

Introduction à l'explicabilité dans les feedback automatisés fournis aux apprenants

RJC EIAH 2022

Esther Félix - 9 mai 2022

Directeurs de thèse :

Julien Broisin (IRIT - TALENT)

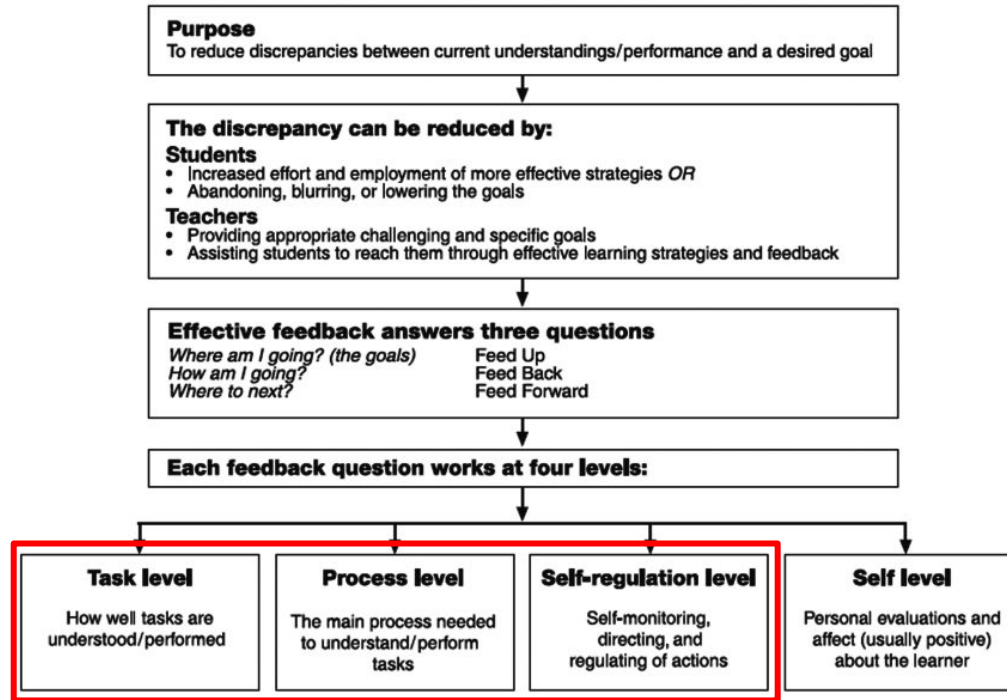
Franck Amadiou (CLLE)

Table des matières

1. Feedback dans le domaine de l'éducation
2. Feedback et explicabilité
3. Conception d'un feedback explicable
4. Expérimentation pour l'apprentissage de la programmation
5. Conclusion et perspectives

1. **Feedback dans le domaine de l'éducation**
2. Feedback et explicabilité
3. Conception d'un feedback explicable
4. Expérimentation pour l'apprentissage de la programmation
5. Conclusion et perspectives

Modélisation du feedback



Feedback dans l'apprentissage de l'informatique

Dans les EIAH consacrés à l'apprentissage de l'informatique, le feedback est :

- la plupart du temps au niveau de la **tâche**
- parfois au niveau des **processus**
- **très rarement au niveau de l'auto-régulation**

(Keuning et al.)

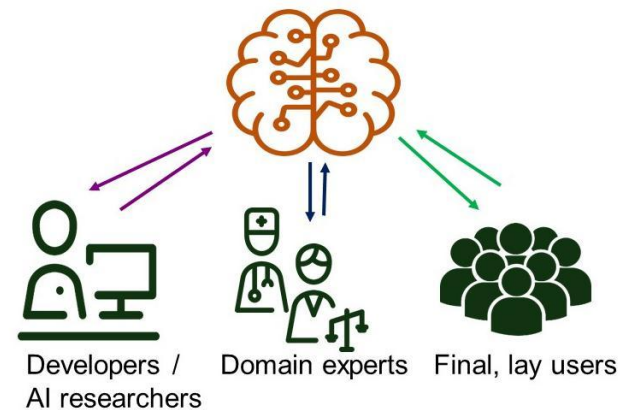
1. Feedback dans le domaine de l'éducation
2. **Feedback et explicabilité**
3. Conception d'un feedback explicable
4. Expérimentation pour l'apprentissage de la programmation
5. Conclusion et perspectives

