated CRule to create a new CNode representing the text within the provided CRule and adds the attribute "text" with US text as value. Finally, the created CNode representing the US text is returned.

6. processActions: The processActions method is responsible for creating CNode objects that represent actions within the CModule. As parameters, it receives the JSON-object of the actions, the CNode-object representing the persona associated with the actions and the unique identifier of the US. Since the edge triggers only refer to the persona and the primary action, a new CNode is created for each primary action that represents the primary action within the provided CRule, an attribute "name" is added and an edge is created from the persona node to the

primary action with the label "triggers". For each secondary action, a new CNode is created to represent the action within the provided CRule and an attribute "name" is added to the action node.

7. checkEntityIsTarget: It receives the name of the entity and the JSON-array with information about target edges. The method iterates through the JSON-array targets, which contains arrays that represent targets edges between actions and entities. It compares the targets entity with the specified entity. If there is a match, it returns true to indicate that the entity is a target.

8. processEntities: It receives as parameters the JSON-object with information about the entities, the CRule object representing the US to which the entities belong and the JSON-array with information about the targets associated with the entities. The method checks whether primary/secondary entities are present, then creates a CNode for each primary/secondary entity and checks whether the entity is present in the target array. If this is the case, its attribute "name" is annotated for deletion. This method is used within the processContainsEdges method to determine whether an entity involved in a contains relationship is also a target of another entity.

9. processContainsEdges: It receives the JSON-object to be processed, the JSON array with information about contains/target edges and the US identifier as parameters. It first checks whether both entities belong to contains edges. If both entities exist, an edge is created between them in CRule with the name "contains". If one of the entities is a target of another entity (as specified in the targets array), the edge is annotated for deletion. If none of the entities is a target, the edge is annotated as "preserve".

Figure 20 is a class diagram that illustrates the attributes, operations and relationships related to the RuleCreator class. To illustrate this step, three example is given in which the RuleCreator class is applied to a dataset (G03) with three collected USs and the resulting artefact is presented.

Example 4.6. *Listing 3 shows the JSON format in relation to user_story_12 and Figure 21 shows the application of the RuleCreator class in this US, which is a transformation rule where the targets and the associated contains relationships are annotated as a <Delete>action and the rest of the nodes and edges are annotated as a <Preserve>action.*

Text of US is: user_story_12: "#G03# As a Staff member, I want to Assign an Application for Detailed Review, so that I can review the for compliance and subsequently approved or denied."
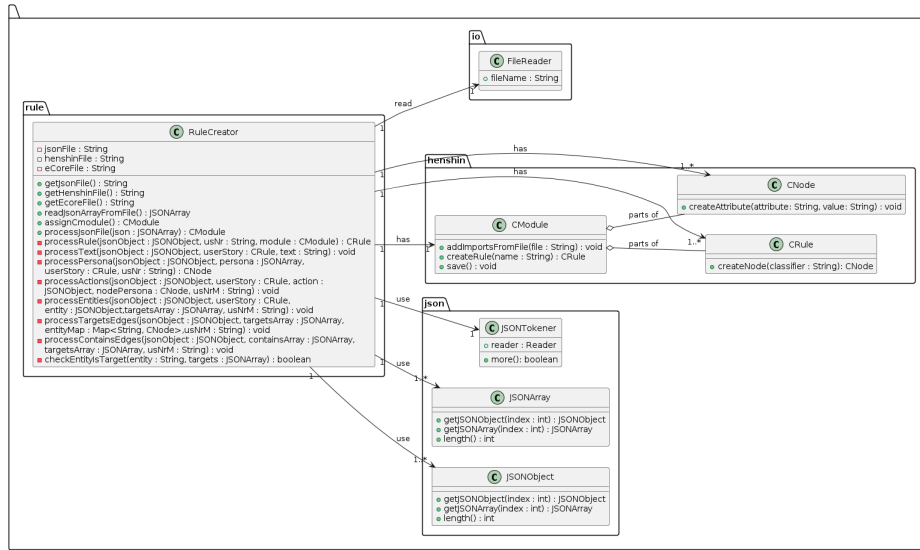
Figure 20: Class diagram of the RuleCreator class and its relationships

```json
{
    "Text": "#G03# As a Staff member, I want to Assign an Application
    for Detailed Review, so that I can review the for compliance and
    subsequently approved or denied.",
    "Persona": ["Staff member"],
    "Action": {"Primary Action": ["Assign"],
    "Secondary Action": ["subsequently approved","denied","review"]},
    "Entity": {"Primary Entity": ["Application"],
    "Secondary Entity": ["compliance","Detailed Review"]},
    "Benefit": "I can review the for compliance and subsequently
    approved or denied",
    "Triggers": [["Staff member","Assign"]],
    "Targets": [["Assign","Application"],["review","compliance"],
    ["subsequently approved","compliance"],["denied","compliance"]],
    "Contains": [["Application","Detailed Review"]],
    "US_Nr": "user_story_12"
},
```

Listing 3: JSON entities Related to user_story_12

As we can see, the "targets" edges and their direct relationships ("triggers" and "contain", if any) are also annotated as <delete>, which is very important to find redundant elements with the CDA tool.

Example 4.7. *Listing 4 shows the JSON entities related to user_story_39 and Figure 22 shows transformation rule generated by RuleCreator class.*

Text of US is: user_story_39: "#G03# As a Plan Review Staff member, I want to Review Plans, so that I can review them for compliance and either approve, or fail or
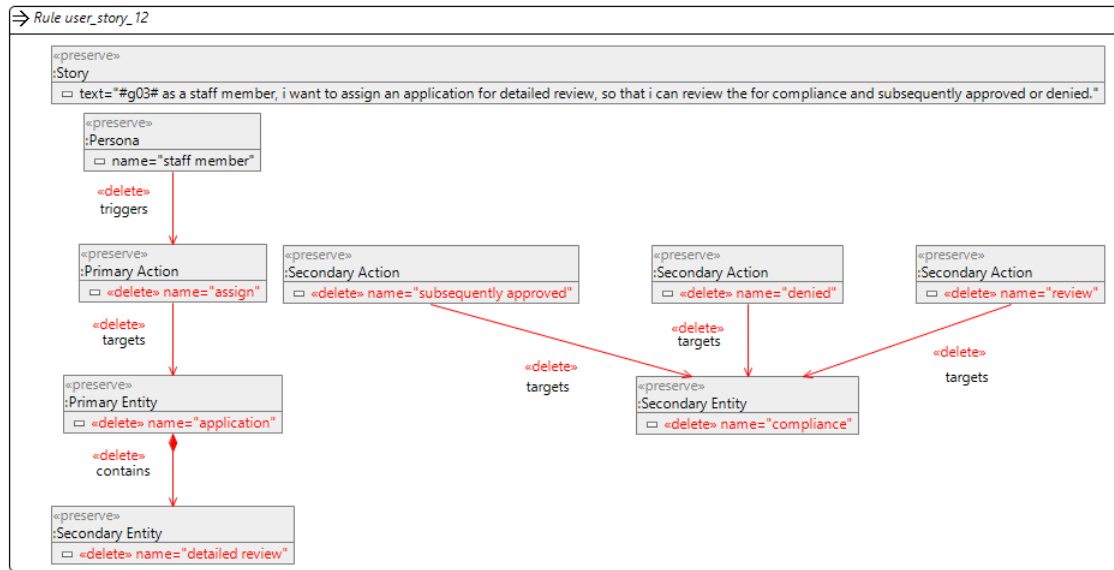
Figure 21: Generated transformation rule related to user_story_12 using RuleCreator class

deny the plans and record any conditions, clearances, or corrections needed from the Applicant."

```
{
  "Text": "#G03# As a Plan Review Staff member, I want to Review
  Plans, so that I can review them for compliance and either approve,
  or fail or deny the plans and record any conditions, clearances, or
  corrections needed from the Applicant.",
  "Persona": ["Plan Review Staff member"],
  "Action": {"Primary Action": ["Review"],
  "Secondary Action": ["deny","review","approve","fail","record"]},
  "Entity": {"Primary Entity": ["Plans"],
   "Secondary Entity":["compliance","plans","clearances","Applicant",
    "corrections","any conditions"]},
  "Benefit": "so that I can review them for compliance and either
  approve, or fail or deny the plans and record any conditions,
  clearances, or corrections needed from the Applicant",
  "Triggers": [["Plan Review Staff member","Review"]],    "Targets":[[
  "Review","Plans"],["review","compliance"],["approve","plans"],["fail"
  ,"plans"],["deny","plans"],["record","clearances"],["record",
  "corrections"],["record","any conditions"]],
  "Contains": [["Applicant","plans"],
   ["plans","clearances"],["plans","corrections"],["plans","any
    conditions"]
  ],
  "US_Nr": "user_story_39"
  },
```

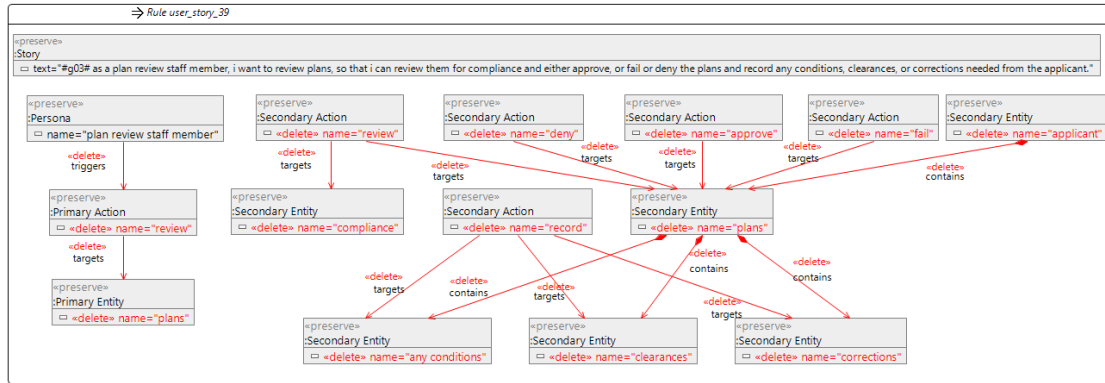Listing 4: JSON entities Related to user_story_39

56

Figure 22: Generated transformation rule related to user_story_39 using RuleCreator class

Attribute "Plan", as we can see, it appears both in the main part as a primary entity and in the benefit part as a secondary entity, forming the various relationships as targets and contains.

Example 4.8. *Listing 5 shows the JSON entities related to user_story_51 and Figure 23 shows transformation rule generated by RuleCreator class.*

Text of US is: user_story_51: "#G03# As an Enforcement Staff member, I want to Issue a Notice of Violation, so that I can provide formal communication to the responsible party."

```
{
    "Text": "#G03# As an Enforcement Staff member, I want to Issue a
    Notice of Violation, so that I can provide formal communication to
    the responsible party.",
    "Persona": ["Enforcement Staff member"],
    "Action": {"Primary Action": ["Issue"],
      "Secondary Action": ["provide"]},
    "Entity": {"Primary Entity": ["Notice of Violation"],"Secondary
    Entity": ["formal communication","responsible party"]},
    "Benefit": "I can provide formal communication to the responsible
    party",
    "Triggers": [["Enforcement Staff member","Issue"]],"Targets": [[
    "Issue","Notice of Violation"],["provide","formal communication"]],
    "Contains": [],
    "US_Nr": "user_story_51"
}
```

Listing 5: JSON entities Related to user_story_51

Last but not least, we have determined that the secondary entity "responsible party" has neither a target nor a contains relationship. Therefore, it is not annotated as
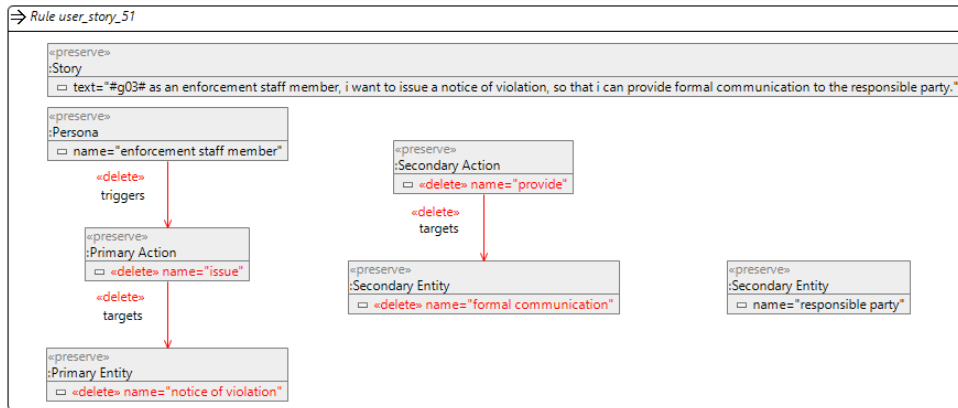
Figure 23: Generated transformation rule related to user_story_51 using RuleCreator class

<delete>, but as <preserve>. This is due to the fact that some phrases have been identified as entities in the CRF tool, but their relationship is not annotated at all, which is problematic for analysing redundancy.

Step 3: Conflicting and Dependency Analysis

After the rules and the corresponding henshin file have been created by the RuleCreator class, we are now able to pass them to the conflict and dependency analysis (CDA) to find potential redundancy pairs.

