LABORATOIRE DE RECHERCHE EN INFORMATIQUE

CentraleSupélec

Colloque WIMS 2018

Les tests adaptatifs

Yolaine Bourda
(Yolaine.Bourda@centralesupelec.fr)

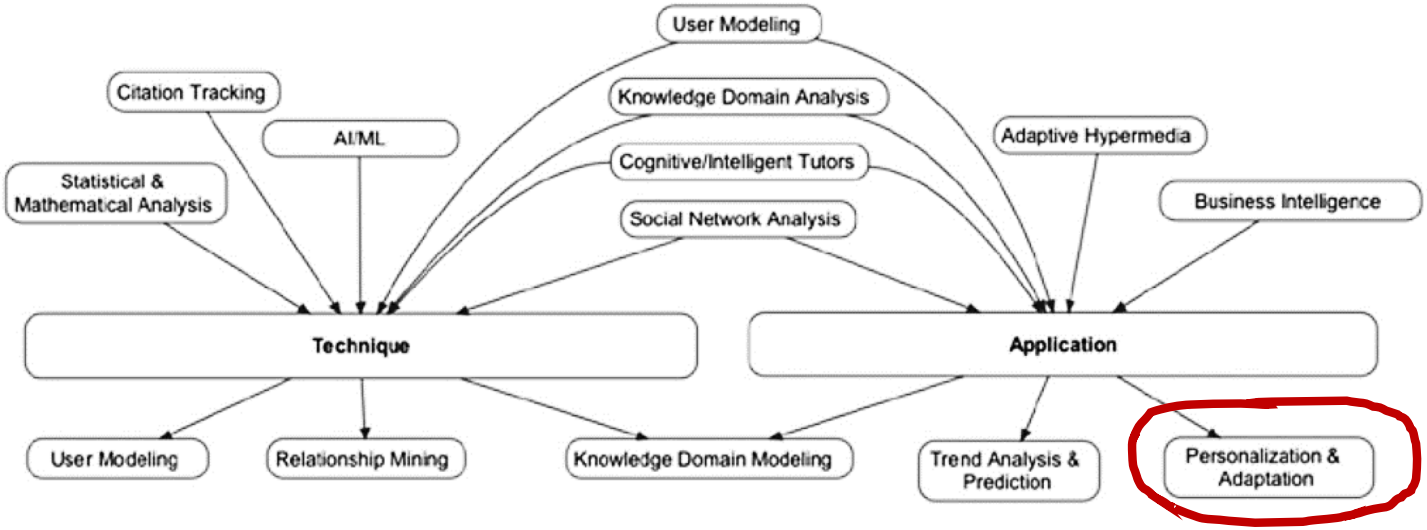
13/06/2018

Personnalisation/Adaptation de l'apprentissage

- ~~*One size fits all*~~
- *by adaptive systems we mean all systems which reflect some features of the user in the user model and apply this model to adapt various aspects of the system to the user* (~ Brusilovsky)
- Historique (rapide partial et partiel)
  - Intelligent Tutoring Systems (ITS)
  - Hypermédias Adaptatifs (AH) (parcours/présentation /contenu)
  - Systèmes de recommandation (ressources/apprenants/forums …)
  - Tests adaptatifs
  - …
  - Systèmes ad-hoc, à base de règles, d'ontologies, de statistiques, structuration a priori vs a posteriori, données +/- massives

Analytique de l'apprentissage

Analytique de l'apprentissage



- Emergence d'une discipline
  - Mieux comprendre l'enseignement, l'apprentissage, la pertinence des contenus, la personnalisation et l'adaptation
  - *Society for Learning Analytics Research*, *International Educational Data Mining Society* et conférences scientifiques

Learning analytics is the measurement, collection, analysis, and reporting of data about learners and their contexts, for the purposes of understanding and optimizing learning and the environments in which it occurs.