Explicabilité : définition

- IA et Machine Learning : besoin d'expliquer les résultats et le fonctionnement de certains algorithmes
 - Pas de définition précise d'une explication :
 - peut prendre différentes formes (**comment** ?)
 - dépend des utilisateurs ciblés, du contexte, de l'objectif de l'explication (**qui, quoi, pourquoi** ?)
- contexte important à prendre en compte

Explicabilité dans le domaine de l'IA


- Recherche en xAI tournée vers les experts
- Explicabilité d'algorithmes de machine learning
- Comment prendre en compte les différents rôles et niveaux d'expertise des utilisateurs ?



WHY		Verification Improvement	Learning Adoption	Compliance with legislation
WHAT		Global model Data representation Why and Why not	Local explanations Why and why not	Why not
HOW		Intrinsic / Post-hoc	Post-hoc Visualization Natural language	Post-hoc Brief Plain language
EVAL		Completeness tests Performance	Test of comprehension Performance Survey of trust	Satisfaction questionnaires

Explicabilité dans les ITS (Intelligent Tutoring Systems)

- XAI appliqué aux ITS
- Conati et al. (2021) : explications “pourquoi” et “comment”
 - effet positif de l’explication sur la perception des conseils par les élèves (plus grande confiance dans le système et intention d’usage)
 - dépend des caractéristiques de l’élève (notamment compétences en lecture)

Why am I delivered this hint?	Why am I predicted to be low learning?	Why are the rules used for classification?
<p>My goal is to help you use the ACSP applet to your full potential.</p> <p>I have been tracking your actions {7} and noticed various patterns {10} which caused me to predict that you are not learning from the ACSP applet from the ACSP applet as effectively as you could. I call this temporary behavior lower learning {12}. One of your actions, Using Reset 4 times, made me present this hint to you.</p>  <p>This is my high-level explanation for delivering your hint. Feel free to explore as much as you want to gain a deeper understanding of this mechanism</p>		

Page “WhyHint” dans un ITS (Conati et al., 2021)

1. Feedback dans le domaine de l'éducation
2. Feedback et explicabilité
3. **Conception d'un feedback explicable**
4. Expérimentation pour l'apprentissage de la programmation
5. Conclusion et perspectives

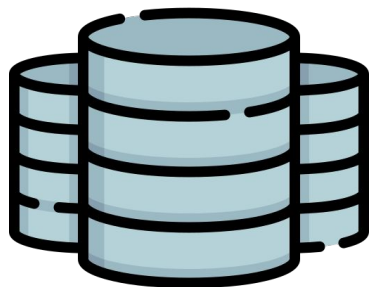
Objectifs

- Prendre en compte les avancées récentes dans le domaine de l'explicabilité + proposer du feedback au niveau de l'auto-régulation
- Apporter de nouvelles connaissances sur du **feedback explicable de haut niveau** pour l'apprentissage de la programmation

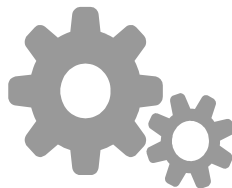
Questions de recherche

- Comment concevoir un feedback explicable, portant sur l'autorégulation et généré automatiquement, pour des programmeurs novices ?
- Quel est l'impact de ce type de feedback sur le comportement, les performances et la perception des élèves ?

Algorithme de clustering



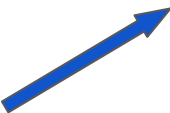
Données récoltées
sur la plateforme



Algorithme de
classification
(non supervisé)



Cluster 1 : profil “essai-erreur”
(beaucoup de soumissions, peu
de temps entre les soumissions)



Cluster 2 : élèves passant
beaucoup de temps à faire de
gros changements dans leur code



Cluster 3 : élèves performants
(peu de soumissions et peu de
temps entre les soumissions)

Fournir le feedback aux étudiants

- Clusters **recalculés chaque semaine** en fonction des données de la semaine précédente
- **Feedback** fourni aux étudiants dans une fenêtre popup au début de chaque TP
- Le feedback hebdomadaire contient un conseil **personnalisé** en fonction du profil (= cluster) de l'élève

Feedback explicable

Feedback explicable
composé de **3 parties** :

- le diagnostic
- les explications
- les conseils

Feedback on your work from the past week

To read carefully,

Edited the 10/12/2021

During the last lab session, we noticed that you have changed your strategy for solving the exercises. However, this strategy can still be improved.

