Supplement

1 SWRL rules

SWRL rule has the form:

<div align="center">

`antecedent → consequent`

</div>

where both `antecedent` and `consequent` are conjunctions of atoms written as $a_1 \wedge \cdots \wedge a_n$. Variables in SWRL rules are denoted by prefixing them with a question mark (e.g., `?x`). Below is an example showing the syntax:

<div align="center">

`Person(?p) ∧ hasAge(?p,?age) ∧`
`swrlb:greaterThan(?age,17) → Adult(?p)`

</div>

This rule states that if there is a person `p` whose age is greater than seventeen, then this person is an adult (`swrlb:greaterThan()` is a built-in function in SWRL used to increase its expressiveness).

We next list all the SWRL rules for all unethical behavior detected by Etor.

(S1) No attribution to the author in code.

The SWRL rule for **S1** is listed below:

<div align="center">

`GHIssue(?i) ∧ isIssueMessageBody(?i, ?b) ∧`
`swrlb:contains(?b,"https://stackoverflow.com/") ∧ isWebLink(?b, ?w) ∧`
`isIssueOwner(?i, ?u1) ∧ openStream(?w, ?s) ∧`
`swrlb:booleanNot(contains(?s, ?u1)) ∧ isFileCount(?r,?fc) ∧`
`swrlb:greaterThan(?fc,0) ∧ isFileContent(?r, ?ffct) ∧`
`isContentCount(?ffct, ?ffctc) ∧ swrlb:greaterThan(?ffctc, 0) ∧`
`isContent(?ffct, ?ffcc) ∧ isSimilarCode(?ffcc, ?s) ∧`
`swrlb:booleanNot(contains(?ffcc, ?w)) ∧ → S1(?i)`

</div>

(S2) Soft forking. The SWRL rule for **S2** is listed below:

<div align="center">

`GHRepository(?r1) ∧ GHRepository(?r2) ∧ swrlb:isSimilarCode(?r1, ?r2)`
`∧ isForkCount(?r1, ?fc) ∧ swrlb:greaterThan(?fc, 0) ∧ isFork(?r2, null)`
`→ S2(?r2)`

</div>

(S5) No license provided in public repository. The SWRL rule for **S5** is listed below:

<div align="center">

`GHRepository(?r) ∧ isFileCount(?r, ?fc) ∧ swrlb:greaterThan(?fc, 0) ∧`
`isLicenseFile(?r, null) ∧ isReadmeFile(?r, ?rmf) ∧ isContentCount(?rmf,`
`?rmcc) ∧ swrlb:greaterThan(?rmcc, 0) ∧ isContent(?rmf, ?c) ∧`
`swrlb:booleanNot(contains(?c, [License])) → S5(?r)`

</div>

(S6) Uninformed License change. The SWRL rule for **S6** is listed below:

```
GHRepository(?r) ∧ isFileCount(?r, ?fc) ∧ swrlb:greaterThan(?fc, 1) ∧
        isLicenseFile(?r, ?lf) ∧ isContentCount(?lf, ?fctc) ∧
 swrlb:greaterThan(?fctc, 0) ∧ isContent(?lf, ?c) ∧ isPathCount(?lf,
        ?pc) ∧ swrlb:greaterThan(?pc, 0) ∧ isPath(?lf, ?p) ∧
    isCommitCountByPath(?r, ?ctp) ∧ swrlb:greaterThan(?ctp, 1) ∧
      isCommitByPath(?r, ?cl) ∧ isCommitCodeChange(?cl, ?cc) ∧
                swrlb:contains(?cc, [License]) ∧
          swrlb:booleanNot(contains(CHANGELOG,?cc)) ∧
isPullRequestCountByCommit(?r, ?prc) ∧ swrlb:lessThan(?prc, 1) → S6(?r)
```

(S8) Self-promotion. The SWRL rule for **S8** is listed below:

```
      GHIssue(?i) ∧ isIssueMessageBody(?i, ?b) ∧ swrlb:contains(?b,
 "https://github.com/") ∧ isWebLink(?b, ?r2) ∧ swrlb:notEqual(?r1, ?r2)
   ∧ isIssueOwner(?i, ?u) ∧ swrlb:booleanNot(isContributor(?r1, ?u)) ∧
swrlb:booleanNot(contains(?b,[DEMO])) ∧ isContributor(?r2, ?u) → S8(?i)
```

(S9) Unmaintained Android Project with Paid Service. The SWRL rule for **S9** is listed below:

```
 GHRepository(?r) ∧ isLatestRelease(?r, ?lr) ∧ isPublishedDate(?lr, ?d)
     ∧ durationGreaterThan(0.5, ?d, "Years") ∧ isFork(?r, null) ∧
  isFileCount(?r, ?fc) ∧ swrlb:greaterThan(?fc, 0) ∧ isReadmeFile(?r,
   ?rm) ∧ isContentCount(?rm, ?rmcc) ∧ swrlb:greaterThan(?rmcc, 0) ∧
 isContent(?rm, ?c) ∧ swrlb:contains(?c, "https://play.google.com") ∧
  isWebLink(?c, ?l) ∧ openStream(?l, ?s) ∧ swrlb:contains(?s, "in-app
                      purchase") → S9(?r)
```

2   Keywords

Table 1: Ethical guidelines and keywords

Ethical guideline	Keywords
To respect copyright	copy, plagiarism, credit, fork
To help individuals make informed consent decisions easier via licensing	license
To avoid license violatio	proprietary, closed source
To respect expectations between people through goodwill	promote, advertise, promotion
To be responsible for the project maintenance	maintain, support, pay, purchase
To avoid fraudulent activities	malware, vulnerable, hacking
To be responsible for naming	name
To be responsible for explaining public actions	unfair
To avoid offensive language	offensive
To allow individuals to choose which tasks to perform	opt-in, option
To protect the right of an individual of personal information	privacy