

# Un modèle générique avec structuration des compétences et facteurs externes pour le Bayesian Knowledge Tracing

Quentin Couland<sup>1</sup>, Philippe Leray<sup>1</sup>, Amine Boulahmel<sup>2</sup>

<sup>1</sup>Nantes Université, École Centrale Nantes, CNRS, LS2N UMR 6004, F-44000 Nantes, France

<sup>2</sup>IMT Atlantique, CNRS, Lab-STICC UMR 6285, F-29238 Brest, France

quentin.couland@ls2n.fr, philippe.leray@ls2n.fr, amine.boulahmel@imt-atlantique.fr

## Abstract

Un des objectifs du projet ANR xCALE est d'être capable d'estimer l'évolution des compétences d'un apprenant au fur et à mesure de son apprentissage, pour ensuite l'aider dans son auto-régulation de cet apprentissage. Dans cet article, nous nous concentrons sur ce suivi des compétences au cours du temps, souvent appelé *Knowledge Tracing*, dans un contexte où le référentiel de compétences peut être structuré, et où plusieurs facteurs externes peuvent aussi influencer sur l'apprentissage. Pour cela, nous proposons un modèle générique étendant les modèles de *Bayesian Knowledge Tracing*, dont nous illustrons l'intérêt sur un jeu de données simulé inspiré de l'enseignement de Mathématiques au collège, puis sur des données réelles issues d'une plate-forme d'apprentissage de la programmation.

Dans le cadre du projet ANR xCALE, nous nous intéressons à l'auto-régulation de l'apprentissage chez les apprenants dans un contexte de MOOC, en confrontant les stratégies d'auto-régulation comme les demandes d'aide, etc. faites par l'apprenant, avec son "contexte" et principalement son niveau de compétences (ou son évolution) (Djelil et al. 2021; Roche et al. 2022).

Pour cela, nous avons tout d'abord besoin de modéliser l'évolution des compétences d'un apprenant en lien avec ses résultats, tâche classiquement définie sous le terme *Knowledge Tracing*.

Dans notre contexte, plusieurs compétences sont mises en œuvre, avec des référentiels de compétences qui peuvent être structurés. Nous avons aussi besoin d'une "souplesse" dans la modélisation permettant de prendre en compte différentes variables de contexte, facteurs externes, influant sur la vitesse d'apprentissage.

Pour cela, nous proposons ici une approche générique de Bayesian Knowledge Tracing faisant évoluer la manière de prendre en compte la structuration des compétences proposée par (Käser et al. 2014a), pour obtenir un modèle (1) plus facilement apprenable ou définissable par expertise, et (2) pouvant incorporer de manière explicite d'autres facteurs dans l'évolution de la vitesse d'apprentissage.

Notre approche est illustrée tout d'abord sur un jeu de données simulé inspiré de l'enseignement de Mathématiques au collège, puis sur des données réelles issues d'une plate-forme d'apprentissage de la programmation dans le cadre du projet xCALE.

## Bayesian Knowledge Tracing

### Définition

Les modèles de *Bayesian Knowledge Tracing* (*BKT*) sont des modèles graphiques probabilistes dynamiques dérivés des modèles de Markov cachés, initialement définis par (Corbett and Anderson 1995) pour suivre au cours du temps le niveau de compétence d'un apprenant en fonction de ses résultats observés. Leur efficacité par rapport à d'autres modèles a été illustrée par exemple par (Baker et al. 2011).

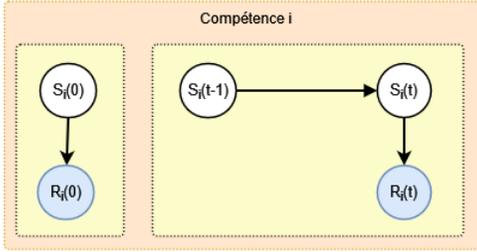


FIGURE 1 – Modèle de Bayesian Knowledge Tracing

Dans leur version la plus simple, représentée par la figure 1, le modèle décrit un processus aléatoire avec deux variables booléennes évoluant au cours du temps.  $S_i(t)$  décrit le niveau de compétence d'un apprenant à chaque instant. Dans la version booléenne la plus simple, la compétence est soit "non acquise" ( $S_i(t) = 0$ ), soit acquise ( $S_i(t) = 1$ ). Il est possible d'envisager des échelles plus riches, mais nous nous concentrerons par simplification au cas booléen dans le reste de cet article.