Tests sommatifs versus formatifs

- Test sommatif : vise à déterminer un score de l'apprenant, sans diagnostic de connaissances
  - Exemple : TOEIC, TOEFL
  - Psychométrie, Item Response Theory
  - Les réponses sont expliquées par des variables cachées continues
  - Quels paramètre mesurer pour construire un modèle permettant de prédire la performance ?
  - Paramètres calculés à partir des données des étudiants

Tests sommatifs versus formatifs

- Test formatif : vise à évaluer l'apprenant, mais aussi à établir un diagnostic de connaissances
  - Retour possible à l'enseignant où à l'apprenant sur les points maitrisés ou non
  - Les réponses sont expliquées par la maîtrise ou non de composantes de connaissances (KC)
  - Généralement des experts associent les KC aux questions
  - Inférer les KC maitrisées permet de prédire la performance

Tests adaptatifs

- Raccourcir la longueur du test …
  - Les tests sont trop longs
  - Les étudiants sont trop testés
  - Eviter la fatigue et l'ennui
- … mais fournir une évaluation précise
- Comment sélectionner k questions qui soient suffisantes pour prédire les résultats sur les autres ?
- Donc k questions qui résument l'ensemble de questions

	Questions							
	1	2	3	4	5	6	7	8
Alice	0	1	1	1	0	0	0	1
Bob	1	0	1	1	0	0	0	1
Charles	1	0	1	0	0	0	0	0
Daisy	1	0	0	1	1	1	1	1
Everett	1	0	0	0	1	0	0	1
Filipe	0	1	0	1	1	1	1	1
Gwen	0	0	0	1	0	0	1	1
Henry	0	0	0	0	1	0	0	1
Ian	1	1	1	1	0	1	1	0
Jill	0	1	1	1	0	0	1	0
Ken	1	1	1	0	1	1	0	1

Applications des tests adaptatifs

- Tests à faible engagement
  - Si formatifs, les apprenants ont un retour sous la forme des KC maitrisées
  - Révisions avant l'évaluation finale
  - Autoévaluations bénéfiques à l'apprentissage
- Test de positionnement au début d'un MOOC
  - *Il semble que vous ne maitrisiez pas les KC 1 et 3 qui sont des prérequis pour ce MOOC*
  - Personnaliser la suite du MOOC en fonction de ces résultats
  - Recommander des ressources complémentaires pour combler les lacunes

Test adaptatifs : hypothèses

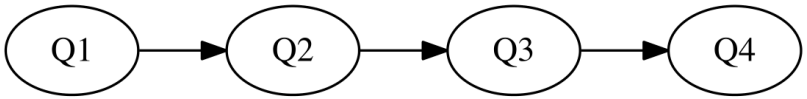
- Test administré électroniquement
- Questions ou tâches binaires
- Le niveau de l'apprenant n'évolue pas pendant le test
- Les questions sont posées une par une et l'apprenant ne peut pas revenir en arrière
- Repose sur deux fonctions :
  - Critère de terminaison
  - Critère de sélection de la question suivante

Algorithme

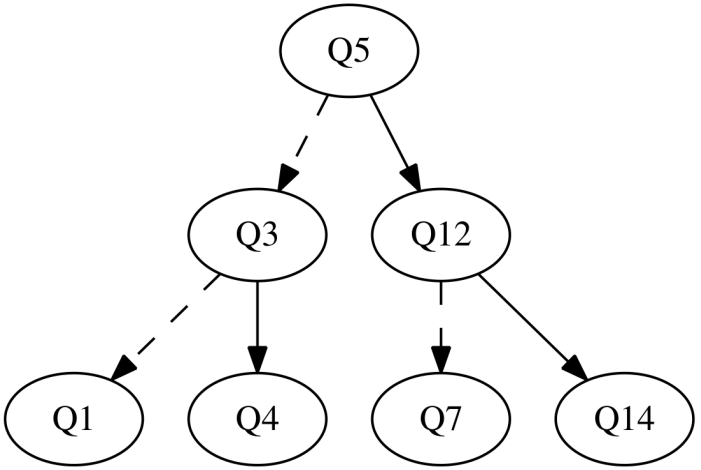
Tant que le critère de terminaison n'est pas vérifié :

- Choisir la question maximisant le critère de sélection de l'item suivant
- La poser au candidat
- Enregistrer le résultat du candidat

Comment personnaliser le test ?



