|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Data Extraction Form** | | | | | | | | | | | | | | | | | | | | | | | |
| **Title** | Augmenting automated game testing with deep reinforcement learning | | | | | | | | | | **Authors(s)** | | Joakim Bergdahl, Camilo Gordillo, Konrad Tollmar, Linus Gissl´en | | | | | | | | | | |
| **Year** | 2020 | | | | | | | | | | **Venue** | | * Journal | | * **Conference** | | | | | * Other \_\_\_\_\_\_\_\_\_\_ | | | |
| **Quality Assessment criteria** | | | | * **QC1** | | | | * QC2 | | | | | * QC3 | | * QC4 | | | | | * QC5 | | | * **QC6** |
| **Inclusion Criteria** | | | | * IC1 | | | | * IC2 | | | | | * IC3 | | | * **IC4** | | | | * IC5 | | | |
| **Exclusion Criteria** | | | | * EC1 | | | * EC2 | | * EC3 | | | | * EC4 | * EC5 | | | | | * EC6 | | * EC7 | | |
| **Approach Used**   * Supervised Machine Learning algorithms * Unsupervised Machine Learning algorithms * Natural language processing * **Deep Learning algorithms (Deep enforcement learning DRL’s)** * Data mining based techniques (relevant   characteristics of BRs and their fixes such as opening time, number of comments, number of times the BR is reopened, number of change sets for BR and the number of files changed and lines modified for fixes or patch)   * Statistical Method (Pearson’s chi-squared test) * Other | | | | | | | | | | **Type of Solution** | | | | | | | | **Yes** | | **No** | | **Unclear** | |
| Novel Technique (Method, Tool, Technique) | | | | | | | | Check mark, Wingdings font, character code 252 decimal. | |  | |  | |
| Evaluation of existing techniques  (Evaluation framework, tool, platform) | | | | | | | |  | | Check mark, Wingdings font, character code 252 decimal. | |  | |
| Supporting techniques | | | | | | | |  | |  | | Check mark, Wingdings font, character code 252 decimal. | |
| **Review dataset** | | **Total number of apps** | | | | N/A | | | | **Evaluation Method Used** | | | | | | | | Close collaboration between researchers and game testers | | | | | |
| **Total number of crawled reviews** | | | | N/A | | | |
| **Year** | | | | 2020 | | | |
| **Research Type Facet**   * Validation Research * **Evaluation Research** * Solution Proposal * **Philosophical Papers** * Opinion Papers * Experience Papers | | | | | **Solution Type**   * **Single** * Hybrid/Integrated | | | | | | | **Contribution**   * Technique * Tool * Comparison * Model * Framework * Prototype * **Taxonomy** | | | | | **Evaluation Strategy**   * **Case Study** | | | | | | |
| **Features used**   * Categorical * **Textual** * Both. | | | | | | |
| **Factors Considered** | | | Navigation of FPS type games, game exploits and bugs, distribution of visited states, and difficulty evaluation | | | | | | | | | | | | | | | | | | | | | |
| **Notes** | | |  | | | | | | | | | | | | | | | | | | | | | |
| **Limitations** | | | * Limited to FPS games * Not all problems can be solved by RL(Reinforcement learning) technique including bugs classification and categorization | | | | | | | | | | | | | | | | | | | | | |
| **Description / Summary** | | | Deep enforcement learning (DRL) can be used to increase test coverage, find exploits, test map difficulty, and to detect common problems that arise in the testing of first-person shooter (FPS) games. RL is better suited for modular integration where it can complement rather than replace existing techniques. Not all problems are better solved with RL and training is substantially easier when focusing on single, well isolated tasks. RL can complement scripted tests in edge cases where human-like navigation, exploration, exploit detection and difficulty evaluation is hard to achieve. | | | | | | | | | | | | | | | | | | | | | |