Cette variable n'est jamais directement observée, mais son évolution au cours du temps est modélisée par une chaîne de Markov, avec une distribution initiale  $P(S_i(0))$  et la matrice de transition  $P(S_i(t) | S_i(t-1) = T_i)$ . Ces distributions sont souvent représentées par les 3 paramètres suivants :

$p_0 = P(S_i(0) = 1)$  est la probabilité initiale qu'un apprenant maîtrise la compétence  $i$  à  $t = 0$ ,

$p_L = P(S_i(t) = 1 | S_i(t-1) = 0)$  est la probabilité que l'apprenant apprenne la compétence entre  $t-1$  et  $t$ ,

$p_F = P(S_i(t) = 0 | S_i(t-1) = 1)$  est la probabilité que l'apprenant oublie la compétence entre  $t-1$  et  $t$  (pour du suivi d'un apprenant sur une période courte, on suppose généralement  $p_F = 0$ ).

La variable  $R_i(t)$  est une autre variable décrivant la réussite éventuelle d'un exercice, observée à l'instant  $t$ . Comme précédemment, nous considérerons ici une échelle d'évaluation booléenne, où l'exercice n'est pas réussi ( $R_i(t) = 0$ ), ou réussi ( $R_i(t) = 1$ ).

Comme dans tout modèle de Markov caché, cette variable est dépendante de la va-

riable cachée, i.e. le niveau de maîtrise de compétence de l'apprenant. Cette dépendance est modélisée par la distribution conditionnelle  $P(R_i(t) | S_i(t))$ , et définie par les deux paramètres suivants :

$p_G = P(R_i(t) = 1 | S_i(t) = 0)$  est la probabilité que l'apprenant donne une bonne réponse alors qu'il ne maîtrisait pas la compétence,

$p_S = P(R_i(t) = 0 | S_i(t) = 1)$  est la probabilité que l'apprenant donne une mauvaise réponse alors qu'il maîtrisait la compétence (il a fait une faute d'inattention par exemple)

Un modèle simple de BKT est donc défini par les 5 paramètres ( $p_0, p_L, p_F, p_G, p_S$ ) qui peuvent être définis par expertise ou appris à partir de données grâce à l'algorithme EM (Dempster, Laird, and Rubin 1977).

De nombreux travaux ont aussi proposé une personnalisation de ces paramètres en fonction de l'apprenant (Baker, Corbett, and Aleven 2008; Pardos and Heffernan 2010; Yudelson, Koedinger, and Gordon 2013; Wang and Heffernan 2012). Le *Knowledge Tracing* s'est aussi emparé des avancées en termes d'apprentissage profond avec les travaux de (Piech et al. 2015) qui ont ensuite donné lieu à de nombreux développements résumés dans l'état de l'art proposé par (Abdelrahman, Wang, and Nunes 2023).

## Utilisation

Les modèles BKT sont généralement utilisés de manière *online* pour estimer à chaque instant  $P(S_i(t) | R_i(0:t))$ , le niveau de matrice de la compétence, à partir des résultats obtenus jusque ici.

La figure 2 montre un exemple de visualisation tiré de (Piech et al. 2015) où l'on suit l'évolution de 5 compétences mathématiques d'un collégien. L'apprenant répond correctement à deux problèmes portant sur les racines carrées, puis échoue à un problème sur l'interception en X. Dans les 47 exercices suivants, il travaille sur des exercices sur les interception en X, en Y, et finalement sur des exercices sur la représentation graphique d'équations linéaires.

Ce type de modèle a été utilisé dans de nombreux systèmes de tutorat intelligents pour modéliser et prévoir les performances (Conati, Gertner, and Vanlehn 2002; Mayo and Mitro-

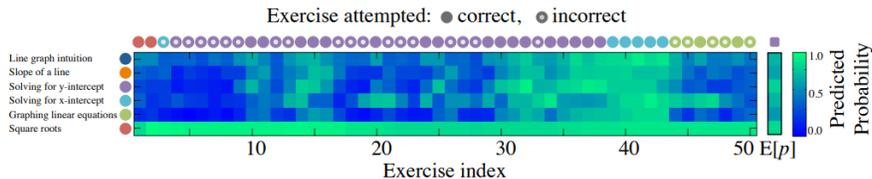


FIGURE 2 – Evolution de 5 compétences mathématiques au fur et à mesure de la réalisation de 50 exercices (figure tirée de (Piech et al. 2015)).

vic 2001), l'état de l'engagement (Baschera et al. 2011) et les objectifs des apprenants (Conati, Gertner, and Vanlehn 2002).

### Améliorations

Dans le cadre du projet xCALE, nous devons prendre en compte différents facteurs dans la vitesse d'acquisition d'une compétence. Quelques travaux de la littérature traitent ponctuellement de questions similaires.

**Prise en compte de référentiels de compétences** Käser et al. (2014a) proposent d'enrichir les modèles de BKT classiques afin de prendre en compte la structuration éventuelle entre compétences.

Pour décrire le fait qu'une compétence  $i$  soit liée dans le référentiel à d'autres compétences  $j...k$ , ils représentent la dépendance probabiliste  $P(S_i(t) | S_i(t-1), S_j(t)...S_k(t))$  par un modèle log-linéaire avec  $2^{n-1}$  paramètres lorsque il y a  $n$  compétences "parentes".