Diagnosis

Indeed, you have made:

- 10 modifications in your code on average, between 2 executions, in an average time of 731 secondes
- 12% of executions with syntax errors.

Explanations

To reduce the amount of time spent making a large number of modifications in your code, you should try to **design the script entirely before executing it**, using the algorithmic notions seen in other modules.

Hints

To help you with the syntax, do not hesitate to read carefully the examples given at the beginning of the instruction sheet.

This feedback helps me understand how I achieved last week exercises (0 = I totally disagree; 10 = I totally agree)

0 1 2 3 4 5 6 7 8 9 10

This feedback will be useful to me to carry out future exercises (0 = I totally disagree; 10 = I totally agree)

0 1 2 3 4 5 6 7 8 9 10

Likert items

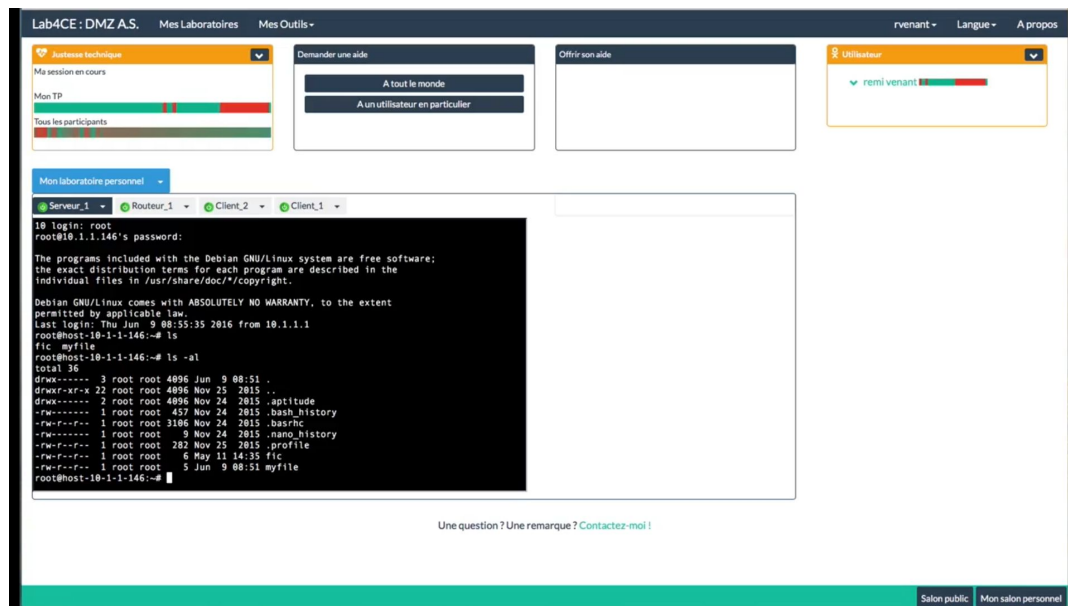
OK

Exemple de feedback avec des statistiques

1. Feedback dans le domaine de l'éducation
2. Feedback et explicabilité
3. Conception d'un feedback explicable
4. **Expérimentation pour l'apprentissage de la programmation**
5. Conclusion et perspectives

Contexte de l'expérimentation

- Participants : 155 étudiants en IUT
- TP d'informatique (programmation Shell)
- Utilisation de la plateforme Lab4CE
- 5 semaines



Interface de la plateforme Lab4CE

Méthodologie

- Étudiants des **3 clusters** divisés équitablement en **3 groupes expérimentaux** :
 - Groupe sans feedback
 - Groupe avec uniquement un diagnostic et des conseils
 - Groupe avec un diagnostic + des conseils + des statistiques (explications)
- Feedback donné à l'étudiant **différent** en fonction de son cluster