Test non adaptatif



Test adaptatif

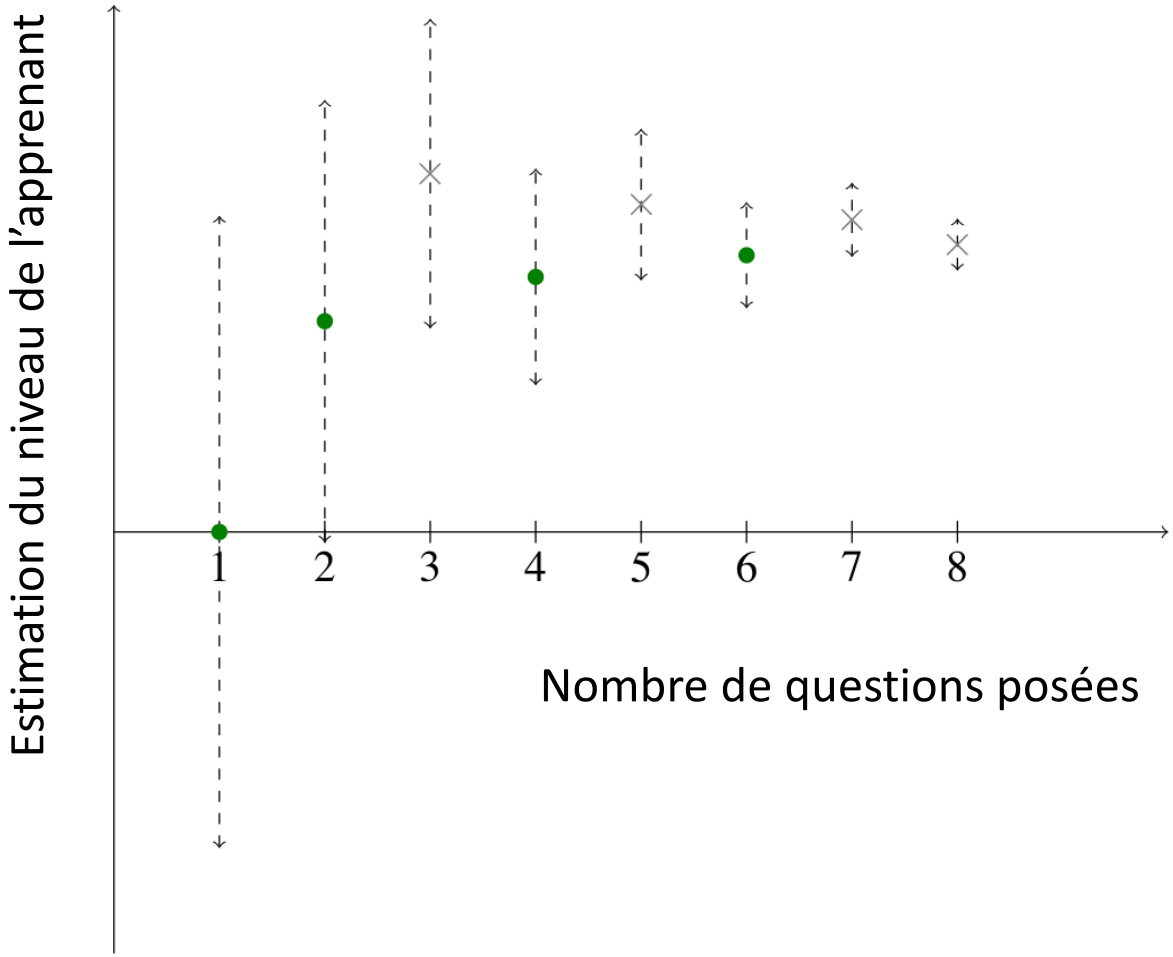
IRT : Item Response Theory

- Modèle le plus simple : modèle logistique à 1 paramètre
- Catégorie fouille de données de l'analytique de l'apprentissage
- Apprenant : 1 paramètre, son niveau
- Question : 1 paramètre, sa difficulté
- Si un apprenant de niveau θ_i veut résoudre une question de difficulté d_j, la probabilité qu'il y arrive :

$$\Pr(\text{« } i \text{ répond à } la\ question\ j \text{ »}) = \Phi(\theta_i - d_j)$$

Où $\Phi(x) = 1/(1 + e^{-x})$ est la fonction logistique

IRT : Item Response Theory



Q-matrices

- Représentation minimale des composantes de connaissance (KC) mises en œuvre dans un test
- Matrice binaire de taille N x K : nombre de questions x nombre de KC
  - $q_{ij} = 0$ si la question i ne fait pas intervenir la KC j
  - $q_{ij} = 1$ si la question i fait intervenir la KC j
- Création manuelle (fastidieux !) ou automatique (mais quel sens donner aux KC trouvées)

Exemple de Q-matrice

	Knowledge components							
	1	2	3	4	5	6	7	8
Item 1	0	0	0	1	0	1	1	0
Item 2	0	0	0	1	0	0	1	0
Item 3	0	0	0	1	0	0	1	0
Item 4	0	1	1	0	1	0	1	0
Item 5	0	1	0	1	0	0	1	1
Item 6	0	0	0	0	0	0	1	0
Item 7	1	1	0	0	0	0	1	0
Item 8	0	0	0	0	0	0	1	0
Item 9	0	1	0	0	0	0	0	0
Item 10	0	1	0	0	1	0	1	1
Item 11	0	1	0	0	1	0	1	0
Item 12	0	0	0	0	0	0	1	1
Item 13	0	1	0	1	1	0	1	0
Item 14	0	1	0	0	0	0	1	0
Item 15	1	0	0	0	0	0	1	0
Item 16	0	1	0	0	0	0	1	0
Item 17	0	1	0	0	1	0	1	0
Item 18	0	1	0	0	1	1	1	0