L'apprentissage des paramètres de ce modèle ne peut plus se faire par EM, et est reformulé comme un problème d'optimisation convexe par (Käser et al. 2014b) nécessitant l'utilisation d'un algorithme d'optimisation de type descente de gradient.

(Chen, Wuillemin, and Labat 2015) proposent une méthode d'apprentissage d'une structure de compétence de type pré-requis basée sur l'extraction de règles d'association probabilistes.

**Prise en compte d'autres facteurs** Certains auteurs se sont également intéressés à l'intégration de facteurs externes afin d'améliorer la précision du système. Par exemple (Pardos and Heffernan 2011) utilisent un noeud (et

un arc) supplémentaire pour représenter la difficulté de chaque question et l'impact de la difficulté sur le résultat de l'évaluation.

### Contributions

Nous proposons dans cet article une amélioration générique des modèles BKT permettant de répondre à différentes situations rencontrées dans le projet xCALE.

Pour cela, nous allons commencer à décrire une amélioration très simple correspondant à une situation courante, mais pas décrite dans la littérature, i.e. le fait que toutes les compétences ne soient pas monopolisées à chaque exercice. Nous décrirons ensuite notre architecture générique permettant de prendre en compte de manière unifiée différents facteurs influant sur la vitesse d'acquisition d'une compétence.

#### Prise en compte de la mise en oeuvre d'une compétence

Dans un modèle de BKT classique, on suppose qu'une compétence est monopolisée à chaque instant (exercice). Ce n'est pas le cas dans notre contexte, où un exercice fait à l'instant  $t$  ne mettra en oeuvre qu'une partie des compétences définies.

Pour cela, nous proposons une première évolution du modèle BKT, décrite dans la figure 3.  $S_i(t)$  décrit toujours le niveau de la compétence  $i$  à l'instant  $t$ , et  $R_i(t)$  l'évaluation de cette compétence réalisée pendant la réalisation de l'exercice. Nous rajoutons la variable booléenne  $Tr_i(t)$  indiquant si la compétence a été mise en oeuvre par l'exercice. Si  $Tr_i(t) = 1$ , la compétence est mise en oeuvre et évaluée par

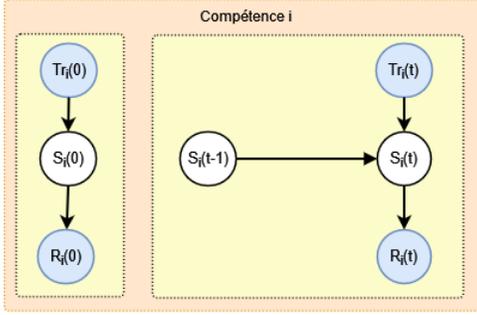


FIGURE 3 – BKT basique avec ajout du trigger

$R_i(t)$ . Si  $Tr_i(t) = 0$ , la compétence n'est pas mise en œuvre et  $R_i(t)$  n'est pas mesurée.

L'ajout de cette variable nous permet de conditionner l'évolution de la compétence au fait qu'elle ait été réellement mise en œuvre lors de l'exercice. Nous définissons ainsi :

$$P(S_i(t) | S_i(t-1), Tr_i(t) = 1) = \mathbb{T}_i$$

pour décrire le fait que le niveau de compétence évolue "classiquement" avec la matrice de transition  $\mathbb{T}_i$  classique d'un BKT, et

$$P(S_i(t) | S_i(t-1), Tr_i(t) = 0) = \mathbb{I}$$

pour décrire le fait que le niveau de compétence n'évolue pas si la compétence n'a pas été mise en œuvre.

### Prise en compte d'un référentiel de compétences structuré et d'autres facteurs

La figure 4 décrit une amélioration du modèle précédent pour prendre en compte deux fonctionnalités différentes pouvant influencer sur la vitesse d'acquisition d'une compétence : (1) la prise en compte d'une structuration entre compétences, et (2) l'existence de facteurs externes pouvant être liés par exemple à l'exercice réalisé (comme son niveau de difficulté).

Pour cela, nous rajoutons une variable latente  $V_i(t)$  décrivant explicitement la vitesse d'apprentissage de la compétence avec plusieurs niveaux de '0' (vitesse nulle) à '4' (vitesse rapide). Cette variable fait la fusion entre les variables  $Tr_i(t)$  et  $F_i(t)$ . Si  $Tr_i(t) = 0$ , la compétence n'est pas mise en œuvre dans