Since the analysis of conflicts and dependencies related to the *attribute* is not yet considered in the CDA API [18], we decided to use the user interface (UI) of the CDA extension of Henshin, which supports analysis of conflict and dependencies of rules through the interactive use of CDA.

To apply CDA to Henshin files, we just need to right-click on the Henshin file and select *Henshin>conflict and Dependency Analysis* from the context menu as shown in Figure 24. A user interface then appears, prompting to select the rule sets to be analysed
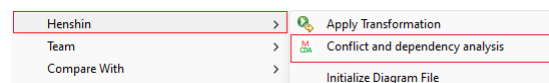

Figure 24: Applying CDA to the selected Henshin file

and the type of analysis. We then select as "*First*" and "*Second Rules*", all rules related to USs. Additionally, as the type of analysis we select "*conflicts*" as illustrated in Figure 25.

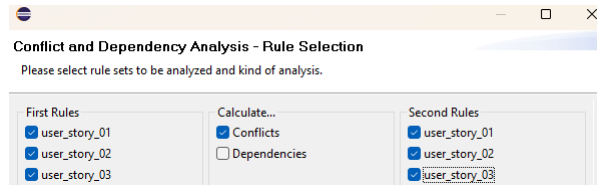[18]https://wiki.eclipse.org/Henshin/conflict_and_Dependency_Analysis

Figure 25: CDA user interface: Selection of rules and type of analysis

On the next page of the CDA UI shown in Figure 26, we specify the depth of analysis that we use with "*Fine granularity*" when selecting "*Create a complete result table*" and "*Create an abstract result table*".

Fine granularity provides a detailed examination of each conflicting rule by listing all conflict reasons. Unlike coarse granularity, which focuses only on minimal conflict reasons, fine granularity includes both minimal and more general conflict reasons. The binary granularity, where simple conflicting rule pairs are listed, may be too simple for complex systems where understanding the nature of the conflict is essential for the solution. We choose "Fine granularity" as the depth of analysis due to the fact that it shows all conflict reasons for each conflicting rule pair. This allows for a deeper understanding of how different model fragments contribute to conflicts.

A conflict reason is a model fragment whose presence leads to a conflict. General conflict reasons result from different combinations of minimal conflict reasons[36].
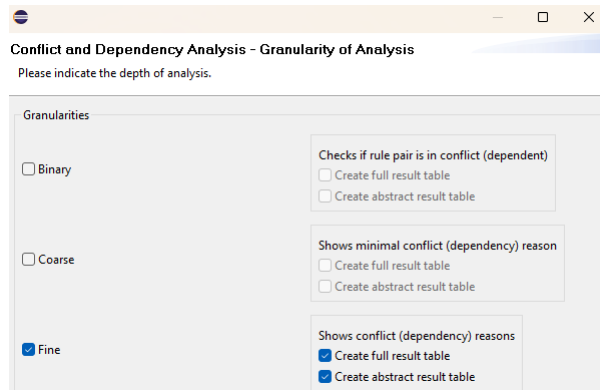

Figure 26: CDA user interface: Selection of report granularity

During the execution of the CDA analysis, the rule pairs is analysed and a conflict analysis is performed. Once the calculation is complete, the results are listed in the "CDA" -> "Result window", as shown in Figure 27. The top entry shows the granularity, which in our case is "Fine". These entries contain the rule pairs that conflict with each other. Each rule pair contains a number of conflict reasons.

Figure 28 shows how the data is saved in the project tree view. The results directory is created in the directory containing the Henshin that was used for the analyses. The new folder name is the date and time at which the analysis was performed. In contrast
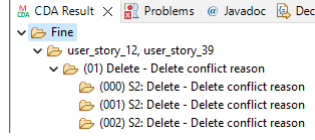
Figure 27: CDA report with fine granularity

to the "*CDA/Results*" view, this folder contains all conflict reasons and atoms together in a rule pair directory.
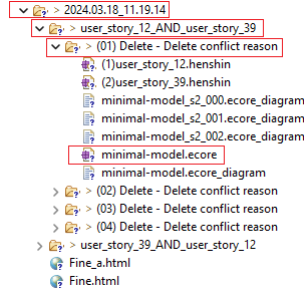
Figure 28: Saving CDA results data in the project structure view

For each conflict reason, there is a "*minimal-model.ecore*" file, that contains packages in which various conflict elements such as "attributes" and "references" (edges) are mapped together and displayed in different packages.

Figure 29 shows the representation of the conflicting attributes and references. An attribute has the property of changing the value and is represented by an arrow "->". The attribute from the first rule is separated from the second rule by an underscore, just as with the nodes.
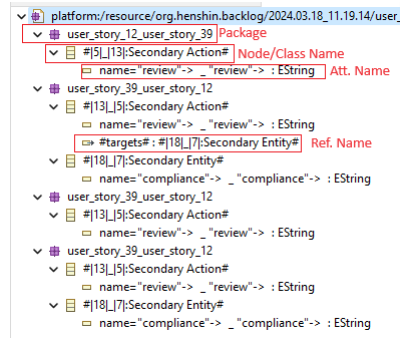
Figure 29: Representation of redundant attributes and references in *minimal-model.ecore* file

Example 4.9. *To illustrate this step, we also pass the transformation rules created in*

60

step 2, which reflect three USs (user_story_12/39/51), to the CDA tool and selecting "conflicts" as conflict type to calculate with "fine granularity" as the depth of analysis. Once the calculation is complete, the results are listed in the "CDA" ->"Results Window" as shown in Figure30. Figure31 shows how the data is saved in the project's tree view, which contains all conflict reasons and atoms together in a rule pair directory. As we can see, the CDA tool has only found redundancy between user_story_12
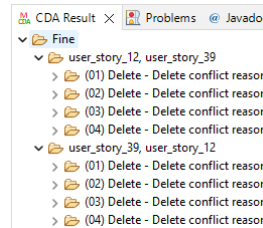

Figure 30: CDA report in relation to three transformation rules with "conflicts" as the type to be calculated and "fine granularity" as the depth of the analysis
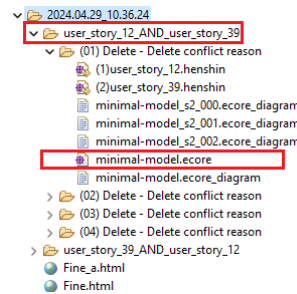
Figure 31: Saving CDA results data in the project's tree view

and user_story_39. This is because there are no redundant clauses between two USs (user_story_12 and user_story_39) and user_story_51.

Regarding the redundant elements specifically, we can refer to the created file "minimal-model.ecore", which is located in the tree view of the project under each conflict reason. Figure 31 show the minimal-model.ecore file related to user_story_12 and user_story_39.

The file Minimal-model.ecore, which refers to user_story_12 and user_story_39, is divided into packages, with each package containing different matches of redundant elements. If there is a redundancy between two elements, this is explicitly indicated by a hash symbol (#). Figure 32 illustrates the redundant elements found in each package.

For example, the attribute "name" with the value "review" in "Secondary Action" and the attribute "targets" with the value "compliance" in "Secondary Entity" are labelled as redundant (both with a hash symbol). The reference to "targets" is also marked as redundant (with a hash symbol), which means that the reference (targets) from "Sec-
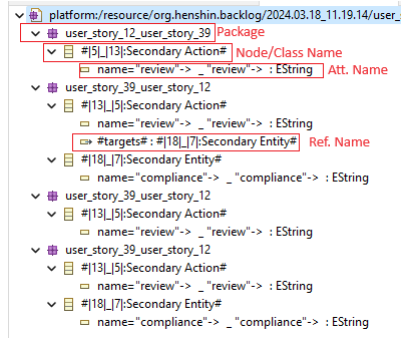
Figure 32: Representation of the redundant elements in each package within *minimal-model.ecore* file

ondary Action" to "Secondary Entity" is also redundant between USs, which is a very important criterion for finding redundancy correctly.

Step 4: Report Extraction

To create a lightweight report for the group or individual in question, we need to extract the key information from the CDA report, e.g. redundancy US-pair, redundancy clauses, count of redundancy clauses in each part of the US (main or benefit part), and create a report as a text file with the following information:

- A table of potential redundant pairs with the number of total redundancy clauses.

- Founded potential redundant US-pairs.

- Redundancy words and clauses of founded US-pairs. Clauses consisting of two words that have one of the relationships triggers, targets or contains.

- Text of US-pairs whose redundancy words are marked with a hash symbol (#).

- Parts of the sentence in which words and clauses are found.

Example 4.10. *After the CDA directory for user_story_12 and user_story_39 is created by CDA tool graphic interface (UI), we pass the location of directory into ReportExtractor class and in order to extracting the important information and save it into the texual report as well as JSON report. Listing 6 illustrates the example of the textual report. In this case, the report only contains one US-pair.*

```
* Table of potential redundancies between user stories and the number of their overlapping
elements

            us_12  us_39 ─────────────  USs identifiers
    us_12   0       1    ─────────────  Total redundant clauses
    us_39   1       0

-----------[Potential Redundant User Stories found]-----------
{user_story_12_AND_user_story_39}
                                                                USs Texts
Redundant clauses within user stories are:
* Secondary Action: review
* Secondary Entity: compliance                    Redundant Clauses
* Targets: Link from "review" to "compliance" is found.

user_story_39: #g03# as a plan review staff member, i want  to review plans, so that i can
#review# them for #compliance# and either approve, or fail or deny the plans and record any
conditions, clearances, or corrections needed from the applicant.

user_story_12: #g03# as a staff member, i want to assign an application for detailed review,
so that i can #review# the for #compliance# and subsequently approved or denied.

The following sentence parts are candidates for possible redundancies between user stories:
                                                                Parts of sentences
user_story_12:  so that i can #review# the for #compliance# and subsequently approved or
denied.
user_story_39:  so that i can #review# them for #compliance# and either approve, or fail or
deny the plans and record any conditions, clearances, or corrections needed from the
applicant.
```

Listing 6: Example of generated textual report for one US-pair

As we can see, the text report consists of a 2x2 table whose first column and first row are US identifiers, and the numbers inside the table are the total number (benefit part + main part) of redundancy elements between two USs; secondly, the redundant clauses related to the redundant US-pair are listed; thirdly, the text of US whose redundant phrases are marked with a hash symbol; finally, the part of the clauses in which redundant elements occur is displayed.

For further evaluation purposes and easy export of the report to another platform such as Excel, a JSON report is created that collects the information about redundant US-pairs separately in a JSON object with the following entries:

- Potential Redundant User Stories: which has stored the US-pair identifier(e.g. "user_story_12_AND_user_story_39").

- Status: consisting of "Main/Beneift Part Redundancy Clauses" and "Total Redundancy Clauses", which store the count of redundancy clauses in the main and benefit part as well as in the total part of the US.

- Entity: which can consist of a "Secondary/Primary Entity" and stores the founded redundant entities.

- Common Targets/Contains: which consists of the "Main Part" and "Benefit Part" entries and only stores the targets/contains relationships that are common between

63

the USs in a particular part of the USs. For example, if there are common redundant targets in the main part of the USs, these are included in the "Main Part" entity of the "Common Targets".

- Text: consisting of two entries, namely "First UserStory" and "Second UserStory", in which the text of the US-pair whose hash symbol has already been applied in redundant phrases is stored.

- Project Number: stores the number of the Project(e.g. "G03").

- Part of Sentence: consists of the entries "First UserStory" and "Second UserStory", in which the part of the US sentences containing redundant clauses is stored.

- All Targets/Contains: which consists of the "Main Part" and "Benefit Part" entries and stores the whole targets/contains relationships that are occurred in the particular part of the USs.

Example 4.11. *Listing 7 illustrates the example of the JSON report regarding user_story_12 and user_story_39.*

```
[
    {
        "Potential Redundant User Stories": "user_story_12_AND_user_story_39",

        "Status": {"Main Part Redundancy Clauses": 0,"Total Redundancy Clauses": 1,
        "Benefit Part Redundancy Clause": 1,},

        "Entity": {"Secondary Entity": ["compliance"]},

        "Common Targets": {"Main Part": [],"Benefit Part": [["review","compliance"]]},

        "Action": {"Secondary Action": ["review"]},

        "Common Contains": {"Main Part": [],"Benefit Part": []},"Targets": [["review",
        "compliance"]],

        "Triggers": [],

        "Text": {"Second UserStory": "user_story_39: #g03# as a plan review staff member,
        i want to review plans, so that i can #review# them for #compliance# and either
        approve, or fail or deny the plans and record any conditions, clearances, or
        corrections needed from the applicant.",
            "First UserStory": "user_story_12: #g03# as a staff member, i want to assign
            an application for detailed review, so that i can #review# the for
            #compliance# and subsequently approved or denied."},

        "Project Number": "g03",

        "Part of Sentence": {"Second UserStory": [" so that i can #review# them for
        #compliance# and either approve, or fail or deny the plans and record any
        conditions, clearances, or corrections needed from the applicant."],
            "First UserStory": [" so that i can #review# the for #compliance# and
            subsequently approved or denied."]},

        "All Contains": {"user_story_39": {"Main Part": [],"Benefit Part": [["applicant",
        "plans"],["plans","clearances"],["plans","corrections"],["plans","any conditions"
        ]]},
            "user_story_12": {"Main Part": [["application","detailed review"]],"Benefit
            Part": []}},

        "All Targets": {"user_story_39": {"Main Part": [["review","plans"]],"Benefit Part"
        : [["review","plans"],["review","compliance"],["approve","plans"],["fail","plans"
        ],["deny","plans"],["record","clearances"],["record","corrections"],["record",
        "any conditions"]]},
        "user_story_12": {"Main Part": [["assign","application"]],"Benefit Part": [[
        "review","compliance"],["subsequently approved","compliance"],["denied",
        "compliance"]]}}
    },

]
```

Listing 7: Example of generated JSON report for one US-pair

In order to extracting a textual report associated with a specific backlog, we implement a class called *ReportExtractor* within the package *org.henshin.backlog.code.report*, which include the following classes form the package *org.eclipse.emf.ecore*[19]. These classes are important for reading the content of minimal-model.ecore:

- org.eclipse.emf.ecore.resource.Resource: A resource of an appropriate type is created by a resource factory; a resource set indirectly creates a resource using such a factory. A resource is typically contained by a resource set, along with related resources.