Item 19	1	1	1	0	1	0	1	0
Item 20	0	1	1	0	1	0	1	0

Description of knowledge components:

1. convert a whole number to a fraction
2. separate a whole number from a fraction
3. simplify before subtracting
4. find a common denominator
5. borrow from whole number part
6. column borrow to subtract the second numerator from the first
7. subtract numerators
8. reduce answers to simplest form

DINA

C est l'ensemble des états latents possibles

- Q-matrice
- Ajout des facteur *guess* g_i et *slip* s_i
  - *Guess* : l'apprenant peut répondre juste par hasard
  - *Slip* : l'apprenant possède les KC mais se trompe en répondant
- Apprenant modélisé par un état latent c de K bits

$$c = (c_1, \dots, c_K) \in C \subseteq \{0,1\}^K$$

- Composante de connaissance k maitrisée si et seulement si $c_k = 1$
- Pr(« réponse correcte à la question i ») = $1 - s_i$ si toutes les composantes de connaissances sont maitrisées, g_i sinon
- Mise à jour bayésienne de l'état latent à chaque réponse
- Problème du trop grand nombre d'états latents (2^K)

GenMA (General Multidimensional Adaptive)

- Théorie des espaces de connaissance (Falmagne 2012)
- Représentation hiérarchique des composantes de connaissances
  - Graphe G = (V, E)
  - V ensemble des KC
  - $(u, v) \in E$ signifie que u doit être maitrisée pour maitriser v
- Permet de réduire drastiquement le nombre d'états latents
- Modèle combiné avec la q-matrice
- Thèse de Jill-Jênn Vie (co-encadrement Popineau, Bourda, Bruillard)