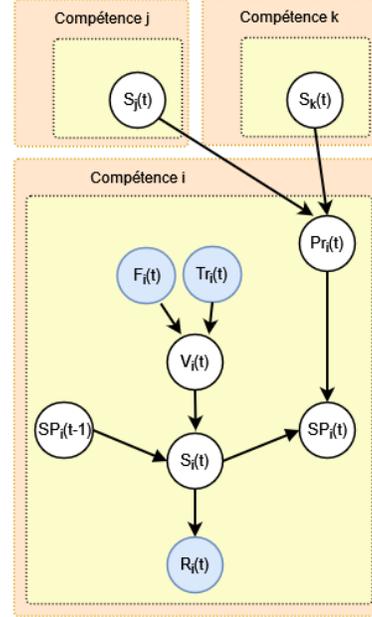


FIGURE 4 – BKT (modèle transition) avec prise en compte d'un référentiel de compétences structuré et de facteurs externes

l'exercice, donc la vitesse d'apprentissage est nulle.

$$P(V_i(t) | F_i(t), Tr_i(t) = 0) = \delta_{V_i(t)=0}$$

Si  $Tr_i(t) = 1$ , la distribution  $P(V_i(t) | F_i(t), Tr_i(t) = 1)$  traduit la dépendance entre la vitesse d'acquisition de la compétence et le facteur externe  $F_i(t)$ . Si  $F_i$  est par exemple la difficulté de l'exercice réalisé, on peut par exemple décrire que la vitesse augmente lorsque on réalise des exercices plus difficiles. Cette partie de la distribution est complètement dépendante du choix des facteurs externes utilisés dans le modèle.

Nous définissons ensuite  $Pr_i(t)$  le niveau de compétence "agrégé" atteint pour compétence  $S_j, \dots, S_k$  parentes de la compétence cible. Nous proposons ici de représenter  $P(Pr_i(t) | S_j(t) \dots S_k(t))$  par une distribution conditionnelle de type NoisyMax (Diez 1993) avec des poids  $w_{ji} \dots w_{ki}$  permettant de définir individuellement les influences de chaque compétence parente sur la compétence  $i$ .

Nous allons ensuite représenter le fait que  $S_i(t)$  le niveau d'une compétence à l'instant  $t$  dépend comme précédemment du niveau de compétence à l'instant précédent  $S_i(t-1)$ , mais aussi de la maîtrise des compétences associées à l'instant précédent  $Pr_i(t-1)$ . Pour cela, nous proposons de définir une nouvelle variable virtuelle  $SP_i(t)$  agrégeant  $S_i(t)$  et  $Pr_i(t)$  (par une fonction déterministe max), puis de définir que  $S_i(t)$  dépend maintenant de  $SP_i(t-1)$  au lieu de  $S_i(t-1)$  pour modéliser le fait que la maîtrise des compétences associées va accélérer l'évolution de la compétence cible.

$$P(S_i(t) | SP_i(t-1), V_i(t) = v) = \mathbb{T}_{iv}$$

où  $\mathbb{T}_{iv}$  seront différentes matrices de transition (une pour chaque valeur  $v$ ), décrivant ainsi plusieurs dynamiques d'apprentissage.

Ainsi lorsque  $v = 0$ ,  $\mathbb{T}_{iv} = \mathcal{I}$ , décrivant le fait que la compétence n'évolue pas lorsque la vitesse d'acquisition est nulle. Pour les vitesses d'acquisition  $v$  non nulle, ces matrices de transition sont définies comme dans le cas classique, par un ensemble de paramètres  $p_{Lv}$  dépendant maintenant de cette vitesse.

En résumé, les paramètres de notre modèle de BKT évoluent donc de la façon suivante :  $p_0$ ,  $p_G$  et  $p_S$  ne changent pas.

$p_L$  qui décrivait la vitesse d'apprentissage de la compétence est remplacé par un vecteur  $p_{Lv}$ , probabilité que l'apprenant acquiert la compétence pour chaque vitesse d'acquisition  $v$ .

Le vecteur  $w_{j_i} \dots w_{k_i}$  décrit les poids de chaque compétence "parente" définis pour la compétence  $i$ .

La distribution conditionnelle  $P(V_i(t) | F_i(t), Tr_i(t) = 1)$  traduit la dépendance entre la vitesse d'acquisition de la compétence et les facteurs externes utilisés dans le modèle.

## Discussion

A la différence de l'approche log-linéaire de Käser et al. dont la représentation de  $n$  compétences parentes nécessitait  $2^{n-1}$  paramètres notre proposition ne nécessite que  $n$  paramètres. De plus, l'algorithme EM classique peut être utilisé pour apprendre les paramètres de ce réseau bayésien dynamique.

Notre proposition permet aussi de prendre en compte de manière générique l'influence de

facteurs externes sur la vitesse d'acquisition des compétences, sans faire exploser la complexité du modèle.

## Cas d'utilisation "Impacted"

Nous allons commencer par illustrer l'intérêt de notre proposition sur un cas d'utilisation fictif, construit initialement dans le cadre d'un projet collaboratif avec la Technopole de la Réunion concernant la pré-incubation d'une startup dans le domaine de l'EdTech.

Pour cela, nous nous sommes intéressés sur l'enseignement des Mathématiques au collège, en nous appuyant sur les ressources fournies par (Eduscol 2016).