---

[19]https://download.eclipse.org/modeling/emf/emf/javadoc/2.7.0/org/eclipse/emf/ecore/

- org.eclipse.emf.ecore.resource.ResourceSet: A resource set manages a collection of related resources and notifies about changes to this collection. It provides a tree of content. A collection of adapter factories supports the search for an adapter via a registered adapter factory.

- org.eclipse.emf.ecore.EObject: EObject is the root of all modelled objects, therefore all method names start with "E" to distinguish the EMF methods from the client methods. It provides support for the behaviour and functions that are common to all modelled objects.

- org.eclipse.emf.ecore.EPackage: A representation of the model object "EPackage".

- org.eclipse.emf.ecore.EClassifier: A representation of the model object "EClassifier".

- org.eclipse.emf.ecore.EClass: A representation of the model object "EClass".

- org.eclipse.emf.ecore.EAttribute: A representation of the model object "EAttribute".

- org.eclipse.emf.ecore.EReference: A representation of the model object "EReference".

Storing Redundancy Items into RedundancyItems Class

To save the redundant elements found by the CDA tool, we implement classes to save the elements according to their type.

Figure 33 is a class diagram that illustrates the attributes, operations and relationships related to the ReundancyItems class to store the redundancy elements provided by the CDA tool. The following classes were created to represent the extracted model object accordingly. All these classes are extensions of the class *RedundancyItems*, which contains all extracted model object from "minimal-model.ecore" such as "EClass", "EAttribute" or "EReference":

- PrimaryAction/SecondaryAction: Which has only saved the EClass specified by "Primary/Secondary Action" and the EAttribute model object with the methods *getType* to retrieve the saved EClass and *getName* to retrieve the saved EAttribute.

- PrimaryEntity/SecondaryEntity: Which has only saved the EClass specified by "Primary/Secondary Entity" and the EAttribute model object with the methods *getType* to retrieve the saved EClass and *getName* to retrieve the saved EAttribute.

- Targets: The EClass specified by "Primary/Secondary Action" as *outgoing edge* and an EAttribute model object as *incoming edge* with "Primary/Secondary Entity". The method *getType* retrieve the stored EClass and method *getName* retrieve the stored EAttribute.

66

Figure 33: Class diagram for the RedundancyItems class and its relationship

- Contains: The EClass specified by "Primary/Secondary Entity" as *outgoing edge* and an EAttribute model object as *incoming edge* with "Primary/Secondary Entity". The method *getType* retrieve the stored EClass and method *getName* retrieve the stored EAttribute.

- Triggers: The EClass specified by "Persona" as *outgoing edge* and an EAttribute model object as *incoming edge* with "Primary Action". The method *getType* retrieve the stored EClass and method *getName* retrieve the stored EAttribute.

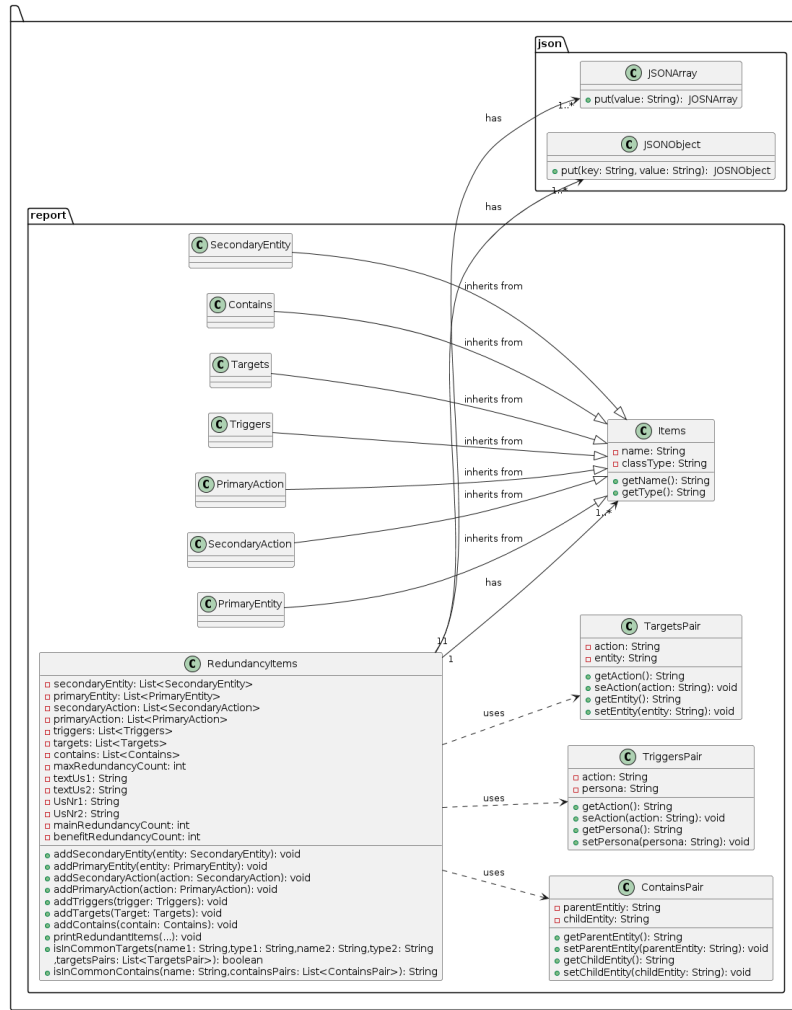- RedundantPair: Stores the identifier of the two USs that were founded as a redundant pair. It also stored the total count of redundancy clauses within the US-pair.

- TargetsPair: Stores effective value of "*Primary/Secondary Action*" and "*Primary/Secondary Entity*" in action and entity fields accordingly.

- ContainsPair: Stores effective value of "*Primary/Secondary Entity*" as "*parent/child entity*" due to the fact that parent entity is a containment of child entity.

- TriggersPair: Stores effective value of "*Persona*" as a persona and "*Primary Action*" as an action.

The class RedundancyItems contains following methods which are important for class ReportExtractor:

- isInCommonContains: This method is designed to determine whether a given entity (specified by its name) is part of a redundant pair listed in the "Contains" array of related USs stored in a JSON file.

  As input, it receives the name of the entity for which we want to check whether it exists in a redundant pair. In addition, a list of redundant pair objects that represent pairs of entities where one entity contains the other. If there is a match with the parent entity, it returns the child entity and vice versa.

- isInCommonTargets: This method is responsible for determining whether a particular action/entity is part of a redundant pair listed in the "Targets" array of related USs stored in a JSON file.

  As input, it receives the name and type of the first element, which is an action, and the name and type of the second element, which is an entity. It also receives a list of TargetsPair objects, which are pairs of common actions and entities between US-pairs.

  For each pair, it checks whether the specified names and types match either the action or entity in the pair. If a match is found, the method returns *true*, which means that the specified elements are part of a redundant pair.

- printRedundantItems: This method is responsible for creating a report on redundancy items based on the data stored in the class instance.

  As input, it takes several parameters, including a "FileWriter" to write to the textual report's file, lists of different pairs of "Targets", "Contains" and "Triggers", and a JSON-object to store the report data in JSON format, which is vital for evaluation.

ReportExtractor Class: Methods related to Extracting Redundancy Items

To extract the redundant elements and save them in a redundancyItems object, we must first iterate through all existing *minimal-mode.ecore* files and extract the redundant reduandcy item from each EPackage.
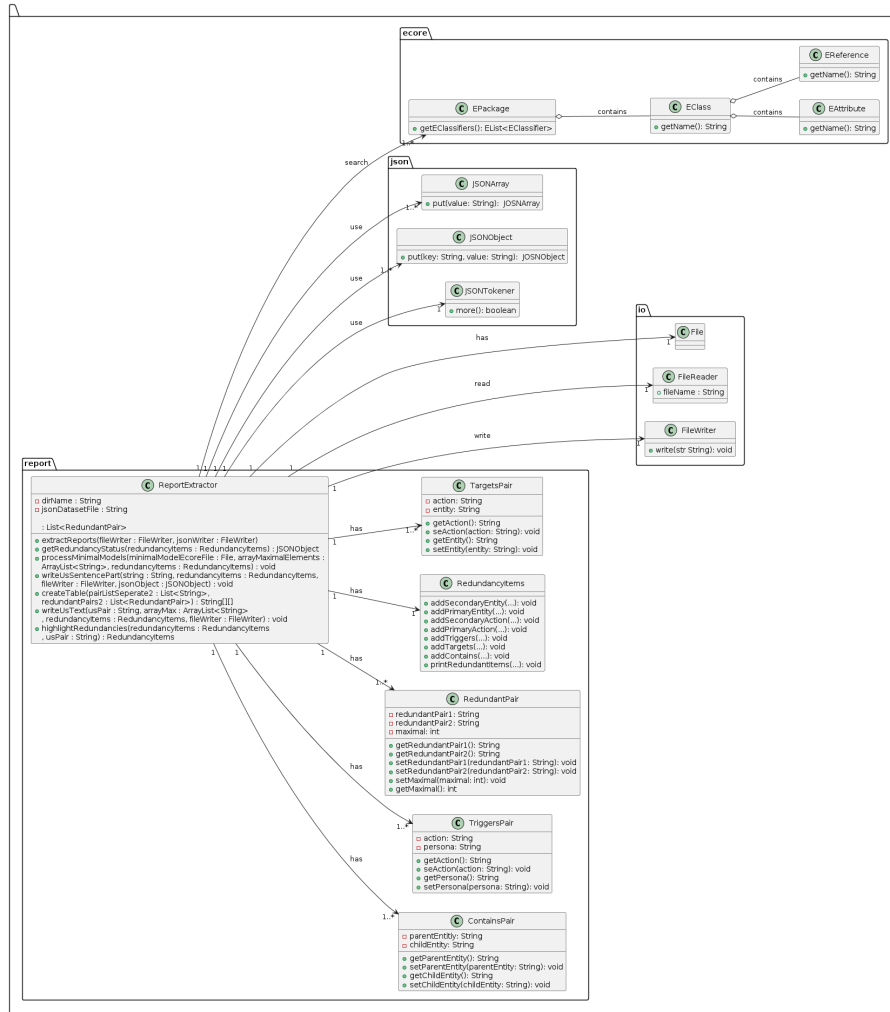
Figure 34: Class diagram for the process of extracting redundancy items

Figure 34 is a class diagram that illustrates the attributes, operations and relationship related to the extraction of redundancy items founded by the CDA tool and their storage in a RedundancyItems object.

The following methods in the ReprortExtractor class are responsible for extracting the redundancies found by the CDA tool:

- extractReports: This method orchestrates the extraction and analysis of created CDA report from a directory containing conflicted US-pairs and their associated reasons(conflict reason), and generates both text and JSON reports for further investigation and processing.

  It receives two FileWriter objects as input, one for writing textual and one for JSON reports.

69

It iterates through each directory in the main directory that represents a conflict pair and uses the *checkIfReportExist* method to make sure that the current US-pair are not already proceeded and the *containsAnd* method to check whether it contains the conjunction "AND" to make sure that it is the valid US-pair name(*e.g.* "user_stroy_12_AND_user_stroy_39" is a valid directory name).

For each valid conflict pair directory, it iterates through the conflict reason directories within the current conflict pair directory and uses the *minimalEcoreExist* method to check whether a *minimal-model.ecore* file exists with reference to a conflict reason.

To identify redundancy items, it uses the method *processMinimalModels* and reads the "minimal-model.ecore" file, processes its content and iterates over the contained EPackages in order to process them further with the method *iteratePackages*, which saves all redundancy items in RedundacyItems object.

In addition, the methods *hasEntitys*, *hasActions* and *hasTargets* are used to check whether the identified elements contain "Primary/Secondary actions" and "Primary/Secondary Entities" with common "Targets" reference. If the identified elements fulfil the criteria, the redundancy pair is included in the report.

It then writes the potentially redundant USs and their clauses to the textual as well as JSON report files for further analysis.

As output, the method returns a list of RedundantPair objects containing information about identified redundancies between US-pairs.

- createOrOverwriteReportFile: The method is responsible for creating or overwriting report file. It first ensures the existence of a report file. If the file doesn't exist, it creates a new one; if it already exists, it overwrites the existing file. Finally, it returns a *FileWriter* to allow writing to the report file.