GenMA : Exemple 1

- Supposons 3 KC A, B, C telles que $A \rightarrow B \rightarrow C$
- Ensemble des états possibles :

000	Ne maîtrise rien
100	Maîtrise seulement A
110	Maîtrise A et B
111	Maîtrise A, B et C

- Pas d'autre état latent possible étant donné les relations de prérequis
- Distribution de probabilité initiale $\pi_0 = 0{,}25$
- $H(\pi_0) = 2$

GenMA : Exemple 1

- Quelle question lui poser parmi Q_A, Q_B ou Q_C ?
- Au début : 75% de chances de répondre à Q_A, 50% à Q_B et 25% à Q_C

	Q_A	Q_B	Q_C
Réponse correcte Etats latents Entropie	75% 100, 110, 111 1,6	50% 110, 111 1	25% 111 1
Réponse incorrecte Etats latents Entropie	25% 000 1	50% 000,100 1	75% 000,100,110 1,6
Entropie moyenne	1,2	1	1,2

- Choix de la question B qui réduit le plus l'entropie

GenMA : Exemple 2

Supposons que l'apprenant maîtrise toutes les KC sauf Banach et Hilbert.

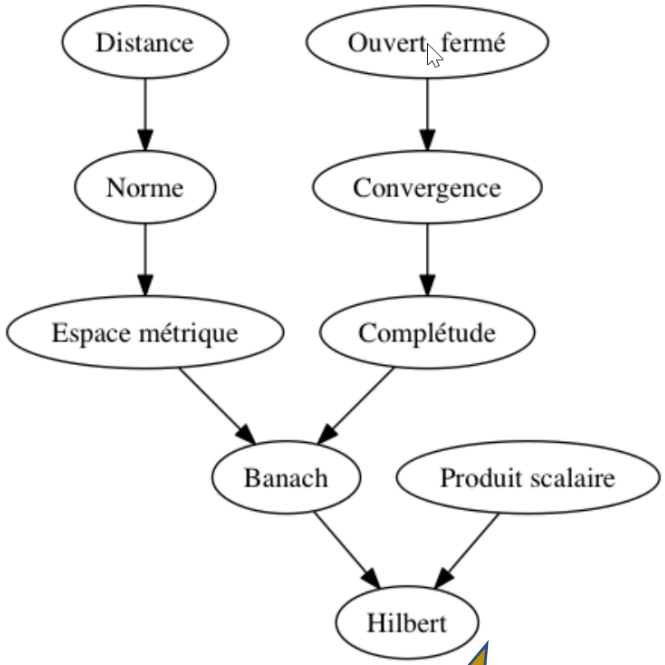
Déroulement d'un test :

- Q1 : est-ce que l'apprenant maîtrise « produit scalaire »?
  - Oui
- Q2 : est-ce que l'apprenant maîtrise « convergence » ?
  - Oui
- Q3 : est-ce que l'apprenant maîtrise « espace métrique » ?
  - Oui

Distribution de probabilité : 0 pour chaque état latent sauf 4, équiprobables avec 0,25 :