## Contexte expérimental

**Compétences** Nous avons choisi de nous concentrer sur trois compétences principales, mobilisées soit individuellement, soit conjointement :  $C_1$  utiliser les nombres,  $C_2$  utiliser le calcul littéral,  $C_3$  démarche scientifique.

L'Education Nationale formalise sur une même échelle les différents niveaux de maîtrise d'une compétence et l'évaluation de cette maîtrise. Dans notre cas d'utilisation, nous avons fait le choix de ne garder que deux niveaux de compétences, 1 : maîtrise insuffisante ou fragile, 2 : maîtrise satisfaisante ou très bonne maîtrise.

**Exercices** Nous avons sélectionné 8 exercices, à partir des mêmes ressources fournies par l'Education Nationale, qui permettent de travailler tout ou partie des 3 compétences précédentes, comme décrit dans la table 1.

Exercice	$C_1$	$C_2$	$C_3$
( $E_1$ ) Calcul avec fractions	X		
( $E_2$ ) Évolution proportion	X		
( $E_3$ ) Proportionnalité	X		
( $E_4$ ) Vente de livres	X		X
( $E_5$ ) Conjecture	X	X	
( $E_6$ ) Programmes de calcul	X	X	
( $E_7$ ) Devinette Math.		X	X
( $E_8$ ) Récupérateur d'eau			X

TABLE 1 – Nos 8 exercices de Mathématiques et les compétences associées

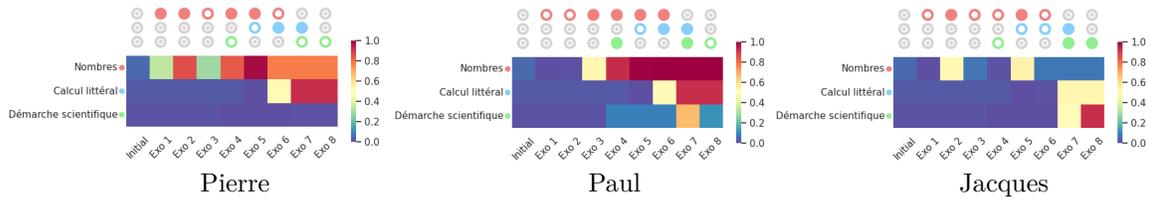


FIGURE 5 – Évolution du niveau de maîtrise des 3 compétences au cours des exercices, pour nos trois apprenants. Un rond gris indique que la compétence n'a pas été mobilisée par l'exercice. Un cercle vide de couleur (resp. un rond plein de couleur) indique que la compétence associée a été mobilisée et évaluée négativement (resp. positivement).

**Apprenants** Dans ce cas d'utilisation, nous définissons 3 apprenants fictifs, Pierre, Paul et Jacques, possédant chacun un profil particulier. Pierre a des difficultés avec la démarche scientifique ( $C_3$ ). Paul n'a aucune difficulté particulière. Jacques est plutôt embêté par l'utilisation des nombres ( $C_1$ ) et le calcul littéral ( $C_2$ ). tout cela s'observe dans leurs résultats obtenus lorsqu'ils ont réalisé chacun des 8 exercices, proposés dans cet ordre, et décrits dans la partie haute de la figure 5, où un cercle vide et un rond plein signifient respectivement que l'évaluation de la compétence est échouée ou réussie.

### Estimation du niveau de compétence, sans structuration des compétences

Nous construisons ici un modèle BKT par apprenant, avec des compétences indépendantes, avec les paramètres  $p_0(C_1) = 0.05$ ,  $p_0(C_2) = 0.02$ ,  $p_0(C_3) = 0.01$ ,  $p_G = 0.1$ ,  $p_S = 0.05$ ,  $p_L = 0.1$ ,  $p_F = 0$ .

La figure 5 nous présente, pour nos 3 apprenants, la probabilité du niveau de maîtrise des 3 compétences, estimé à  $t=0$ , puis après chaque exercice.

On retrouve bien le fait que Pierre semble bien maîtriser la compétence  $C_1$  (nombres), même s'il échoue sur les exercices 3 et 6, son niveau de maîtrise atteint les 80%. Idem pour le calcul littéral, où un échec suivi de deux réussites montrent pour l'instant qu'il est sur la bonne voie. Par contre, ses trois échecs sur les exercices liés à la démarche scientifique montrent qu'il ne maîtrise effectivement pas cette compétence.

De son côté, Paul démarre lentement sur la compétence Nombre, mais ces 4 derniers exercices validés l'amène à une très bonne maîtrise de la compétence. Idem pour le calcul littéral. La démarche scientifique est une compétence moins maîtrisée dès le début, et ses deux réussites puis un échec ne permettent pas encore de montrer une bonne maîtrise.

Jacques a du mal avec la compétence Nombres, il alterne réussites et surtout échecs, et ne montre pas de bonne maîtrise. Par contre pour le calcul littéral, deux échecs suivis d'une réussite montrent pour l'instant qu'il est sur la bonne voie. Idem pour sa progression en termes de démarche scientifique.