- checkIfReportExist: This method takes two parameters, namely US-pair and the list of all previously processed pairs in the CDA report directory. It returns *true* if the US-pair was found in the pairList, which means that a report with the specified pairs has already been executed and therefore does not need to be executed again.

- minimalEcoreExist: This method checks the existence of a *minimal-model.ecore* file using a conflict pair and a conflict reason generated by CDA tool.

  As input, it receives a conflict pair and a conflict reason. Using these parameters, the method constructs the file path to the minimal-model.ecore file and checks whether the file exists under the constructed path.

  If the ECore file exists, the method returns *true*, indicating that the minimal-model.ecore file exists for the examined conflict pair and conflict reason. If the file does not exist, *false* is returned.

- containsAnd: This method ensures that the folder name is identified with "_AND_", as the report generated by CDA tool is formatted like "user_story_<digit >_AND_

user_story<digit >". It returns *true* if the examined directory contains "AND", otherwise *false* is returned.

- processMinimalModels: This method reads a Minimal Model Ecore file, processes its content and iterates over the contained EPackages in order to process them further with the *iteratePackages* method.

  As parameters, it receives a File object that represents examined minimal model ecore file, an array list in which the names of the redundant elements are stored, and a RedundancyItems object that is used to handle redundant elements.

  First, a ResourceSet and a ResourceFactoryRegistry corresponding to the minimal model Ecore file are set up and a Resource object is created from the Ecore file; the *iteratePackages* method is called for each EPackage.

- iteratePackages: This method identifies redundancies within the attributes and references of EClasses in examined minimal model, stored them accordingly and updates the RedundancyItems object.

  It takes several parameters, such as the EPackage to be iterated over, an array list in which the names of the redundant elements are stored, and a RedundancyItems object.

  It iterates through every EClassifier in the minimal package that contains EClasses and checks whether "#" is present in EClass; if this is the case, EClass has been recognised as a conflict by CDA tool.

  If an attribute is found, the class of the conflicting attribute is determined and added to the corresponding element within RedundancyItems (e.g. Primary/Secondary Action/Entity).

  Each EReference is then iterate through in the EClass. Depending on the reference name, the reference is added to the corresponding element within RedundancyItems (e.g. Triggers, Targets, Contains). The method is completed once all EClassifiers within the specified EPackage have been processed.

- hasActions: This method is useful for using the data stored in RedundancyItem to determine whether certain actions are present in a list of redundant entries, in order to later check whether the identified elements fulfil the criteria and the redundancy pair is included in the report.

  It checks if there's a match with name of any primary/secondary action stored in the RedundancyItems object. If match is found, it immediately returns *true*.

- hasEntitys: This method is used to determine whether certain secondary/primary entities are present in the RedundancyItems object based on their name. For each item, it checks whether there is a match with the name of a primary/secondary entity stored in the RedundancyItems object.

If a match is found, *true* is returned immediately, in order to later check whether the identified elements fulfil the criteria and the redundancy pair is included in the report.

- hasTargets: This method uses the content of a RedundancyItems object to determine whether targets are present in a list of founded redundant elements.

  As input, it receives an array list with the names of the redundant elements and an object of the type RedundancyItems, which contains a collection of founded "Targets".

  The method checks whether there is a match with the name of any targets reference stored in the RedundancyItems object.

  If a match is found between the elements, the method immediately returns *true*, which means that at least one targets reference is present, in order to later check whether the identified elements fulfil the criteria and the redundancy pair is included in the report.

- readJsonArrayFromFile: This method provides the ability to read JSON data from a file and convert it into a JSONArray object, handling cases where the file is empty or does not exist.

  It receives the file path of the JSON file to be read as input. An attempt is made to open the specified file and check whether the file is empty or does not exist. If the file exists and is not empty, it reads the JSON data from the file and creates a JSON array object from the JSON data read from the file and returns the JSON array object with the JSON data.

- getRedundancyStatus: This method add statistics like count of main/benefit/total redundancies into JSON report. It receive as input redundancyItems and as output write the count of main/benefit/total redundancies into JSON report already defined in method highlightRedundancies.

Methods for Extracting Data from Dataset of Backlog

To get information about USs such as text, redundant clauses in targets/contains/triggers references between US-pairs, we need to extract data from the dataset stored in the JSON file.

The following methods in the ReprortExtractor class are responsible for extracting data from dataset within JSON file:

- readJsonArrayFromFile: This method receives a JSON file as input. After reading, the JSON content is tokenised, parsed into a JSON array and the parsed JSON array is returned.

- getUssTexts: This method ensures that the text of the specified US-pair is retrieved from the JSON file and properly assigned to the RedundancyItems object for further processing. It receives a US-pair and RedundancyItems as input.

  It reads a JSON array from a file using the *readJsonArrayFromFile* method, iterates over each JSON object in the array and compares the extracted US identifier with the US identifier extracted from the input US-pair. If a match is found, the text of the first and second USs is set in the redundancyItems object.

- getCommonTargets: Used to determine overlaps in "Targets" references (pairs of actions and entities) between specified US-pairs from a JSON file.

  It receives the identifiers of the US-pair as input. After it finds the US objects, it compares each pair of entries (actions and entities) in "Targets" references of the two USs and looks for matches.

  The output is the matches, i.e. the list of common pairs of actions and entities, in "Targets" reference.

- getCommonContains: Used to determine overlaps in "Contains" references (pairs of entities) between specified US-pairs from a JSON file.

  It receives the identifiers of the US-pair as input. After it finds the US objects, it compares each pair of entities in "Contains" references of the two USs and looks for matches.

  The output is the matches, i.e. the list of common pairs of entities, in "Contains" reference.

- getCommonTriggers: Used to determine overlaps in "Triggers" references (pairs of personas and actions) between specified US-pairs from a JSON file.

  It receives the identifiers of the US-pair as input. After it finds the US objects, it compares each pair of entries (personas and actions) in "Triggers" references of the two USs and looks for matches.

  The output is the matches, i.e. the list of common pairs of personas and actions, in "Triggers" reference.

Methods for Highlighting Words using Hash Symbol(#)

In order to distinguish redundancy words between a US-pair, in text of each US, we decide to highlighting redundant words using hash symbol like " ... #<word># ... ".

The following methods in the ReprortExtractor class are responsible for highlighting redundancy words:

- writeUsText: This method reads the text of examined US-pair, highlights redundant element between them using *highlightRedundancies* method, writes the highlighted text to a file, and records information about the redundant elements.

As input it receive a US-pair, a list of redundant pairs to store redundant pairs in it, an object RedundancyItems containing stored redundancy elements and a FileWriter object used for writing output.

It extract the identifier of the two USs from US-pair, retrieves the text of the USs from JSON file and add them to RedundancyItems, invoke the *highlightRedundancies* method to identify and highlight redundants between the USs, it writes the highlighted text of each US to the FileWriter. Finally, sets the redundant pair and count of redundancy clauses.

- highlightRedundancies: This method identifies redundancies between US-pair, applies hash symbols using *applyHashSymbols* method to highlight common items and updates the redundancy counts in the redundancyItems object.

  It takes two parameters, redundancyItems and US-pair, which represents the pair of USs to be analysed.

  It checks whether both USs contain a main clause part or whether one of them has a benefit part or whether both USs also have a benefit part.

  It applies hash symbols to common elements that only occur in the part of the USs that occurs in the same part (e.g. only main or only benefit part of the USs).

  In each condition, it checks if there are redundancy clauses in the main part, then persona is also highlighted using *applyHashSymbolPersona* method. It also updates the count of main/benefit/total redundancies and sets the changed text of USs. Finally, it returns the updated redundancyItems object.

- applyHashSymbols: This method is used to mark certain words within a substring with hash symbols (#) at the beginning and end to ensure that they are distinguishable and can be easily identified or processed later.

  It takes a substring in which replacements are to be made and a field of matches containing the words to be surrounded with hash symbols. First, the field of matches is sorted in descending order of length and processed accordingly to avoid adding hash symbols to unwanted clauses.

  Example 4.12. *For example, let's assume that we have "data" and "data format" as redundancy elements. If we continue first with "data" and then with "import data", "import data" will be replaced by "import #data#", which is not desired.*

- hasMoreThanFour/SixHashSymbols: These methods receive a text from the US as input. They are used to check whether there are redundant clauses in the main part of the sentence (it can be one or two clauses). If so, *true* is returned.

- applyHashSymbolPersona: This method identifies common "Triggers" references, marks them with hash symbols and returns the changed text parts together with the count of redundant triggers references.

As input, it receives a list of common triggers between US-pair, RedundancyItems and the parts of the USs. It iterates through the list of common triggers and checks whether both elements(persona and primary action) of the triggers are present in both parts.

It then increments the redundancy count to keep track of the count of redundant triggers. The output returned, is the text of USs containing the manipulated text parts with hash symbols and the redundancy count in main part.

- applyHashSymbolTargets: This method identifies common targets references between two parts of USs, marks them with hash symbols and returns the changed text parts together with the count of redundant targets references.

  As input, it receives a list of common targets between the US-pair, RedundancyItems and the parts of the USs.

  It iterates through the list of common targets and checks whether both elements (primary/secondary action and primary/secondary entity) of the targets are present. It then increments the redundancy count in examined part of USs to keep track of the count of redundant targets.

  The output returned is the text of USs containing the manipulated text parts with hash symbols and the redundancy count in examined part(main or benefit).

- applyHashSymbolContaians: This method identifies common contains references between two parts of USs, marks them with hash symbols and returns the changed text parts together with the count of redundant contains references.

  As input, it receives a list of common contains between the US-pair, RedundancyItems and the parts of the USs. It iterates through the list of common contains and checks whether both elements of the contains are present in both parts. It then increments the redundancy count to keep track of the count of redundant contains.

  The output returned is the text of USs containing the manipulated text parts with hash symbols and the redundancy count in examined part(main or benefit).

Methods related to Report Parts of USs

To show which part of the USs with redundancy words occur between a US-pair, we have split the individual parts of the USs and included them in the report. The following methods in the ReprortExtractor class are responsible for splitting and reporting the parts of the USs :

- splitUsText: This method is used to split the text of two USs into separate sections based on the occurrence of redundancy clauses.

  The input is the text of the first and second US and their corresponding identifiers, a FileWriter for writing to a file and a JSON object for processing JSON data.

It splits each US text into three parts using commas and saves the result in arrays. It iterates over parts of the first and second USs and searches for occurrences of hash symbol pairs. For each part, the number of hash symbol pairs found is counted. Finally, all parts of the records that contain hash symbols are written to a text file and a JSON file as well.

- writeUsSentencePart: This method facilitates the extraction and storage of highlighted USs parts from US-pair in a textual report file for further analysis.

  As input, it receives the US-pair, RedundancyItems and a FileWriter allowing the extracted USs parts to be written. It receives the text of the USs from the reundancyItems object and calls the *splitUsText* method to split the US-pair texts into USs parts with highlighted elements.

  The extracted USs parts are also written to the text and JSON report files, using the FileWriter for further processing and analysis.

Example 4.13. *Take, for example, the following US-pair:*

user_story_60: #g22# as an it staff member, I want to #know# #how# the #data# is #used#, so that I can determine what kind of basic services and functionalities are required.

user_story_04: #g22# as a data manager, I want to #know# #how# the #data# is #used#, so that I can develop more detailed usage and support scenarios with researchers.

The following sentence parts are candidates for possible redundancies between user stories:

user_story_04: I want to #know# #how# the #data# is #used#

user_story_60: I want to #know# #how# the #data# is #used#

Methods related to Creating Table

The summary of potential redundancy between US-pairs is presented in a table, which makes it easy to find out which US-pairs have been identified as potentially redundant US-pairs and how many redundancy elements there are in total. The following methods in the ReprortExtractor class are responsible for creating a table at the beginning of the text report:

- writeTable: This method is used to write a table of potential redundancies between USs and the count of their total redundant clauses *createTable* method.

As input, it receives a File object into which the table is inserted and a list of redundancy pairs containing information about redundant clauses in US-pars.

It reads the existing content of the textual report file into a StringBuilder. It creates a table to display the potential redundancies between USs and the count of total redundancy clauses.

The table headers and contents are generated based on the redundant pairs. It calculates the maximum width for each column in the table to ensure proper formatting. Finally, the table content is written to the FileWriter, followed by the existing content stored in the report's StringBuilder.

- createTable: This method prepares the content for the table in which potential redundancies between USs are displayed, taking into account the total redundancy count between each US-pairs based on the RedundantPair objects provided.

  As input, it receives a list of unique US-pairs for which the table is to be created and a list of RedundantPair objects containing information about redundant clauses between US-pairs.

  It initialises a two-dimensional array containing the contents of the table. The size of the table is determined by the count of unique pairs of USs plus one for the header row and column.

  It fills the header row and the first column of the table with unique pairs of US-pairs, replacing "user_story" with "us" for the purpose of brevity.

  It calculates the maximum redundancy count between each US-pair by calling the method *getTotalRedundanciesFromPair*. Finally, it fills the table with the total redundancy count.

  The output is a two-dimensional array representing the contents of the table, with each cell containing the maximum redundancy count between the corresponding pair of USs.

- getTotalRedundanciesFromPair: This method makes it easier to retrieve the total number of redundancies between a US-pair from a list of RedundantPair objects.

