Un nouvel outil pour l'évaluation des compétences des étudiants : les tests de concordance de scripts

Ecole Centrale de Nantes

Morgan Magnin (ECN/IRCCyN), Guillaume Moreau (ECN/CERMA) – Enseignants-chercheurs

{Morgan.Magnin|Guillaume.Moreau}@ec-nantes.fr - 02 40 37 {16 28|68 47} - http://eat-tice.ec-nantes.fr

Les formations professionnalisantes telles qu'un cursus d'ingénierie doivent permettre aux étudiants d'acquérir non seulement les connaissances, mais aussi les compétences nécessaires à l'exercice de leur future profession. En marge des dispositifs permettant de vérifier la bonne acquisition des savoirs (exercices, questions de cours, etc.), il est donc primordial de disposer d'outils pour valider le savoir-faire des élèves.

Afin de cibler certaines compétences primordiales, l'Ecole Centrale de Nantes a participé à un projet UNIT sur les tests de concordance de scripts (TCS), une forme de généralisation des questionnaires à choix multiples (QCM), jusque là réservés au domaine médical pour l'évaluation des capacités de diagnostic des étudiants en médecine [1]. Comme leur nom le laisse penser, les TCS sont basés sur la théorie cognitive dite « des scripts », qui se focalise sur les réseaux de connaissance qu'utilisent les individus en situation de résolution de problème.

Le principe d'un TCS est le suivant : une situation problématique, suivie d'une hypothèse pertinente, est présentée. Une nouvelle information est alors introduite, et l'étudiant est invité à évaluer en quoi l'ajout de cette donnée augmente, diminue ou est sans influence sur la qualité de l'hypothèse initiale. Le problème étant complexe, et partiellement décrit, les analyses des experts varient. Il n'existe donc pas une unique « bonne » réponse. De fait, l'établissement du score se fait en comparant les réponses de l'étudiant à celui d'un panel de référence (constitué d'experts du domaine ayant préalablement effectué le test). L'étudiant reçoit un crédit de points proportionnel au nombre d'experts du panel qui ont choisi la même réponse.

Dans un premier temps, nous avons adapté les TCS pour évaluer les compétences des étudiants en informatique. Depuis juin 2009, nous les avons expérimentés dans le cadre de différents enseignements au niveau L3 et M2, pour des groupes de 30 à 120 étudiants. Cette expérience [2], la première menée en dehors du secteur médical, nous a permis de d'établir les forces et les limites des TCS en informatique, une analyse qui ouvre la porte à un usage plus large de ce dispositif dans d'autres disciplines.

Le recours aux TCS pour l'évaluation du raisonnement en contexte d'incertitude soulève généralement deux freins : d'une part l'expertise présumée des experts, d'autre part le peu d'informations retirées sur le processus de raisonnement de l'étudiant. La première limite peut être surmontée en composant avec soin le panel d'experts. Pour ce faire, nous avons sollicité des personnes avec un diplôme de niveau Master et au moins trois ans d'expérience pratique dans le domaine. La seconde limite est plus délicate à prendre en compte, car des experts peuvent avoir des processus de raisonnement (i.e. des scripts) différents, mais aboutir à des conclusions similaires. Nous avons dès lors choisi d'enrichir les TCS d'un champ « Justification » permettant à l'élève d'argumenter sa réflexion.

Les TCS ne sont pas adaptés à toutes les situations : ils visent plutôt des configurations incertaines et/ou ouvertes, dans lesquelles on n'a pas forcément accès à toutes les données. Leur valeur ajoutée est plus limitée dans des configurations parfaitement définies. Ils vont dénoter d'un savoir-faire par rapport à une situation incertaine, là où les QCM « classiques » vont permettre d'évaluer un savoir par rapport à une situation très bien définie. C'est un outil complémentaire d'approches plus traditionnelles. Les TCS sont particulièrement adaptés dans des domaines tels que l'analyse des systèmes d'information, le droit informatique ou la programmation (débogage). Ils s'appliquent efficacement pour la conception et le diagnostic. Ils permettent d'évaluer (ou d'aider à l'auto-évaluation) les étudiants de manière pertinente, efficace et rapide.