**Prise en compte d'un référentiel de compétences structuré** La figure 6 nous montre l'évolution du niveau de maîtrise de nos 3 compétences, pour une suite de 16 exercices, avec trois référentiels de compétences. La partie (1) présente le scénario "baseline", sans dépendance entre compétences. Les parties (2) et (3) utilisent le fait que  $C_1$  et  $C_2$  influent sur  $C_3$ , avec des poids de plus en plus forts entre (2) et (3).

La série d'exercices commence par mettre en œuvre  $C_1$ , puis  $C_2$ , avec plus ou moins de succès. Sans dépendance entre compétences, nous voyons que pendant ce temps là (les 11 premières itérations) le niveau de maîtrise de  $C_3$  reste à son niveau initial dans le scénario 1, alors qu'il augmente légèrement dans les 2 autres scénarios, et d'autant plus que les poids sont plus élevés (scénario 3). Lorsque la compétence  $C_3$  est directement mise en œuvre dans les derniers exercices, on voit alors que le ni-

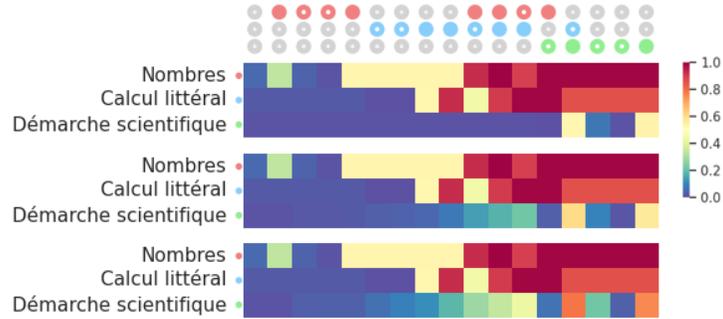


FIGURE 6 – Évolution du niveau de maîtrise des 3 compétences au cours des exercices, pour trois référentiels de compétences : (1) compétences indépendantes (2) avec influence des deux premières compétences vers la dernière (démarche scientifique) et des poids 0.02 et 0.05, et (3) idem avec des poids plus élevés (0.05 et 0.1).

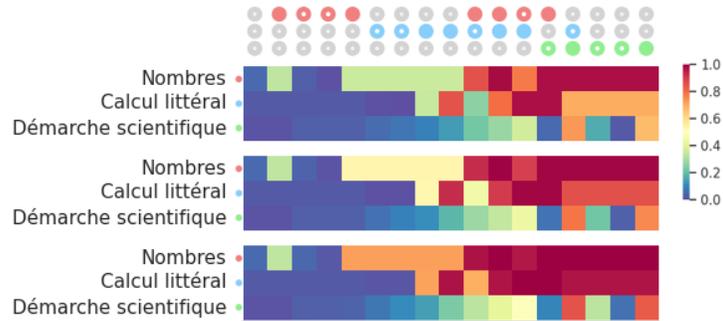


FIGURE 7 – Évolution du niveau de maîtrise des 3 compétences au cours des exercices, pour trois scénarios liés aux facteurs externes : (1) avec des exercices faciles (2) de difficulté moyenne, et (3) difficiles.

veau de maîtrise associé continue à être plus élevé dans les deux scénarios modélisant ces influences.

**Prise en compte d'autres facteurs** La figure 7 nous propose l'évolution du niveau de maîtrise de nos 3 compétences, avec un référentiel de compétences structuré, pour la même suite de 16 exercices, mais avec maintenant trois scénarios liés à la prise en compte de facteurs externes, ici la difficulté des exercices. Les exercices sont tous supposés faciles dans le premier scénario, puis de difficulté moyenne et difficiles dans les scénarios suivants.

Les résultats obtenus illustrent bien la pro-

priété de dépendance entre la difficulté de l'exercice et la vitesse d'acquisition de la compétence décrite maintenant dans le modèle : plus l'apprenant travaille sur des exercices compliqués, plus son niveau de maîtrise augmente rapidement.

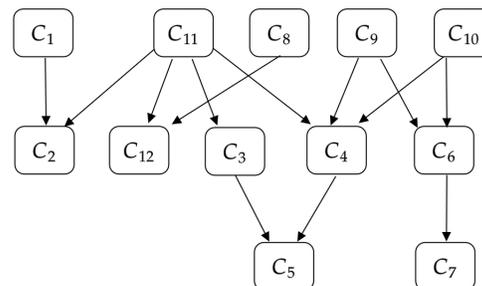
## Cas d'utilisation "xCALE"

### Contexte expérimental

France-IOI, membre du consortium du projet ANR xCALE, est une association proposant plusieurs plate-formes d'apprentissage de la programmation et de l'algorithmique, dont