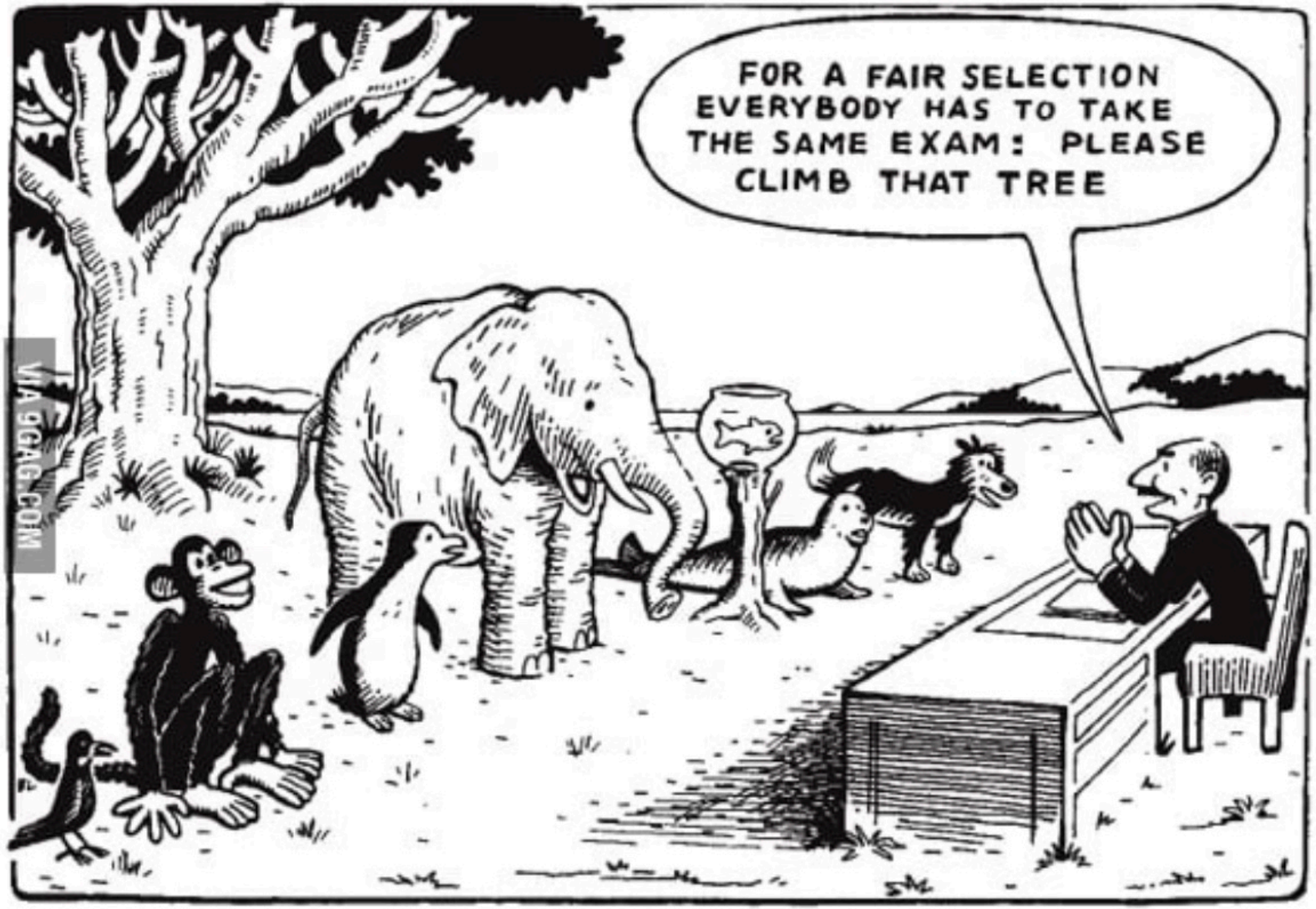
  - apprenant maîtrise tout
  - apprenant maîtrise tout sauf Complétude, Banach, Hilbert
  - apprenant maîtrise tout sauf Banach, Hilbert
  - apprenant maîtrise tout sauf Hilbert

- Q4 : est-ce que l'apprenant maîtrise « Banach » ?
  - Non
- Q5 : est-ce que l'apprenant maîtrise « complétude » ?
  - Oui



5 questions posées au lieu de 9

GenMA

- Testé sur des données réelles de MOOC
  - Pas besoin d'apprentissage, seulement graphe de prérequis
  - Arrêt lorsque la certitude est supérieure à 95%
  - Slip et guess = même valeur ϵ (si ϵ augmente, le nombre de questions augmente)
  - Diminution du nombre de questions

https://9gag.com/gag/4239400