

DeepExpert : vers une Intelligence Artificielle autonome et explicable

Introduction

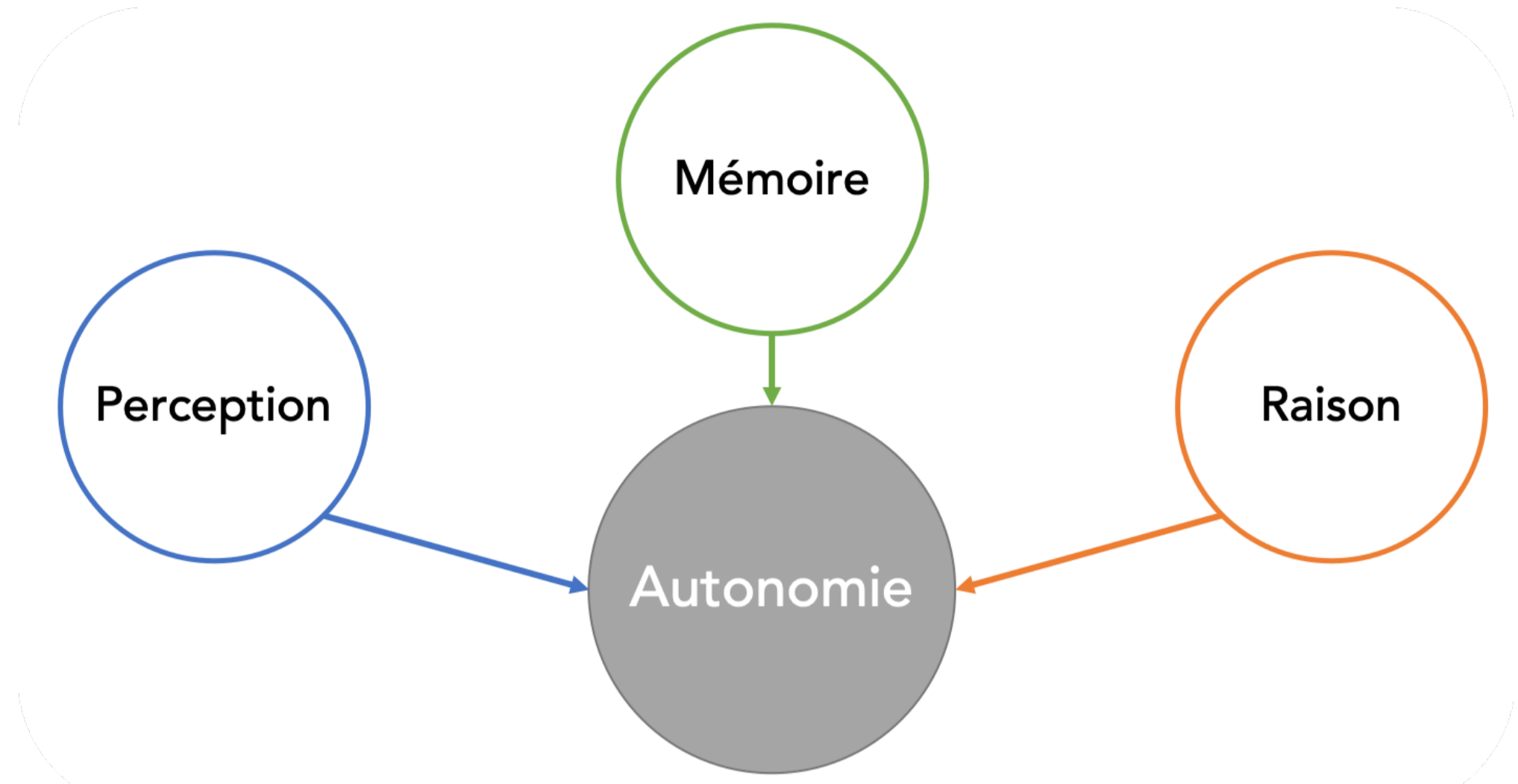


Figure 1 – Les fonctions cognitives de l'Homme [1] appuyant l'autonomie

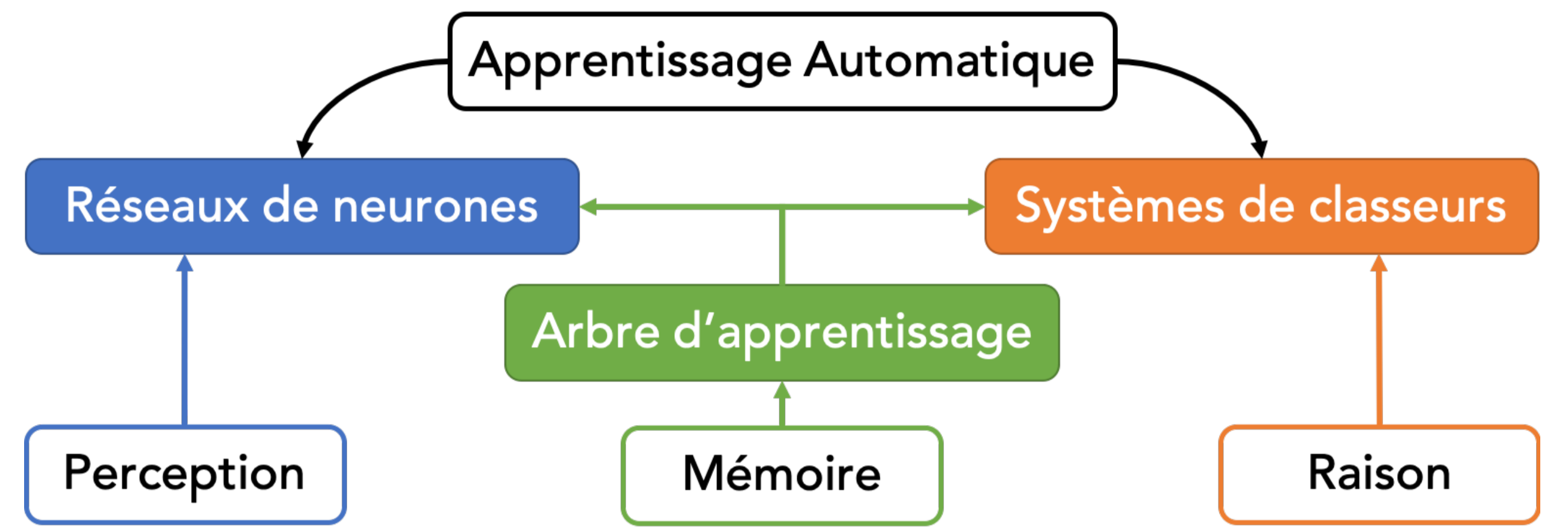


