Motivation: Learning to relate comment and code elements is critical to automated systems for generating code and comments and detecting inconsistent code and comments.

Task: Given a noun phrase (NP) in a comment, we classify the relationship between the NP and each candidate code token in the corresponding method as either associated or not associated.

Challenge #1: Obtaining labeled data

Noisy supervision: Mining simultaneous code/comment updates from GitHub commits

GitHub

```java
/** @return the highest value from the list of scores */
public int getScore() {
    return Collections.max(scores);
}
```

```java
/** @return the lowest value from the list of scores */
public int getScore() {
    return Collections.min(scores);
}
```

Challenge #2: Reasoning across two distinct language representations

Salient features: Capturing characteristics of code, comments, and the relationship between them