  As input, it receives a list of RedundantPair objects containing information about redundant elements in US-pairs, where the first and second USs are to be compared.

  It iterates through each RedundantPair object in the RedundantPairs list and checks for each RedundantPair object whether the examined US-pair matches as redundant US-pair.

  If a matching pair is found, the maximum redundancy number stored in this RedundantPair object is returned. If no matching pair is found, *zero* is returned, indicating that there are no redundancies between the examined US-pair.

Step 5: Report Evaluation

The Evaluation class, part of the org.henshin.backlog.code.evaluation package, was developed to determine the level of redundancy in USs based on JSON reports. This class provides methods to evaluate whether two USs are either fully or partially redundant, analysing different application components of these USs.

Figure 35 is a class diagram that illustrates the attributes, operations of the Evaluation class and its relationship to other classes.
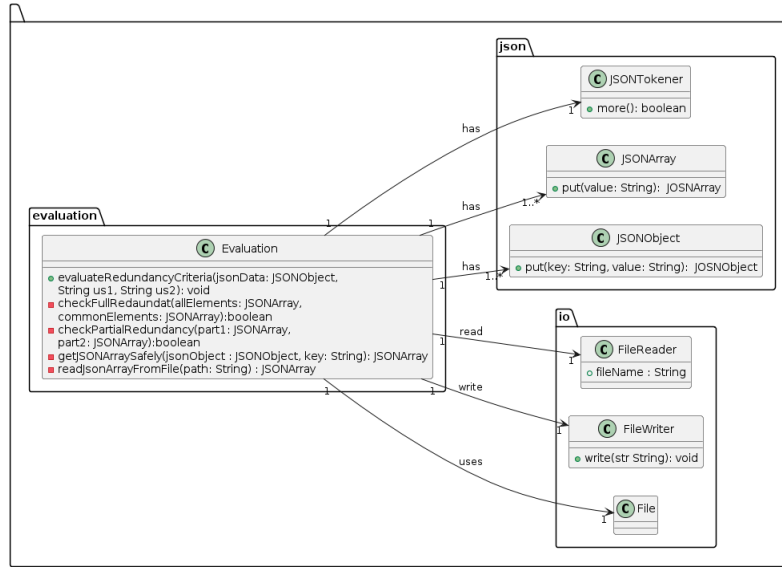
Figure 35: Class diagram related to Evaluation class

The evaluation process involves a complex logic to determine whether user stories are redundant. This includes:

- Checking whether the arrays are empty or contain similar elements.

- Comparing the individual elements in the arrays for both USs to determine if they fully match (full redundancy) or if they have some common elements (partial redundancy).

Example 4.14. *For the two US-pairs of dataset G03, we apply the evaluation class to determine whether there is redundancy in the main or benefit part, and if so, what type of redundancy is recognised(full or partially).*

As shown in Listing 8, four entries are added to the JSON report, namely "Main Partially Redundant", "Benefit Part Fully Redundant", "Main Part Fully Redundant", "Benefit Partially Redundant" as "Status" which their value is whether true or false.

```
"Status":
{"Main Part Redundancy Clauses": 0,
"Main Part Partially Redundant": false,
"Total Redundancy Clauses": 1,
"Benefit Part Fully Redundant": false,
"Benefit Part Redundancy Clause": 1,
"Main Part Fully Redundant": false,
"Benefit Part Partially Redundant": true
},
```

Listing 8: Example of generated entries in JSON report regarding evaluation of level of redundancy in main or benefit part

In the case of user_story_12 and user_story_39, the entry "Benefit Partially Redundant" was marked as true, which means that US-pair in benefit parts are partially redundant.

The class performs these checks by iterating through the JSON arrays of Triggers, Targets and Contains and comparing each element with those in the common sections to determine redundancy.

The class offers a detailed mechanism to evaluate redundancy through following methods:

- *evaluateRedundancyCriteria*: This method compares elements such as "Triggers", "Targets" and "Contains" in the "Main" and "Benefit" parts of USs. The method checks whether these clauses are fully redundant (i.e. they are identical between USs) or partially redundant (i.e. they share some clauses but are not fully identical).

  As input, it receives the JSON object for a specific US-pair.

  In particular, it checks various conditions, e.g. whether common clauses are present, whether targets or contains are empty and whether the clauses in the stories are fully or partially identical.

  As output, the method updates the JSON report with new keys indicating the redundancy status:

  Main Part Fully Redundant: A Boolean value indicating whether the "main part" of the USs is fully redundant.

  Main Part Partially Redundant: A Boolean value that indicates whether the "main part" of the USs is partially redundant.

  Benefit Part Fully Redundant: A Boolean value that indicates whether the "benefit part" of the USs is fully redundant.

  Benefit Partially Redundant: A Boolean value that indicates whether the "benefit part" of the USs is partially redundant.

- checkPartialRedundancy: The method processes these arrays to identify partial redundancy by checking whether elements from arrays of elements match elements from another array.

It receives the two arrays of elements as input, iterates through each element of the first input array and compares each of these elements with all elements in the second input array. A match counter records how often elements from the first input array find a match in the second input array.

As output, method returns a Boolean value. True: indicates that there is at least one matching element pair between the two JSON arrays, which means partial redundancy. False: indicates that there are no matching elements, indicating no partial redundancy between these specific parts of the USs.

Example 4.15. *Assume we have the following JSON arrays for two different USs:*

First JSON array: [["login", "button"], ["help", "link"]]
Second JSON array: [["logout", "button"], ["help", "link"]]

The method would determine that the second element of the first array ("help", "link") matches the second element of the second array. As there is at least one match, checkPartialRedundancy would return true, indicating partial redundancy.

- getJSONArraySafely: This method is designed to handle JSON operations safely by retrieving JSON array from a JSON object without the risk of throwing exceptions if the specified key does not exist.

  As input it receives the JSON object from which the JSON array needs to be extracted and a key corresponding to the JSON array within the provided JSON object.

  The method first checks whether the provided key exists in the JSON object. If the key exists, it retrieves the JSON array associated with that key. If the key does not exist, it returns an empty JSON array.

Example 4.16. *Let's assume that a JSON object representing USs details should contain the key for "Benefit Part", which does not exist:*
{
"Common Targets" : {
"Main Part" : [["login","button"],["help","link"]] }
}

If the "Benefit Part" in the "Common Targets" needs to be accessed that does not exist, this is handled safely with getJSONArraySafely by returning an empty array to avoid runtime errors.

- readJsonArrayFromFile: The method enables JSON data to be read from a file and converted into a JSON array object. It takes the file path as input and attempts to

open the file, checking whether the file is empty or missing. If the file is available and contains data, this data is read, converted into a JSON array object and returned.

Error Handling

There were erroneous data in datasets that force us to handle them correctly. Therefore, we implement/use the following exceptions to accurately distinguish and handle them.

The following classes, which extend the *Exception* class, are used in the *org.henshin .backlog.code.rule* package and relate to error handling related to the JSON entries in the dataset of the backlogs and to the Ecore meta-model required to create the rules based on it:

- EmptyOrNotExistJsonFile: Is triggered if the JSON file could not be found in the file system.

- ActionInJsonFileNotFound: Is triggered if the entry *Action*, which contains *Primary/Secondary Actions*, is not present in the JSON file and its absence should be reported.

- EntityInJsonFileNotFound: Is triggered if the entry *Entity*, which contains *Primary/Secondary Entity*, is not present in the JSON file and its absence should be reported.

- PersonaInJsonFileNotFound: Is triggered if the entry *Persona* does not exist in the JSON file for a specific US.

- TextInJsonFileNotFound: Is triggered if the entry *Text* does not exist in the JSON file for a specific US.

- UsNrInJsonFileNotFound: Is triggered if the entry *Us_Nr* does not exist in the JSON file for a specific US.

- EdgeWithSameSourceAndTarget: Refers to the creation of edges in graph transformation rules and is triggered if the source and target of the edge have already been created, then the duplicate edge should be avoided.

- TargetsInJsonFileNotFound: Is triggered if the entry *Targets* does not exist in the JSON file for a specific US.

- ContainsInJsonFileNotFound: Is triggered if the entry in *Contains* is not present as *Primary/Secondary Entity* in the JSON file and its absence should be reported.

- TriggersInJsonFileNotFound: Is triggered if the entry *Triggers* does not exist in the JSON file for a specific US.

- EcoreFileNotFound: Is triggered if the required ECore meta-model file could not be found and should also be reported.

Within the *org.henshin.backlog.code.report* package, special classes that extend the *Exception* class are designed to solve problems related to the CDA report directory, which encapsulates all US-pairs together with the associated conflict reasons. These classes include:

- CdaReportDirIsEmpty: This exception is called if the CDA report directory is found but has no content.

- CdaReportDirIsNotADirectory: This exception is thrown in scenarios where the path provided for the CDA report directory is either not a directory (e.g. it is a file) or the specified path does not lead to a directory.

- CdaReportDirNotFound: This exception is triggered if the CDA report directory cannot be found within the specified path.

Limitations

There are technical limitations that are causing us to change our implementation strategy. The following limitations should be clarified at the beginning:

- Use Eclipse version 2023-03, as Henshin version 4 cannot be installed with the latest version of Eclipse.

- Working with Java as the programming language, as the Henshin and CDA APIs are only available in the Java programming language.

- CDA API is not yet implemented to take into account conflicts and dependencies for attributes that are crucial for our approach. This forces us to use the CDA graphical user interface(GUI) instead of the CDA API.

- Lack of Henshin documentation regarding methods and classes, which makes it time consuming to understand the methods and make the right decision.

4.4   Test

In this section, we aim to validate certain functionalities, check the system requirements and ensure reliability and robustness of implemented classes and methods. As a test strategy, we perform unit tests with *JUnit* version 4[20] as version 4 is more suitable and compatible with Eclipse version 2023-03.

We also used EclEmma[21], a code coverage tool integrated into the Eclipse IDE, to ensure thorough testing and coverage. By using EclEmma, we were able to systematically measure the effectiveness of our test suites and determine the test coverage for each individual class.

---

[20]https://junit.org/junit4/
[21]https://www.eclemma.org/

Configuration of Test Environment

In the main project *org.henshin.backlog*, we create a separate package called *org.henshin. backlog.test*. This package contains two Java classes namely *ReportExtractorTest.java* and *RuleCreator_Test.java*, each of which corresponds to the Java source codes accordingly.

Scope of Testing

The scope of the test depends on the system requirements and the two most important implemented classes *ReportExtractorTest.java*, *RuleCreator_Test.java,Evaluation.java* and their methods. The implemented error handling classes are also tested.

Test Cases and their Code Coverage

We describe the specific test cases that are performed during the tests. Each test case contains a description of the test scenario, the data provided and the expected result. For refining and improving our test cases, we used a code coverage report created by EclEmma in order to increase coverage and ensure a more reliable, error-resistant application.

Table 11 shows the test cases for the RuleCreator.java class and Figure 36 shows the code coverage.

Test Case	Supplied Data	Expected Outcome	Description
testAssignCmodule	Assing a dummy ECore model file	Through an exception: *EcoreFileNotFound*	Check whether the ECore model already exists and CModule is correctly assigned
ActionNotExist (test-ProcessJsonFile)	JSON file with a US without "Action" entry	Through an exception: *ActionInJsonFileNot-Found*	Check whether there is an entry "Action" in the related US in JSON file
EntityNotExist (testProcessJsonFile)	JSON file with a US without "Entity" entry	Through an exception: *EntityInJsonFileNot-Found*	Check whether there is an entry "Entity" in the related US in JSON file
PersonaNotExist (test-ProcessJsonFile)	JSON file with a US without "Persona" entry	Through an exception: *PersonaInJsonFileNot-Found*	Check whether there is an entry "Persona" in the related US in JSON file
TargetsNotExist (test-ProcessJsonFile)	JSON file with a US without "Targets" entry	Through an exception: *TargetsInJsonFileNot-Found*	Check whether there is an entry "Targets" in the related US in JSON file
ContainsNotExist (test-ProcessJsonFile)	JSON file with a US without "Contains" entry	Through an exception: *ContainsInJsonFileNot-Found*	Check whether there is an entry "Contains" in the related US in JSON file
ContainsInTargets (test-ProcessContainsEdges)	JSON file with a US which the entity in "Contains" entry also belong to Targets	The entity in Contains should be annotated as <Delete>	Check whether entity in Targets have connection as Contains, if so entity in Contains should also annotated as <Delete>

