

Mécanismes de compression d'information en mémoire de travail

Stage proposé par Sophie Portrat (psychologie cognitive) et Benoît Lemaire (informatique)
Laboratoire de Psychologie et NeuroCognition (CNRS UMR 5105)
Université Pierre-Mendès-France, Grenoble

Sophie.Portrat@upmf-grenoble.fr ou Benoit.Lemaire@upmf-grenoble.fr

Le système cognitif doit faire face à un monde qui est extrêmement complexe, mais qui est, néanmoins, hautement régulier. Ces régularités sont cruciales (Chater, 1999).

Détecter ces régularités de notre environnement est en effet très important pour pouvoir percevoir et agir de manière efficace. Si nous vivions dans un monde aléatoire, nous ne pourrions ni comprendre, ni prédire. Étonnamment, ces mécanismes de détection de régularités ont été peu étudiés en ce qui concerne la capacité de la mémoire de travail.

Depuis Miller (1956), de nombreux travaux ont cherché à estimer la capacité de notre mémoire de travail. On a parlé d'une capacité de 7 ± 2 unités, on parle plus volontiers maintenant de 4 unités. Mais les stimuli qui sont présentés dans ces expériences sont volontairement dénués de régularités, afin de se concentrer uniquement sur la capacité. Par exemple, on fait mémoriser aux participants des séquences d'objets qui sont toutes soigneusement indépendantes les unes des autres. Or, les régularités permettent de compresser l'information, ce qui est crucial dans une perspective de stockage. Il est donc probable que des mécanismes cognitifs de compression de l'information existent, qui nous permettent d'augmenter les performances de mémorisation.

On sait par exemple que, dans un flux de données, les variations de fréquences des objets qui composent ce flux permettent de compresser l'information. Si tous les objets apparaissent avec la même fréquence, il n'y a aucune raison qu'ils soient encodés avec des codes de tailles différentes. Mais s'il y a des variations de fréquences, les objets les plus fréquents peuvent être codés avec des codes de tailles plus petites, ce qui permet donc de gagner de la place. En d'autres termes, les objets très fréquents peuvent être codés avec des codes très petits et les objets peu fréquents peuvent sans inconvénient être codés avec des codes plus longs. Au total, on aura gagné de la place...

Par exemple, une séquence comme ADCBBCADDCBACADB, dans laquelle les 4 objets A,B,C et D ont la même fréquence ($\frac{1}{4}$) est ainsi beaucoup moins compressible qu'une séquence comme ABAADBACDABACBAA qui comporte des fréquences variables ($\frac{1}{2}$ pour A, $\frac{1}{4}$ pour B et $\frac{1}{8}$ pour C et D). Si on code chaque lettre par des chiffres binaires (ce qui est le langage de codage le plus simple qu'on puisse imaginer), on attribuera des codes de taille 2 dans le premier exemple (comme A=00, B=01, C=10, D=11), ce qui prendra au total $16 \times 2 = 32$ bits. En revanche, on pourra faire mieux dans le second exemple grâce aux fréquences variables (par exemple, A=0, B=10, C=110, D=111) et coder la séquence avec seulement $8 \times 1 + 4 \times 2 + 2 \times 3 + 2 \times 3 = 28$ bits. Ce sont d'ailleurs des mécanismes similaires qui sont utilisés dans les logiciels de compression de fichiers.

Si on montre que ces mécanismes de compression existent en mémoire de travail, ce que des articles récents (p. ex., Brady et al., 2009) laissent penser, alors la mémoire de travail devrait être exprimée plus justement en terme de quantité d'information plutôt qu'en terme de nombre d'unités, comme c'est actuellement le cas.

L'idée de ce stage est donc de s'intéresser à la capacité de la mémoire de travail, mais en étudiant les mécanismes de compression de l'information à l'œuvre. Des expériences existent déjà au sein du laboratoire (Robinet, Portrat & Lemaire, 2011) ainsi des modèles computationnels qui simulent l'activité cognitive (Robinet et al., in press).

L'encadrement de ce stage est délibérément pluri-disciplinaire. Selon le parcours et les envies de l'étudiant, on pourra orienter sa contribution vers une approche expérimentale du problème ou vers une approche de modélisation.

Références

- Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in visual working memory: using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology: General*, 138(4), 487-502.
- Chater, N. (1999). The search for simplicity : A fundamental cognitive principle? *The Quarterly Journal of Experimental Psychology A*, 52 , 273-302.
- Miller, G. A. (1956). The magical number seven plus or minor two : some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- Robinet, V., Lemaire, B., Gordon, M.B. (in press). MDLChunker: a MDL-based Cognitive Model of Inductive Learning. *Cognitive Science*.
- Robinet, V., Portrat, S., Lemaire, B. (2011). Working memory capacity as a quantity of information. *42th Meeting of the European Mathematical Psychology Group*, Paris, August 2011.