	Être capable d'
$C_1$	appeler une fonction sans argument
$C_2$	appeler une fonction avec argument
$C_3$	écrire une boucle répéter x fois
$C_4$	écrire une boucle tant-que
$C_5$	imbriquer deux boucles
$C_6$	utiliser un <i>if</i>
$C_7$	utiliser un <i>if-else</i>
$C_8$	utiliser les opérateurs arithmétiques
$C_9$	utiliser les opérateurs de comparaison
$C_{10}$	utiliser les opérateurs logiques
$C_{11}$	utiliser une variable
$C_{12}$	incrémenter une variable

(a) les 12 compétences mises en œuvre



(b) structuration des compétences

FIGURE 8 – Cas d'utilisation xCALE : référentiel de compétences utilisé

la plate-forme Quick-Pi<sup>1</sup>.

12 compétences (décrites dans la table 8a) sont potentiellement mobilisées dans les exercices proposés sur cette plate-forme. La figure 8b décrit la structuration entre compétences identifiées par les experts de la plate-forme.

Nous avons sélectionné 16 activités des deux premiers parcours de Quick-Pi. Chacune de ces activités existe sous 3 formes différentes de difficulté croissante, et pouvant mettre en œuvre des compétences différentes, mais seule la compétence la plus "complexe" sera alors évaluée. Cela nous amène donc à 48 exercices différents.

Nous avons extrait 14 apprenants ayant réalisé le parcours 1 de Quick-Pi. Les résultats recueillis sont gradués de 0 (compétence non acquise à 3 (exercice de niveau 3 réussi).

### Paramètres du modèle

Nous construisons un modèle BKT pour chaque apprenant, avec les paramètres suivants :  $p_G = 0.1$ ,  $p_S = 0.05$ ,  $p_F = 0$ ,  $p_{Lv} = [0.1, 0.3, 0.6]$  et  $p_0 = [0.1, 0.08, 0.02]$  est la probabilité initiale des trois niveaux de maîtrise d'une compétence. Les poids utilisés pour les dépendances entre compétences sont constants,  $w = 0.05$ .

### Résultats

La figure 9 nous présente de premiers résultats concernant l'estimation du niveau de maîtrise des 12 compétences au cours du temps

1. <https://quick-pi.org/>

pour l'un des apprenants au fur et à mesure de son activité. Les exercices réalisés font travailler les compétences  $C_1$  et  $C_2$  sur presque toute la durée de l'expérience.  $C_3$  est mobilisée et évaluée pendant les premières itérations, alors que  $C_4$  à  $C_6$  sont plutôt mobilisées dans les itérations suivantes.

Sauf à la première itération,  $C_1$  n'est pas évaluée directement, mais sa mobilisation sur la durée fait que son niveau de maîtrise augmente.

$C_2$  est d'abord évaluée négativement, puis très bien évaluée et continue à être mobilisée, et qui fait que son niveau de maîtrise augmente encore plus vite.

Les performances concernant  $C_3$  sont plus variables au début, mais s'améliorent ensuite.

Les autres compétences, moins directement évaluées, comme  $C_5$  ou  $C_7$  ont un niveau de maîtrise qui s'améliore au fur et à mesure de la pratique, mais qui redescend logiquement lorsque la compétence est ensuite évaluée négativement.

### Conclusion

Nous avons proposé dans cet article une approche générique de Bayesian Knowledge Tracing permettant de modéliser de manière unifiée différents facteurs influant sur la vitesse d'acquisition d'une compétence.

Notre approche permet à la fois de prendre en compte les structurations éventuelles entre compétences d'une manière plus concise que



FIGURE 9 – Cas d'utilisation xCALE : estimation du niveau de maîtrise des 12 compétences au cours du temps pour l'un des apprenants, en fonction de son activité. Pour chaque compétence, les périodes en noir montrent les instants où la compétence est mobilisée, et le résultat éventuel de l'évaluation.

l'approche de (Käser et al. 2014a) et d'incorporer de manière explicite d'autres facteurs dans l'évolution de la vitesse d'apprentissage comme la difficulté de l'exercice réalisé.

L'intérêt de cette proposition a été illustrée dans le cadre d'un cas d'utilisation fictif "Impacted", puis avec de premiers résultats obtenus sur les données réelles fournies par FranceIOI dans le cadre du projet xCALE.

Les paramètres de nos modèles sont pour l'instant estimés par expertise, mais une prochaine étape consistera à apprendre ces paramètres à partir de données réelles, et à pouvoir effectuer des comparaisons avec certains benchmarks, même s'il semble en exister peu présentant des référentiels de compétence structurés connus.

La dernière perspective concerne plus directement le projet xCALE pour lequel notre objectif est de confronter les stratégies d'auto-régulation comme les demandes d'aide, etc. faites par l'apprenant pendant son apprentis-

sage, avec ses niveaux de compétences estimés (ou son évolution) pour être capable (1) d'estimer leur cohérence et (2) de recommander d'éventuelles autres stratégies.