Répartition des groupes expérimentaux

Cluster 1	Conseil personnalisé selon le profil 1 + statistiques explicatives	Conseil personnalisé selon le profil 1	Message générique à propos du thème du TP
Cluster 2	Conseil personnalisé selon le profil 2 + statistiques explicatives	Conseil personnalisé selon le profil 2	
Cluster 3 (associé aux meilleures performances)	Conseil personnalisé selon le profil 3 + statistiques explicatives	Conseil personnalisé selon le profil 3	
	Groupe 1 (feedback = conseil + explication)	Groupe 2 (feedback = conseil uniquement)	Groupe 3 (pas de feedback)

Données collectées

- Évolution des **clusters** au fil des semaines
- Pour chaque feedback reçu, note du feedback par l'étudiant sur une **échelle de Likert**
- **Questionnaire** à la fin des 4 semaines
- **Focus Group** avec 10 étudiants 2 mois plus tard

Résultats

- **Pas d'impact significatif** du feedback avec explications par rapport au feedback sans explication, sur le comportement comme sur la perception du feedback
 - explications possibles : charge cognitive + élevée avec les explications, problèmes d'interprétation/de compréhension
- Focus group et questions ouvertes du questionnaire : pistes d'amélioration

Résultats du focus group

- Besoins des élèves concernant les statistiques fournies :
 - plus de précision
 - besoin d'un référentiel (fourni par le professeur ou par les autres élèves)
 - ne pas inclure de statistiques sur le temps

Pistes d'amélioration

- Formuler l'information avec des éléments similaires en termes de sémantique
- Fournir dans les explications des données de base concernant les caractéristiques de l'algorithme utilisé
- Ne pas utiliser dans les conseils des mots ou de phrases de sens opposé qui augmentent la difficulté de compréhension
- Afficher le diagnostic et les conseils au premier coup d'œil, tandis que les explications doivent être cachées au premier abord mais rapidement accessibles
- Ne pas présenter certains éléments, ou les formuler de manière positive

1. Feedback dans le domaine de l'éducation
2. Feedback et explicabilité
3. Conception d'un feedback explicable
4. Expérimentation pour l'apprentissage de la programmation
5. **Conclusion et perspectives**

Conclusion et perspectives

- Besoin de trouver un équilibre entre le nombre d'informations utiles à donner et la compréhension des élèves
- Prise en compte des caractéristiques individuelles des élèves dans le feedback donné (par exemple connaissances et compétences préalables)
- Explorer les pistes fournies par le focus group (besoins des élèves concernant les statistiques fournies)
- Tableau de bord pour les professeurs et pour les étudiants

Merci pour votre attention

Contact :
esther.felix@irit.fr

Bibliography

- Arya, Vijay, Rachel K. E. Bellamy, Pin-Yu Chen, Amit Dhurandhar, Michael Hind, Samuel C. Hoffman, Stephanie Houde, et al. «One Explanation Does Not Fit All: A Toolkit and Taxonomy of AI Explainability Techniques ». *arXiv:1909.03012 [cs, stat]*, 14 septembre 2019. <http://arxiv.org/abs/1909.03012>.
- Conati, C., K. Porayska-Pomsta, et M. Mavrikis. « AI in Education needs interpretable machine learning: Lessons from Open Learner Modelling ». *ArXiv*, 2018. <https://api.semanticscholar.org/CorpusID:49555964>
- Conati, Cristina, Oswald Barral, Vanessa Putnam, et Lea Rieger. « Toward Personalized XAI: A Case Study in Intelligent Tutoring Systems ». *Artificial Intelligence* 298 (1 septembre 2021): 103503. <https://doi.org/10.1016/j.artint.2021.103503>.
- Keuning, Jeuring and Heeren. 2018. A Systematic Literature Review of Automated Feedback Generation for Programming Exercises. *ACM Trans. Comput. Educ.* 19, 1, Article 3 (March 2019), 43 pages. <https://doi.org/10.1145/3231711>
- Hattie, John, et Helen Timperley. « The Power of Feedback ». *Review of Educational Research* 77, n° 1 (mars 2007): 81-112. <https://doi.org/10.3102/003465430298487>.
- Miller, Tim. « Explanation in Artificial Intelligence: Insights from the Social Sciences ». *Artificial Intelligence* 267 (1 février 2019): 1-38. <https://doi.org/10.1016/j.artint.2018.07.007>.
- Ribera, Mireia and Àgata Lapedriza. “Can we do better explanations? A proposal of user-centered explainable AI.” *IUI Workshops* (2019). <https://api.semanticscholar.org/CorpusID:84832474>