Nous comptons évaluer désormais élargir le dispositif et appliquer les TCS dans d'autres disciplines des sciences de l'ingénieur.

[1] Charlin B., Gagnon R., Kazi-Tani D. and Thivierge R. (2006) "*Le test concordance comme outil d'évaluation en ligne du raisonnement des professionnels en situation d'incertitude*". Revue internationale des technologies en pédagogie universitaire. 2(1).

[2] M. Magnin et G. Moreau. "*Utilisation des Tests de Concordance de Scripts pour l'evaluation en informatique*". Communication scientifique. Dans : 7eme Colloque International Technologies de l'Information et de la Communication pour l'Enseignement (TICE), 8 pages, Nancy-Metz, Decembre 2010.

Annexes

Scénario	Pour payer ses personnels dans une société d'intérim, le responsable de la paie édite la liste des heures effectuées par personne à partir des éléments contenus dans le système d'information. Lors de cette édition, Xavier Martin apparaît deux fois.
Hypothèse	Vous supposez qu'il y a un doublon dans la base de données.
Nouvelle donnée	Les dates de naissance de ces occurrences sont différentes.
Échelle de Likert	Quel est l'effet de cette nouvelle donnée sur l'hypothèse ? [ ] rend l'hypothèse beaucoup plus probable [ ] rend l'hypothèse plus probable [ ] n'a aucun effet sur l'hypothèse [ ] rend l'hypothèse moins probable [ ] rend l'hypothèse beaucoup moins probable

Tableau 1 • Exemple de TCS appliqué au domaine des systèmes d'information (compétence ciblée : identifier un problème d'accès concurrent)

Commentaires sur le TCS présenté sur le Tableau 1 : a priori, ce test peut laisser penser à un doublon dans la base de données. Une date de naissance différente conduit naturellement à deux hypothèses (d'ailleurs non exclusives) : une simple homonymie ou une erreur sur une ou plusieurs dates de naissance. Ce dernier cas n'exclut pas le doublon, mais ne le valide pas totalement non plus : il n'est pas impossible, même si c'est très peu probable, d'avoir affaire à deux personnes de même nom et de dates de naissance identiques. Enfin, on peut aussi penser à une erreur de saisie des noms. Il est assez facile de tirer des cas dérivés de ce premier test.

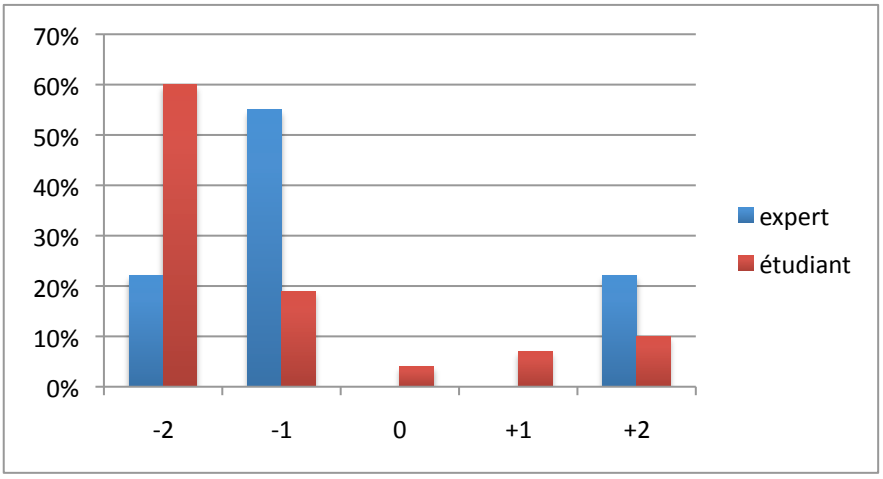

Figure 1 : Répartition des réponses des experts et des étudiants sur le TCS présenté dans le Tableau 1. L'abscisse représente l'éventualité choisie : -2 correspond à « rend l'hypothèse beaucoup moins probable », -1 « moins probable », 0 « n'a aucun effet sur l'hypothèse », etc. Réponses obtenues auprès d'un panel de 10 experts et 110 étudiants.