## Remerciements

Ce travail est financé dans le cadre du projet ANR xCALE (ANR-20-CE38-0010). Le cas d'utilisation "Impacted" a été partiellement financé dans le cadre d'un Contrat de Collaboration de Recherche (20210519-01-IRR-GC-104) avec la Technopole de la Réunion. Nous remercions aussi Mathias Hiron et FranceIOI pour l'accès aux données utilisées dans le cadre du second cas d'utilisation.

## Références

- [1] Abdelrahman, G. ; Wang, Q. ; and Nunes, B. 2023. Knowledge tracing : A survey. *ACM Comput. Surv.* 55(11).

- [2] Baker, R. S. J. d. ; Pardos, Z. A. ; Gowda, S. M. ; Nooraei, B. B. ; and Heffernan, N. T. 2011. Ensembling predictions of student knowledge within intelligent tutoring systems. In *User Modeling, Adaption and Personalization*, 13–24. Springer.
- [3] Baker, R. S. J. d. ; Corbett, A. T. ; and Alven, V. 2008. More accurate student modeling through contextual estimation of slip and guess probabilities in bayesian knowledge tracing. In *Intelligent Tutoring Systems*, 406–415. Springer.
- [4] Baschera, G.-M. ; Busetto, A. G. ; Klingler, S. ; Buhmann, J. M. ; and Gross, M. 2011. Modeling engagement dynamics in spelling learning. In *International Conference on AI in Education*, 31–38. Springer.
- [5] Chen, Y. ; Wullemmin, P.-H. ; and Labat, J.-M. 2015. Discovering prerequisite structure of skills through probabilistic association rules mining. In *Proc. of EDM 2015*, 117–124.
- [6] Conati, C. ; Gertner, A. ; and Vanlehn, K. 2002. Using bayesian networks to manage uncertainty in student modeling. *User modeling and user-adapted interaction* 12(4) :371–417.
- [7] Corbett, A. T., and Anderson, J. R. 1995. Knowledge tracing : Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction* 4(4) :253–278.
- [8] Dempster, A. P. ; Laird, N. M. ; and Rubin, D. B. 1977. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)* 1–38.
- [9] Diez, F. 1993. Parameter adjustment in Bayes networks. The generalized noisy OR-gate. In *Proc. of UAI95*, 99–105. Washington D.C. : Morgan Kaufmann, San Mateo, CA.
- [10] Djelil, F. ; Gilliot, J.-M. ; Garlatti, S. ; and Leray, P. 2021. Supporting Self-Regulation Learning Using a Bayesian Approach. Some Preliminary Insights. In *IJCAI-21, Workshop AI for Education*.
- [11] Eduscol. 2016. Ressources pour l'évaluation des mathématiques (cycle 4). Technical report, Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche.
- [12] Käser, T. ; Klingler, S. ; Schwing, A. G. ; and Gross, M. 2014a. Beyond knowledge tracing : Modeling skill topologies with bayesian networks. In *International Conference on Intelligent Tutoring Systems*, 188–198. Springer.
- [13] Käser, T. ; Schwing, A. ; Hazan, T. ; and Gross, M. 2014b. Computational education using latent structured prediction. In *Artificial Intelligence and Statistics*, 540–548.
- [14] Mayo, M., and Mitrovic, A. 2001. Optimising its behaviour with bayesian networks and decision theory. *International Journal of AI in Education* 12 :124–153.
- [15] Pardos, Z. A., and Heffernan, N. T. 2010. Modeling individualization in a bayesian networks implementation of knowledge tracing. In *User Modeling, Adaptation, and Personalization*, 255–266. Springer.
- [16] Pardos, Z. A., and Heffernan, N. T. 2011. Kt-idem : introducing item difficulty to the knowledge tracing model. In *International conference on user modeling, adaptation, and personalization*, 243–254. Springer.
- [17] Piech, C. ; Bassen, J. ; Huang, J. ; Ganguli, S. ; Sahami, M. ; Guibas, L. J. ; and Sohl-Dickstein, J. 2015. Deep knowledge tracing. In *Proc. of NIPS 2015*.
- [18] Roche, M. ; Pentecouteau, H. ; Leray, P. ; Djelil, F. ; Bertrand, E. ; Eneau, J. ; Gilliot, J.-M. ; and Lameul, G. 2022. L'autorégulation des apprentissages dans une formation pour adulte. l'exemple de la demande d'aide. In *Congrès international d'Actualité de la Recherche en Éducation et en Formation (AREF)*.
- [19] Wang, Y., and Heffernan, N. T. 2012. The student skill model. In *Intelligent Tutoring Systems*, 399–404. Springer.
- [20] Yudelson, M. V. ; Koedinger, K. R. ; and Gordon, G. J. 2013. Individualized bayesian knowledge tracing models. In *International Conference on AI in Education*, 171–180. Springer.