Test Case	Supplied Data	Expected Outcome	Description
TriggersNotExist (test-ProcessJsonFile)	JSON file with a US without "Triggers" entry	Through an exception: *TriggersInJsonFileNot-Found*	Check whether there is an entry "Triggers" in the related US in JSON file
DuplicateTriggers (test-ProcessContainsEdges)	JSON file with a US with duplicate Triggers entries	Through an exception: *EdgeWithSameSource AndTarget*	Verify whether there is duplicated entries in Triggers JSON-array
TextNotExist (testProcessJsonFile)	JSON file with a US without "Text" entry	Through an exception: *TextInJsonFileNot-Found*	Check whether there is an entry "Text" in the related US in JSON file
UsNrNotExist (testProcessJsonFile)	JSON file with a US without "Us_Nr" entry	Through an exception: *UsNrInJsonFileNot-Found*	Check whether there is an entry "US_Nr" in the related US in JSON file
UndefindedEntity (testProcessContainsEdges)	Specify an entity that is not contained as "Entity" in the JSON file, but appears as "Contains"	Through an exception: *EntityInJsonFileNot-Found*	Check whether the entity that appears in the *Contains* entry has already been identified as an entity
PrimaryActionNotFound (testProcessTargetsEdges)	Specify a primary action that is not contained as "Primary Action" in the JSON file, but appears as "Targets"	Through an exception: *ActionInJsonFileNot-Found*	Check whether the primary action that appears in the *Targets* entry has already been identified as a primary action
PrimaryEntityNotFound (testProcessTargetsEdges)	Specify a primary entity that is not contained as "Primary Entity" in the JSON file, but appears as "Targets"	Through an exception: *EntityInJsonFileNot-Found*	Check whether the primary entity that appears in the *Targets* entry has already been identified as a primary entity
SecondaryActionNotFound (testProcessTargetsEdges)	Specify a secondary action that is not contained as "Secondary Action" in the JSON file, but appears as "Targets"	Through an exception: *ActionInJsonFileNot-Found.class*	Check whether the secondary action that appears in the *Targets* entry has already been identified as a secondary action
SecondaryEntityNotFound (testProcessTargetsEdges)	Specify a secondary entity that is not contained as "Secondary Entity" in the JSON file, but appears as "Targets"	Through an exception: *EntityInJsonFileNot-Found.class*	Check whether the secondary entity that appears in the *Targets* entry has already been identified as a secondary entity
UndefindedEntity (testProcessTargetsEdges)	Specify an entity that is not contained as "Entity" in the JSON file, but appears as "Targets"	Through an exception: *EntityInJsonFileNot-Found*	Check whether the entity that appears in the *Targets* entry has already been identified as an entity
ActionNotFound	Provision of a JSON file that contains actions in *Targets* that are not entered in primary/secondary action	Through an exception: *EntityInJsonFileNot-Found*	Verify whether the action that appear in the *Targets* entry has already been identified as an action
ReadJsonArrayFromFile	Assign a dummy JSON file	Through an exception: *EmptyOrNotExistJson-File*	Check whether the JSON file already exists

Test Case	Supplied Data	Expected Outcome	Description
EmptyOrNotExistJsonFile	Assign an empty JSON file	Through an exception: *EmptyOrNotExistJson-File*	Check whether the JSON file is empty

Table 11: Test cases for RuleCreator class

RuleCreator

Element	Missed Instructions	Cov.	Missed Branches	Cov.	Missed	Cxty	Missed	Lines	Missed	Methods
● main(String[])		0 %		n/a	1	1	3	3	1	1
● processJsonFile(JSONArray)		100 %		100 %	0	10	0	53	0	1
● processEntities(JSONObject, CRule, JSONObject, JSONArray, Map, String)		100 %		100 %	0	7	0	31	0	1
● processActions(JSONObject, CRule, JSONObject, CNode, Map, String)		100 %		100 %	0	5	0	21	0	1
● processTargetsEdges(JSONObject, JSONArray, Map, Map, String)		100 %		100 %	0	4	0	18	0	1
● processContainsEdges(JSONObject, JSONArray, JSONArray, Map, String)		100 %		90 %	1	6	0	15	0	1
● createRules(String)		100 %		n/a	0	1	0	7	0	1
● processPersona(JSONObject, JSONArray, CRule, String)		100 %		n/a	0	1	0	4	0	1
● processText(JSONObject, CRule, String)		100 %		n/a	0	1	0	4	0	1
● readJsonArrayFromFile()		100 %		100 %	0	2	0	6	0	1
● getJsonFileAbsolutePath()		100 %		100 %	0	2	0	4	0	1
● checkEntityIsTarget(String, JSONArray)		100 %		100 %	0	3	0	6	0	1
● getEcoreFileAbsolutePath()		100 %		100 %	0	2	0	4	0	1
● assignCmodule()		100 %		100 %	0	2	0	5	0	1
● RuleCreator(String, String, String)		100 %		n/a	0	1	0	5	0	1
● processRule(JSONObject, String, CModule)		100 %		n/a	0	1	0	2	0	1
● getJsonFile()		100 %		n/a	0	1	0	1	0	1
● getHenshinFile()		100 %		n/a	0	1	0	1	0	1
● getEcoreFile()		100 %		n/a	0	1	0	1	0	1
Total	5 of 923	99 %	1 of 66	98 %	2	52	3	191	1	19

Figure 36: Code coverage related to class RuleCreator

Table 12 shows the test cases for the ReportExtractor.java class and Figure 37 shows the code coverage.

Test Case	Supplied Data	Expected Outcome	Description
testEmptyDirectroy	Assing a dummy directory	Through an exception: *CdaReport-DirIsEmpty.class*	Check if CDA Report directory is empty
testEmptyJSONFile	Assing an empty JSON dataset file	Through an exception: *EmptyOrNotExistJson-File*	Check if JSON dataset file is empty
testCdaDirNotDirectroy	Assing a file instead of CDA directory	Through an exception: *CdaReportDirIsNotA Directory*	Check if assigned path is a directory
testCdaDirectroy	Assing a not readable directory	Through an exception: *CdaReportDir NotFound*	Check whether CDA directory is accessible
testFinalReportDir	Assign an empty dataset directory	Return *false* if the assigned directory is empty	Check if directory of datasets are empty
testMinimalEcoreExist	Assign a CDA directory without minimal-ecore file	Return *false* if the minimal-ecore file not found	Check whether the Ecore file already exist in CDA directory
testWriteTable	Provide a CDA report for a US-pair	A table should be created in the report file	Check whether the table for US-pairs has already been created

Test Case	Supplied Data	Expected Outcome	Description
completeMajorElements Edge (testExtractReports)	Provision of a CDA report for US-pair with exactly one "Targets" edge	Inclusion of this US-pair in the text report with the nodes "Action", "Entity" and "Targets" edge	Verifies *extractReports* method when all major elements are present in the CDA report
completeMajorElements Upper Edge (testExtractReports)	Provide a CDA report for a US-pair with at least one "Targets" edge and redundancy elements such as Triggers	Generated textual report contains information about Secondary Entities, Secondary Actions, Targets, and Triggers	Verifies *extractReports* method when all major elements are present in the CDA report and the upper edge case is reached
notCompleteMajor Elements (testExtractReports)	Provide a CDA report for a US-pair without "Targets" edge, but with action and entity	The US-pair should not reported	Verifies the behavior of the *extractReports* method when not all major elements are present in the input data
getBenefitPart RedundanciesElements (testExtractReports)	Provision of a CDA report for a US-pair that only contains redundancy clauses in the "Benefit" part	Check whether the count of redundancy clauses in the benefit part of the USs matches the value "Benefit Part Redundancy Clause" specified in the JSON_Report file	Verifies the behaviour of the *extractReports* method when there are redundancy elements in the benefit part of USs
getMainPartRedundancies Elements (testExtractReports)	Provision of a CDA report for a US-pair that only contains redundancy clauses in the "Main" part	Check whether the count of redundancy clauses in the main part of the USs matches the value "Main Part Redundancy Clause" specified in the JSON_Report file	Verifies the behaviour of the *extractReports* method when there are redundancy elements only in the main part of USs
getTotalRedundancy Elements (testExtractReports)	Provision of a CDA report for a US-pair that contains redundancy clauses in the "Main" and "Benefit" parts	Check whether the count of redundancy clauses in the main and benefit parts of the USs matches the value "Total Redundancy Clause" specified in the JSON_Report file	Verifies the behaviour of the *extractReports* method when there are redundancy elements in the main and benefit parts of USs
highlightPersona (testExtractReports)	Providing a CDA report for a US-pair with redundancy clause in "Triggers" edge (from Persona to Primary Action)	The persona should only be marked with hash symbol if there is a redundant clause in the main part	Checks the behaviour of the *extractReports* method when highlighting redundant persona in USs
BenefitInBothUss (testhigHlightRedundancies)	Provision of a CDA report for a US-pair where both USs have the "benefit" part	Search both the main and the benefit parts of USs and mark redundancy clauses with a hash symbol if they occur	Verifies the behaviour of the *highlightRedundancies* method when both USs have benefit part
noBenefitInBothUss (testhigHlightRedundancies)	Provision of a CDA report for a US-pair where both USs do not have the "benefit" part	Search only the main part of USs and mark redundancy clauses with a hash symbol	Verifies the behaviour of the *highlightRedundancies* method when both USs don't have benefit part

Test Case	Supplied Data	Expected Outcome	Description
noBenefitInUs1 (testhigHlightRedundancies)	Provision of a CDA report for a US-pair where only the second US have the "benefit" part	Search only the main part of USs and mark redundancy clauses with a hash symbol	Verifies the behaviour of the *highlightRedundancies* method when only the second US have benefit part
noBenefitInUs2 (testhigHlightRedundancies)	Provision of a CDA report for a US-pair where only the first US have the "benefit" part	Search only the main part of USs and mark redundancy clauses with a hash symbol	Verifies the behaviour of the *highlightRedundancies* method when only the first US have benefit part
ContainInBenefitPart (testHighlightRedundancies)	Provide a CDA report for a US-pair with redundancy clauses of "Contains" within "Benefit" part	The entities included in Contains should be marked with hash symbol	Checks the behaviour of the *highlightRedundancies* method when highlighting redundant entities included in the Contains
ContainInMainPart (testHighlightRedundancies)	Provide a CDA report for a US-pair with redundancy clauses of "Contains" within "Main" part	The entities included in Contains should be marked with hash symbol	Checks the behaviour of the *highlightRedundancies* method when highlighting redundant entities included in the Contains
TargetsInMainPart (testHighlightRedundancies)	Provide a CDA report for a US-pair with redundancy clauses with more than one "Targets" within the "Main" part	Founded redundancy clauses should also be marked with a hash symbol, if the main part contains more than one redundancy clause as Targets	Check the behaviour of the *highlightRedundancies* method if there is more than one redundancy clause as Targets in main part
TargetsInBenefitPart (testHighlightRedundancies)	Provide a CDA report for a US-pair with redundancy clauses with more than one "Targets" within the "Benefit" part	Founded redundancy clauses should also be marked with a hash symbol, if the benefit part contains more than one redundancy clause as Targets	Check the behaviour of the *highlightRedundancies* method if there is more than one redundancy clause as Targets in benefit part

Table 12: Test cases for ReportExtractor class

Table 13 shows the test cases for the Evaluation.java class and Figure 38 shows the code coverage.

Test Case	Supplied Data	Expected Outcome	Description
testEmptyOrNotExist JsonFile	Provision of a dummy JSON file	Through an exception:*EmptyOrNotExist JsonFile*	Check whether the provided JSON file already exists and is not empty
testJsonFile	Provision of a JSON file which is not accessible	Through an exception:*IOException*	Check whether the provided JSON file is accessible

Test Case	Supplied Data	Expected Outcome	Description
main/benefitFull_Case1	Provide a JSON file with US-pair that have no common textitContains entries but have common *Targets* entries in main/benefit part	In JSON report the entry "Main/Benefit Part Fully Redundant" should be *true*	Check whether US-pair without *Contains* entries can also be fully redundant in main/benefit part
main/benefitFull_Case2	Provide a JSON file with US-pair that have not common textitContains entries but one US has entries as Contains in main/benefit	In JSON report the entry "Main/Benefit Part Fully Redundant" should be *true*	Check whether US-pair without common *Contains* entries can also be fully redundant in main/benefit part
main/benefitFull_Case3	Provide a JSON file with US-pair that have common textitContains entries but one US have additional entries as Contains in main/benefit part	In JSON report the entry "Main/Benefit Part Fully Redundant" should be *true*	Check whether USs with additional *Contains* entries from common contains can also be fully redundant in main/benefit part
main/benefitFull_Case4	Provide a JSON file with US-pair that have exactly common textitContains entries as well as common *Targets* in main/benefit part	In JSON report the entry "Main/Benefit Part Fully Redundant" should be *true*	Check whether USs with *Contains* and *Targets* entries can also be fully redundant in main/benefit part
mainBenefitFull	Provide a JSON file with US-pair that all clauses in the main and benefit parts are identical	The JSON report the entries "Main Part Fully Redundant" and "Benefit Part Fully Redundant" should be *true*	Check that all clauses in the main and benefit part are identical, main and benefit part are evaluated as fully redundant
mainPartial_Case1	Provide a JSON file with a US-pair whose entries in *Targets* are partially identical	In the JSON report, the entry "Main Part Partially Redundant" should be *true*	Check whether USs with partially redundant entries in *Targets* are evaluated as partially redundant
mainPartial_Case2	Provide a JSON file with a US-pair whose entries in *Contains* are partially identical	In the JSON report, the entry "Main Part Partially Redundant" should be *true*	Check whether USs with partially redundant entries in *Contains* are evaluated as partially redundant
benefitParial	Provide a JSON file with US-pair where only some clauses in *Targets* and *Contains* entries are common	In the JSON report, the entry "Benefit Part Partially Redundant" should be *true*	Check whether USs with partially common *Targets* and *Contains* entries are evaluated as partially redundant
mainFullBenefitPartial	Provide a JSON file with US-pair where all clauses in the main part and some in the benefit part are identical	In the JSON report, the entry "Main Part Fully Redundant" and "Benefit Part Partially Redundant" should be *true*	Check whether the main part is evaluated as fully redundant if all clauses are identical and the benefit part is evaluated as partially redundant if some clauses are identical

Test Case	Supplied Data	Expected Outcome	Description
mainPartialBenefitFull	Provide a JSON file with US-pair where some clauses in the main part and all in the benefit part are identical	In the JSON report, the entry "Main Part Partially Redundant" and "Benefit Part Fully Redundant" should be *true*	Check whether the main part is evaluated as partially redundant if some clauses are identical and the benefit part is evaluated as fully redundant if all clauses are identical

Table 13: Test cases for Evaluation class

4.5   Evaluation

The purpose of this section is to assess the effectiveness of our approach in analysing redundancy between USs. This section serves several key objective:

- Demonstrating the validity and reliability of our proposed methodology for identifying and categorising redundancies between USs.