Commentaire sur la figure 1 : nous avons constaté que, sur certaines questions, les experts étaient moins catégoriques que les étudiants. Il en va ainsi de notre question présentée sur le Tableau 1. Par rapport à l'effet de la nouvelle donnée sur l'hypothèse formulée, 20% des experts la jugent « beaucoup moins probable » tandis que 60% des experts préfèrent un avis plus nuancé en répondant seulement « moins probable ». Les étudiants, eux, choisissent, pour 60% d'entre eux, la formulation « beaucoup moins probable » et 19% l'hypothèse « moins probable ». On assiste là à une inversion presque parfaite des réponses entre enseignants et étudiants. Si les deux réponses vont dans le même sens (l'hypothèse paraît confirmée par la

nouvelle donnée), leur gradation n'est pas envisagée de la même manière par les experts d'une part et les étudiants d'autre part. Cette tendance des étudiants à choisir des éventualités plus fortes que les experts s'est retrouvée confirmée sur les autres questions. Échaudés par des cas pratiques souvent difficiles à dénouer, les experts ont certainement plus tendance à une certaine réserve que des étudiants habitués, de par leur formation passée (une très large majorité de nos élèves sont issus de classes préparatoires scientifiques), à des problèmes « parfaitement posés » et à répondre à des problèmes.

Scénario	Pierre et Paul ont écrit chacun la moitié d'un programme en C dans deux fichiers x.c et y.c. Une fois compilé correctement et les liens édités sans problème, le programme plante.
Hypothèse	Vous supposez que Pierre a tort car il est habituellement moins fort en programmation.
Nouvelle donnée	Paul a vérifié le fonctionnement de toutes ses fonctions.

Tableau 2 • Exemple de TCS appliqué au domaine de la programmation (compétence ciblée : identifier le point de faiblesse d'un processus de programmation) – l'échelle de Likert est omise pour des raisons de place.

Commentaire sur le TCS présenté sur le Tableau 2 : l'exemple propose une démarche de type diagnostic. Il s'agit de comprendre les causes du non-fonctionnement d'un programme. Il y a deux grandes hypothèses restantes ici : soit c'est Paul qui n'a pas réalisé de tests unitaires suffisants sur les fonctions qu'il a écrites, soit c'est l'intégration qui a posé problème.

Scénario	Vous venez d'être nommé Correspondant Informatique et Liberté dans votre entreprise. Parmi les adresses e-mail figurant dans les fichiers de prospection publicitaire de l'entreprise, vous soupçonnez que certaines ont été collectées sans l'accord préalable des personnes à recevoir des messages commerciaux. Mais les fichiers ne contiennent aucune information sur le fait que les personnes ont - ou non - donné leur consentement.
Hypothèse	Vous supposez qu'il est nécessaire de supprimer l'intégralité des fichiers d'adresses e-mail.
Nouvelle donnée	Vous apprenez que ces adresses e-mail appartiennent majoritairement à des personnes morales. Le reste des adresses e-mail appartiennent à d'anciens clients de l'entreprise.

Tableau 3 • Exemple de TCS appliqué au domaine du droit dans la société de l'information (compétence ciblée : analyser une situation sensible vis-à-vis du droit du traitement de l'information et du commerce électronique) – l'échelle de Likert est omise pour des raisons de place.

Commentaire sur le TCS présenté sur le Tableau 3 : le TCS présenté ici illustre qu'en situation professionnelle, il arrive souvent que nous ne possédions pas l'intégralité des informations nécessaires à la prise d'une décision parfaitement déterministe dans le domaine juridique. Il importe donc de raisonner sur les éléments à notre disposition, en étant le plus proche possible de l'esprit du législateur. Ici, la nouvelle donnée nous apprend que, potentiellement, 100% du fichier de prospection publicitaire est légal (le droit français accorde l'autorisation de prospecter par courriel - sans accord préalable - des personnes morales et d'anciens clients n'ayant pas manifesté leur refus par une demande de désinscription). Les réponses des experts se répartissent dès lors entre « rend l'hypothèse moins probable » et « beaucoup moins probable », en fonction de la sensibilité de chacun (sachant qu'il manque une information pour conclure définitivement : est-ce que certaines personnes dans le fichier avaient manifesté leur volonté de ne plus recevoir de courriels ?).