

Évaluation différentielle des projets de programmation

D. Renault
ENSEIRB-Matmeca
1er juillet 2026

- Des étudiants doivent réaliser un projet de programmation.
 - ▶ le S5 à l'ENSEIRB-Matmeca sur 8 semaines,
- Ex. :
 - ▶ ~ 45 équipes de 2 étudiants,
 - ▶ 8 encadrants + un grand Manitou
- On désire évaluer **uniformément** leurs capacités de programmation.
 - ▶ Pas évident avec de multiples encadrants,
 - ▶ Pas évident avec des LLMs qui traînent.
- On désire aussi respecter l'esprit du développement de **projet**.
 - ▶ Évaluation sur un travail en équipe,
 - ▶ Développement sur un temps "long" (> exercice de TD),
 - ▶ Minimum exigé de réflexion / conception
- On désire que cela reste (relativement) **simple à corriger**.

Idée générale

- ❶ Prévoir le sujet du projet avec une possibilité d'**extension**.
- ❷ Terminer les projets une semaine avant la date originale.
- ❸ Faire que les élèves aient à ce moment là un code “propre”.
- ❹ Utiliser la séance supplémentaire pour leur faire **réaliser l'extension**.
- ❺ Évaluer le travail sur le **différentiel** sur cette séance.

Le projet du S5 à l'ENSEIRB-Matmeca

- Les sujets des dernières années :
 - ▶ Royal Dédale
 - ▶ Bourgade
 - ▶ Sandwich
 - ▶ Naval Batal
- Conçus en deux parties distinctes :
 - ▶ Une version de base (avec pseudo-code et décomposition des tâches) ;
 - ▶ Un système d'extensions du sujet (nommés **achievements**).
- Méthode :
 - ▶ Initialement, seule la version de base est visible ;
 - ▶ Les achievements sont rendus publics au fur et à mesure des projets.
- L'idée est de pousser les élèves à reprendre / modifier / étendre des composants de leur code.

Exemple d'extension (pendant le projet)



- On part d'un graphe vide $H \times W$;
- On ajoute à chaque tour une tuile de manière à rester connexe ;
- Les tuiles sont tirées d'un ensemble de base.

Version de base : l'ajout se fait par translation

Extension : on autorise aussi les rotations

Exemple d'extension (séance finale)

Définitions :

- Chaque tuile possède une couleur $\{B, G, O, V\}$;
 - Le **mot** d'une partie est la suite des couleurs des tuiles posées ;
 - Le **score** d'une partie est égal au nombre de paires ordonnées de lettres distinctes dans ce mot ;
 - ▶ BB : 1 paire BB, score = 1
- Ex :
- ▶ BVB : 2 paires BV et VB, score = 2
 - ▶ BVVVVVVVB : 3 paires BV, VB et VV, score = 3

Exigences :

- L'exécutable du projet doit pouvoir être exécuté en **color mode**.
- En color mode, à chaque fois qu'une tuile est posée, l'exécutable affiche la chaîne "COLOR:<col>"
- En color mode, à la fin de la partie, l'exécutable affiche la chaîne "COLOR:Score=<score>"

Difficultés de conception d'une extension

- Difficulté de programmation raisonnable
 - ▶ Non localité des modifications
 - ▶ Stockage des données nécessaires
- Difficulté algorithmique raisonnable
 - ▶ Au moins polynomial, mais pas exponentiel
- Facilité à séparer les effets du reste du code

Le jour prévu pour réaliser l'extension :

- ➊ S'assurer que les codes des élèves compilent et passent leurs tests ;
- ➋ Placer une étiquette (tag) git sur le code dans cet état ;
- ➌ Placer les élèves dans les conditions d'un examen machine ...
 - ▶ Travail sur les machines pédagogiques
 - ▶ Accès restreint à Internet
 - ▶ Pas de téléphone portable
- ➍ ... mais aussi d'une séance de projet "normale" :
 - ▶ Travail en équipe sur le projet
 - ▶ Présence des encadrants
 - ▶ Toute forme de discussion autorisée
- ➎ < Travailler >
- ➏ Les élèves placent un tag git sur leur code lorsqu'ils estiment avoir terminé.

Analyse des résultats

- Le travail est celui réalisé entre les tags `$START_TAG..$END_TAG`
- Utilisation de `diff2html` pour visualiser les diffs :
 - ▶ <https://diff2html.xyz>
 - ▶ <https://github.com/rtpessoa/diff2html>

```
diff2html --su closed -o stdout -f html -s side \  
  --title $(basename $(pwd)) -- \  
  -w $START_TAG..$END_TAG > diff.html
```

- Réalisation d'un script Python (via LLM) pour la vérification de la spécification.

Exemples de résultats : le bon et le mauvais diff

(cf. fichiers HTML)

Exemples de résultats : le bon et le mauvais diff

Un “bon” diff :

```
— id: ein5-proj1-27147
— members: veronique, davina
— repository: https://thor.enseirb-matmeca.fr/git/ein5-proj1-27147
— compilation: OK
— executable_exists: OK
— execution: OK
— specification: OK
  — the games always altern in colors, and are not very exhaustive
— code_inspection:
  — Presence of 'color.[ch]' OK
  — Computation of the score OK
  — Handling of getopt
  — Using malloc to handle NUM_TURNS and valgrind OK
```

Exemples de résultats : le bon et le mauvais diff

Un “moins bon” diff :

```
— id: ein5-proj1-27151
— members: dupond, dupont
— repository: https://thor.enseirb-matmeca.fr/git/ein5-proj1-27151
— compilation: OK
— executable_exists: OK
— execution: OK
— specification: KO
  — Invalid COLOR line format: 'COLOR -B'
  — Invalid COLOR line format: 'COLOR : score='
  — No computation of the score
— code_inspection:
  — Use of char to handle colors
  — Use of strcmp to compare pairs of colors BAD
  — Handling of getopt
  — Multiple duplications in addition to the code of players BAD
```

Résultats obtenus

- 18 équipes (39%) réussissent complètement l'exercice, jusqu'à valider la spec avec OK ou OK BUT,
- 5 équipes (11%) qui ont un calcul de score correct, mais qui ne respectent pas la spécification,
- 13 équipes (28%) qui ont un code qui s'exécute mais qui donne un score incorrect ou manquant,
- 3 équipes (7%) qui ont un code qui s'exécute mais qui donnent un score incorrect et font des erreurs de segmentation,
- 7 équipes (15%) qui ne permettent pas de vérifier quoi que ce soit parce que je ne peux pas compiler leur code ou l'exécuter une seule fois

Conclusion / Discussion

- Applicable à de nombreux types de projets
- S'insère (plus ou moins) naturellement à la fin des projets
- Évaluation sur une quantité **réduite** de code
- Possibilité de **discussion** pendant la séance (encadrants + élèves)
- Pas vraiment un examen, et pas vraiment un TD
- Demande un ensemble de **préparations** à l'avance :
 - ▶ Prévoir une extension du sujet de projet
 - ▶ Préparer les élèves à la séance finale