- Presenting the results of our analysis across multiple datasets, highlighting the prevalence of "Full and Partial Redundancy" for both the main and benefit parts of USs.

- Gaining insights into the nature and extent of redundancies between USs and its potential impact on software development processes.

- Comparing results from different data sets, identifying patterns or trends and analysing variations in redundancy levels.

Methodology

In this section, we recap the methodology employed to analyse redundancy between USs. We utilized a systematic approach that involved several key steps:

- Data Collection: We collected a total 19 datasets containing USs from various software development projects.

- Preprocessing: Using the generated JSON report file for each dataset, we create a VBA script called *extractFromJSONFiles*[22] to iterate through all the JSON report files and extract the information such as "US-pair", "US-text", "total redundancy clauses", "main redundancy clauses", "benefit redundancy clauses" and "project number" and collect them all in an Excel file to perform further analyses.

---

[22]https://github.com/amirrabieyannejad/USs_Annotation/tree/main/Skript/extractFromJSONFiles

ReportExtractor

Element	Missed Instructions	Cov.	Missed Branches	Cov.	Missed	Cxty	Missed	Lines	Missed	Methods
splitUsText(String, String, String, String, FileWriter, JSONObject)		100 %		100 %	0	9	0	27	0	1
reportExtractor(String[])		100 %		100 %	0	2	0	15	0	1
createTable(List, List)		100 %		100 %	0	4	0	13	0	1
getAbsoluteDirPath()		100 %		100 %	0	6	0	15	0	1
getUssTexts(String, RedundancyItems)		100 %		100 %	0	4	0	13	0	1
hasActions(ArrayList, RedundancyItems)		100 %		100 %	0	6	0	8	0	1
hasEntitys(ArrayList, RedundancyItems)		100 %		100 %	0	6	0	8	0	1
applyHashSymbols(String, String[])		100 %		100 %	0	2	0	4	0	1
getTotalRedundanciesFromPair(List, String, String)		100 %		100 %	0	6	0	6	0	1
checkIfReportExist(String, List)		100 %		100 %	0	2	0	7	0	1
hasTargets(ArrayList, RedundancyItems)		100 %		100 %	0	4	0	5	0	1
minimalEcoreExist(String, String)		100 %		100 %	0	2	0	4	0	1
getAllTargetsInBenefit(String, String, String, JSONObject)		100 %		100 %	0	2	0	8	0	1
getAllContainsInBenefit(String, String, String, JSONObject)		100 %		100 %	0	2	0	8	0	1
getAllTargetsInMain(String, String, String, JSONObject)		100 %		100 %	0	2	0	7	0	1
getAllContainsInMain(String, String, String, JSONObject)		100 %		100 %	0	2	0	7	0	1
getAbsoluteFinalReportDir()		100 %		100 %	0	2	0	4	0	1
readJsonArrayFromFile(String)		100 %		100 %	0	2	0	6	0	1
processElements(String, String, String, JSONObject)		100 %		100 %	0	3	0	4	0	1
getFinalReportDir(Path)		100 %		100 %	0	2	0	4	0	1
hasMoreThanFourHashSymbols(String)		100 %		100 %	0	2	0	2	0	1
containsAnd(String)		100 %		100 %	0	2	0	3	0	1
lambda$0(int)		100 %		100 %	0	2	0	1	0	1
lambda$2(int)		100 %		100 %	0	2	0	1	0	1
highlightRedundancies(RedundancyItems, String, List, List, List, JSONObject)		100 %		87 %	3	13	0	104	0	1
getCommonTargets(String, JSONObject)		100 %		87 %	4	17	0	69	0	1
getCommonContains(String, JSONObject)		100 %		87 %	4	17	0	68	0	1
writeTable(File, List)		100 %		87 %	2	9	0	28	0	1
iteratePackages(EPackage, ArrayList, ArrayList, RedundancyItems)		100 %		85 %	4	16	0	32	0	1
getCommonTriggers(String)		100 %		81 %	4	12	0	27	0	1
extractReports(FileWriter, FileWriter)		100 %		80 %	6	16	0	59	0	1
applyHashSymbolTargetsMain(List, List, RedundancyItems, String, String, JSONObject)		100 %		75 %	3	7	0	27	0	1
applyHashSymbolContaiansMain(List, RedundancyItems, String, String, JSONObject)		100 %		70 %	3	6	0	21	0	1
applyHashSymbolTargetsBenefit(List, List, RedundancyItems, String, String, JSONObject)		93 %		66 %	4	7	2	26	0	1
applyHashSymbolPersona(List, RedundancyItems, String, String)		100 %		66 %	2	4	0	13	0	1
processMinimalModels(File, ArrayList, ArrayList, RedundancyItems)		100 %		60 %	4	6	0	11	0	1
applyHashSymbolContaiansBenefit(List, RedundancyItems, String, String, JSONObject)		92 %		58 %	5	7	2	23	0	1
createOrOverwriteReportFile(File)		63 %		50 %	1	2	2	6	0	1
hasMoreThanSixHashSymbols(String)		85 %		50 %	1	2	0	2	0	1
main(String[])		0 %		n/a	1	1	2	2	1	1
writeUsText(String, ArrayList, List, RedundancyItems, FileWriter, List, List, List, JSONObject)		100 %		n/a	0	1	0	17	0	1
getRedundancyStatus(RedundancyItems)		100 %		n/a	0	1	0	5	0	1
writeUsSentencePart(String, RedundancyItems, FileWriter, JSONObject)		100 %		n/a	0	1	0	6	0	1
getClassName(String)		100 %		n/a	0	1	0	2	0	1
ReportExtractor(String, String)		100 %		n/a	0	1	0	4	0	1
getRefName(String)		100 %		n/a	0	1	0	2	0	1
getAttName(String)		100 %		n/a	0	1	0	2	0	1
getUsName1(String)		100 %		n/a	0	1	0	1	0	1
getUsName2(String)		100 %		n/a	0	1	0	1	0	1
getDirName()		100 %		n/a	0	1	0	1	0	1
getJsonDatasetFile()		100 %		n/a	0	1	0	1	0	1
Total	35 of 3.863	99 %	50 of 357	85 %	51	231	8	738	1	51

Figure 37: Code coverage related to class ReportExtractor

- Identification of main and benefit parts: Each US was divided into its main part, which is the core functionality, and its benefit part, which describes the value to the persona.

- Redundancy analysis: A separate redundancy analysis was carried out between the main parts of US-pair and the benefit parts of US-pair. The redundancy analysis was categorised as either "Full or Partial Redundancy".

Dataset of Backlogs

For a comprehensive assessment, we applied our approach to 19 backlog datasets presented by Mosser et al.[23]. They applied the CRF approach to these publicly available requirements datasets[10].

---

[23]https://github.com/ace-design/nlp-stories

Evaluation

Element	Missed Instructions	Cov.	Missed Branches	Cov.	Missed	Cxty	Missed	Lines	Missed	Methods
evaluateRedundancyCriteria(JSONObject, String, String)		98 %		79 %	30	76	0	99	0	1
checkFullRedaundat(JSONArray, JSONArray)		94 %		70 %	3	6	0	12	0	1
checkPartialRedundancy(JSONArray, JSONArray)		100 %		100 %	0	8	0	14	0	1
readJsonArrayFromFile(String)		100 %		100 %	0	2	0	8	0	1
getJSONArraySafely(JSONObject, String)		100 %		100 %	0	2	0	1	0	1
Evaluation()		100 %		n/a	0	1	0	2	0	1
Total	8 of 613	98 %	34 of 178	80 %	33	95	0	136	0	6

Figure 38: Code coverage related to class Evaluation

It is also worth noting that some backlog datasets (g02, g13, g17, g27) did not follow the expected sentence structure, which is why we did not include them in the evaluation results to avoid unexpected behaviour. Table 14 shows the project number of each dataset and the count of USs.

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
Project Nr.	g03	g04	g05	g08	g10	g11	g12	g14	g16	g18	g19	g21	g22	g23	g24	g25	g26	g27	g28	Total USs
Total USs	57	51	53	66	97	73	54	67	66	102	137	69	83	56	53	100	100	114	60	1458

Table 14: Project number and count of USs contained in each backlog dataset

Assessment of Result: High-Level Overview

Based on the datasets provided, it was found that 1,851 redundancy clauses were identified across all projects, with the highest count found in the backlog G19 dataset (925 clauses), indicating a significant presence of redundancy in the USs of this project. Table 15 shows the total number of redundancy clauses found in each dataset.

The main parts of the USs have less redundancy (17 clauses), which means that there are 17 cases where the main part of the US-pairs is exactly the same. This level of redundancy indicates common functionality of the US-pairs, which is a sign of overlapping features.

A much higher occurrence of partial redundancy in the main parts indicates that there are common elements between the US-pairs, but still enough differences to avoid a complete match.

Fewer instances of partial redundancy in the benefit parts (46 clauses) indicate that the USs diverge in their benefit clauses, they are more tailored to specific project outcomes.

The higher count of full redundancies compared to partial redundancies in the benefit parts is interesting as it indicates that the expected objectives of certain features are often repeated in the USs. This suggests that the project is aiming for a set of common goals regardless of the specific functionality.

Table 15 shows the amount of full and partial redundancies in the main and benefit parts.

Project Nr.	Total Redundancy Clause	Main Part Redundancy Clause	Benefit Part Redundancy Clause	Main Part Redundancy Clause (Partial)	Main Part Redundancy Clause (Full)	Benefit Part Redundancy Clause (Partial)	Benefit Part Redundancy Clause (Full)
G03	2	0	2	0	0	2	0
G04	6	4	2	2	0	2	0
G05	61	30	31	15	0	2	14
G08	95	31	64	18	0	5	18
G10	30	25	5	14	0	1	2
G11	20	11	9	8	0	10	0
G12	6	4	2	2	0	0	2
G14	18	11	7	5	2	6	1
G16	5	5	0	2	1	0	0
G18	283	274	9	157	6	3	2
G19	925	925	0	512	0	0	0
G21	31	24	7	13	1	2	3
G22	105	67	38	26	3	9	13
G23	12	12	0	6	0	0	0
G24	32	24	8	10	0	1	4
G25	89	89	0	38	2	0	0
G26	19	18	1	9	0	1	0
G27	22	19	3	8	1	2	1
G28	90	90	0	30	1	0	0
Grand Total	1851	1663	188	875	17	46	60

Table 15: Detail about full and partial redundancies related to main or benefit parts

Assessment of Result: Detailed Insights

In the dataset of Backlog G19, the clause "have alfred" was repeated in all USs, indicating that "alfred" is the end product and not a specific functionality. This suggests that the quality of the USs should be considered, so that they do not contain generic information that results in a lot of redundant information(false positives) being provided.

In the main parts of the USs, 1,663 clauses were redundant, while only 188 redundancy clauses were found in the benefit parts. This means that some USs are not often written in a standardised format for the description of core functions, resulting in a higher count of redundant clauses in the main part.

Example 4.17. *In the dataset of backlog G18 (274 clauses out of 283 in the main parts), for example, this is due to the clauses "have ability" in the main part, which are unnecessarily repeated in most USs:*

user_story_92: #g18# as a #researcher# I want to have #the #ability# to search files by file type and format.

user_story_80: #g18# as a #researcher# I want to have #the #ability# to attach standard metadata for behavioural observations (and video) so that my data can be searched and understood later.

Actually, the clause "have ability" in USs should be deleted and no redundancy should be possible at all:

user_story_92: #g18# as a researcher I want to search for files by file type and format.

user_story_80: #As a researcher, I want to attach standard metadata for behavioural observations (and videos), so that my data can be searched and understood later.

There are also USs with the same functionality, but one provides more details about the functionality. In other words, one US is contained within another, and we refer to them as redundant US-pairs, as deleting the US with less detail has no negative impact on the system. Sometimes it is necessary to merge these two USs to obtain a single detailed US.

Example 4.18. *In the dadaset of backlog G18, for example, we have two US-pairs that are marked as full redundancy between user_story_12 and user_story_11 as well as user_story_13 and user_story_11:*

user_story_12: "#g18# as a #researcher#, I want to #upload# #files# prior to having them #attached# to a #log book page# using the web interface."

user_story_11: "#g18# as a #researcher#, I want to #upload# #files# prior to having them #attached# to a #log book page#."

user_story_13: "#g18# as a #researcher#, I want to #upload# #files# prior to having them #attached# to a #log book page# using a mapped network drive."

As we can see, the user_story_11 is an incomplete version of two other USs, and deleting it has no negative impact in the system, due to the fact that its goal is achieved and fulfilled by two other USs.

Assessment of Result: Aggregated Analysis

We also look for an aggregate for the count of redundant clauses in both the main and the benefit parts of the USs.

In the main parts, the large count of full redundancies, especially where two clauses are present (6 clauses), indicates a high degree of similarity in the functionalities described.

In the benefit part of the USs, there is a considerable count of cases where full redundancy prevails, especially where two clauses are inserted (24 clauses). The large count of full redundancies with two clauses indicates that benefits are often described in a standardised way in the different USs.

Figure 39 illustrates the aggregation of the count of redundancy clauses that occur in USs.
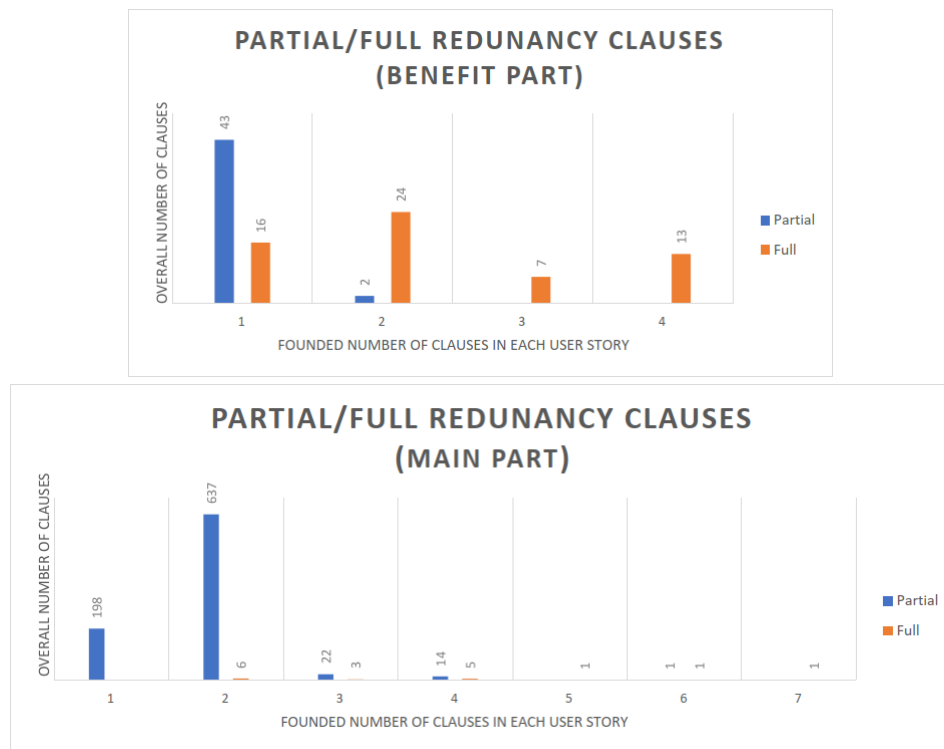

Figure 39: The aggregation between the count of redundancy clauses occured in USs

4.6   Conclusion

In this study, we developed and applied a comprehensive approach that combines the CRF tool, Henshin and the CDA tool to systematically identify and report redundancies in USs within software development projects.

By carefully analysing 19 different backlog datasets, our method not only separated USs into main and benefit parts for nuanced examination, but also facilitated the distinction between full and partial redundancies within these parts.

Our results reveal a crucial finding: the effectiveness of redundancy detection is significantly influenced by the quality of the USs. Well-formulated USs that do not contain unnecessary clauses(e.g. the repetition of "end product" in all USs) significantly improve the accuracy of redundancy detection. This emphasises the importance of rigorous US formulation as a prerequisite for effective redundancy analysis.

If the main part is evaluated to be full redundant, then we have a US-pair that is functionally identical and we can merge the US-pair into one US. In the case of full redundancy in the benefit part, this means that the US-pairs belong to the same goal and objective, to which they should be categorised for better accessibility and understanding.

A notable trend emerged from our analysis: the benefit parts are more often fully redundant than the main parts of USs. This indicates that multiple USs often strive for different functions that contribute to a common system aspect or goal.

Recognising such redundancies not only helps to consolidate functionally identical US-pairs into single, more compact USs, but also to group US-pairs together, improving accessibility and understanding of project backlogs.

In summary, our study confirms the central role of a syntactic analysis approach in detecting and managing redundancies in project backlogs, thereby contributing to the rationalisation of software development processes.

While the quality of USs plays a critical role in the success of this approach, the insights gained from this research provide valuable guidance for both current practices and future research in software project management.

However, our investigation has also shown that we can only consider USs with syntactic redundancy. If they are indeed US-pairs with the same functionality but using different words and clauses to achieve the same goal, we cannot detect this with this approach.

This finding shows that the distinction between actual redundancy and mere superficial similarity can be further refined, which leads us to analyse the USs semantically.

In the next section, we therefore present a method for analysing conflicts and dependencies between USs in semantic way.

References

[1] A. Al-Hroob, A. T. Imam, and R. Al-Heisa. The use of artificial neural networks for extracting actions and actors from requirements document. *Information and Software Technology*, 101:1–15, 2018.

[2] T. Arendt, E. Biermann, S. Jurack, C. Krause, and G. Taentzer. Henshin: advanced concepts and tools for in-place EMF model transformations. In *Model Driven Engineering Languages and Systems: 13th International Conference, MODELS 2010, Oslo, Norway, October 3-8, 2010, Proceedings, Part I 13*, pages 121–135. Springer, 2010.

[3] M. Arias, A. Buccella, and A. Cechich. A Framework for Managing Requirements of Software Product Lines. *Electronic Notes in Theoretical Computer Science*, 339:5–20, July 2018. ZSCC: 0000025.

[4] S. Arulmohan, M.-J. Meurs, and S. Mosser. Extracting Domain Models from Textual Requirements in the Era of Large Language Models. MDEIntelligence (colocated with ACM/IEEE 26th International Conference on . . . , 2023.

[5] C. F. Baker, C. J. Fillmore, and J. B. Lowe. The berkeley framenet project. In *COLING 1998 Volume 1: The 17th International Conference on Computational Linguistics*, 1998.

[6] D. M. Berry and E. Kamsties. Ambiguity in requirements specification. In *Perspectives on software requirements*, pages 7–44. Springer.

[7] K. Born, T. Arendt, F. Heß, and G. Taentzer. Analyzing conflicts and dependencies of rule-based transformations in Henshin. In *Fundamental Approaches to Software Engineering: 18th International Conference, FASE 2015, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2015, London, UK, April 11-18, 2015, Proceedings 18*, pages 165–168. Springer, 2015.

[8] E. Cambria and B. White. Jumping NLP curves: A review of natural language processing research. *IEEE Computational intelligence magazine*, 9(2):48–57, 2014.

[9] P. Carlshamre, K. Sandahl, M. Lindvall, B. Regnell, and J. N. och Dag. An industrial survey of requirements interdependencies in software product release planning. In *Proceedings Fifth IEEE International Symposium on Requirements Engineering*, pages 84–91. IEEE, 2001.

[10] F. Dalpiaz. Requirements Data Sets (User Stories), 2018.07.28.

[11] F. Dalpiaz, A. Ferrari, X. Franch, and C. Palomares. Natural Language Processing for Requirements Engineering: The Best Is Yet to Come. 35:115–119, 2018.

[12] A. Gomez, G. Rueda, and P. P. Alarcón. A systematic and lightweight method to identify dependencies between user stories. In *Agile Processes in Software Engineering and Extreme Programming: 11th International Conference, XP 2010, Trondheim, Norway, June 1-4, 2010. Proceedings 11*, pages 190–195. Springer, 2010.

[13] D. Greer and G. Ruhe. Software release planning: an evolutionary and iterative approach. 46:243–253, 2004.

[14] A. Habel and K.-H. Pennemann. Correctness of high-level transformation systems relative to nested conditions. *Mathematical Structures in Computer Science*, 19(2):245–296, 2009.

[15] A. C. D. H. J. Hartmanis, T. Henzinger, J. H. N. J. T. Leighton, and M. Nivat. Monographs in Theoretical Computer Science An EATCS Series. 2006.

[16] J. H. Hausmann, R. Heckel, and G. Taentzer. Detection of conflicting functional requirements in a use case-driven approach: a static analysis technique based on graph transformation. In *Proceedings of the 24th international conference on software engineering*, pages 105–115, 2002.

[17] P. Heck and A. Zaidman. A quality framework for agile requirements: a practitioner's perspective. *arXiv preprint arXiv:1406.4692*, 2014.

[18] H. Kastenberg and A. Rensink. Model checking dynamic states in GROOVE. In *Model Checking Software: 13th International SPIN Workshop, Vienna, Austria, March 30-April 1, 2006. Proceedings 13*, pages 299–305. Springer, 2006.

[19] K. Kipper, A. Korhonen, N. Ryant, and M. Palmer. Extending VerbNet with Novel Verb Classes. In *LREC*, pages 1027–1032, 2006.

[20] A. Korhonen and T. Briscoe. Extended lexical-semantic classification of English verbs. In *Proceedings of the Computational Lexical Semantics Workshop at HLT-NAACL 2004*, pages 38–45, 2004.

[21] J. D. Lafferty, A. McCallum, and F. C. N. Pereira. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. In C. E. Brodley and A. P. Danyluk, editors, *Proceedings of the Eighteenth International Conference on Machine Learning (ICML 2001), Williams College, Williamstown, MA, USA, June 28 - July 1, 2001*, pages 282–289. Morgan Kaufmann, 2001.

[22] B. Levin. *English verb classes and alternations: A preliminary investigation*. University of Chicago press, 1993.

[23] O. I. Lindland, G. Sindre, and A. Solvberg. Understanding quality in conceptual modeling. *IEEE software*, 11(2):42–49, 1994.

[24] O. Liskin, R. Pham, S. Kiesling, and K. Schneider. Why we need a granularity concept for user stories. In *Agile Processes in Software Engineering and Extreme Programming: 15th International Conference, XP 2014, Rome, Italy, May 26-30, 2014. Proceedings 15*, pages 110–125. Springer, 2014.

[25] K. Logue and K. McDaid. Handling uncertainty in agile requirement prioritization and scheduling using statistical simulation. In *Agile 2008 Conference*, pages 73–82. IEEE, 2008.

[26] G. Lucassen, F. Dalpiaz, J. M. E. van der Werf, and S. Brinkkemper. Improving agile requirements: the quality user story framework and tool. *Requirements engineering*, 21:383–403, 2016.

[27] T. Mens, G. Taentzer, and O. Runge. Analysing refactoring dependencies using graph transformation. *Software & Systems Modeling*, 6:269–285, 2007.

[28] G. A. Miller. WordNet: a lexical database for English. *Communications of the ACM*, 38(11):39–41, 1995.

[29] M. Moens, M. Steedman, et al. Temporal ontology and temporal reference., 2005.

[30] S. Mosser, C. Pulgar, and V. Reinharz. Modelling Agile Backlogs as Composable Artifacts to support Developers and Product Owners. *Journal of Object Technology (JOT)*, 2022.

[31] E. Paja, F. Dalpiaz, and P. Giorgini. Managing security requirements conflicts in socio-technical systems. In *Conceptual Modeling: 32th International Conference, ER 2013, Hong-Kong, China, November 11-13, 2013. Proceedings 32*, pages 270–283. Springer, 2013.

[32] C. A. Prolo. Generating the XTAG English grammar using metarules. In *COLING 2002: The 19th International Conference on Computational Linguistics*, 2002.

[33] I. K. Raharjana, D. Siahaan, and C. Fatichah. User stories and natural language processing: A systematic literature review. *IEEE access*, 9:53811–53826, 2021.

[34] W. N. Robinson. Integrating multiple specifications using domain goals. *ACM SIGSOFT Software Engineering Notes*, 14(3):219–226, 1989.

[35] T. Sedano, P. Ralph, and C. Péraire. The product backlog. In *2019 IEEE/ACM 41st International Conference on Software Engineering (ICSE)*, pages 200–211. IEEE, 2019.

[36] D. Strueber. Henshin/Conflict and Dependency Analysis, 2023.02.11.

[37] M. Tichy, C. Krause, and G. Liebel. Detecting performance bad smells for henshin model transformations. *Amt@ models*, 1077, 2013.

[38] G. Varró, F. Deckwerth, M. Wieber, and A. Schürr. An algorithm for generating model-sensitive search plans for EMF models. In *Theory and Practice of Model Transformations: 5th International Conference, ICMT 2012, Prague, Czech Republic, May 28-29, 2012. Proceedings 5*, pages 224–239. Springer, 2012.

[39] Y. Wautelet, S. Heng, S. Kiv, and M. Kolp. User-story driven development of multi-agent systems: A process fragment for agile methods. 2017.

[40] Y. Wautelet, S. Heng, M. Kolp, and I. Mirbel. Unifying and extending user story models. In *Advanced Information Systems Engineering: 26th International Conference, CAiSE 2014, Thessaloniki, Greece, June 16-20, 2014. Proceedings 26*, pages 211–225. Springer, 2014.

[41] M. Wynne, A. Hellesoy, and S. Tooke. *The cucumber book: behaviour-driven development for testers and developers.* Pragmatic Bookshelf, 2017.