Figure 2 – Couplage visant à développer l'autonomie des ordinateurs se basant sur les fonctions cognitives [2,3]

Gagner en autonomie par l'expérience

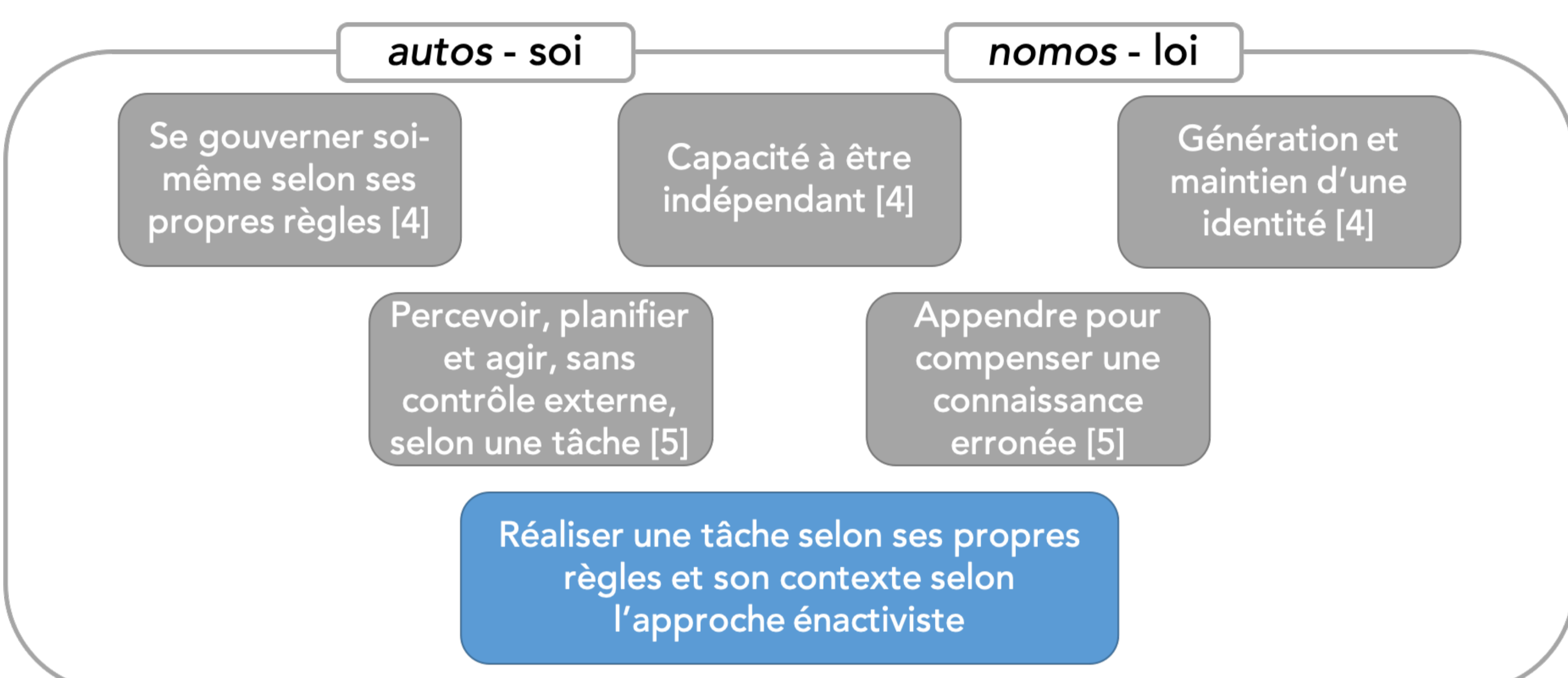


Figure 3 – Différentes vision de l'autonomie et notre définition

Mise en œuvre de DeepExpert

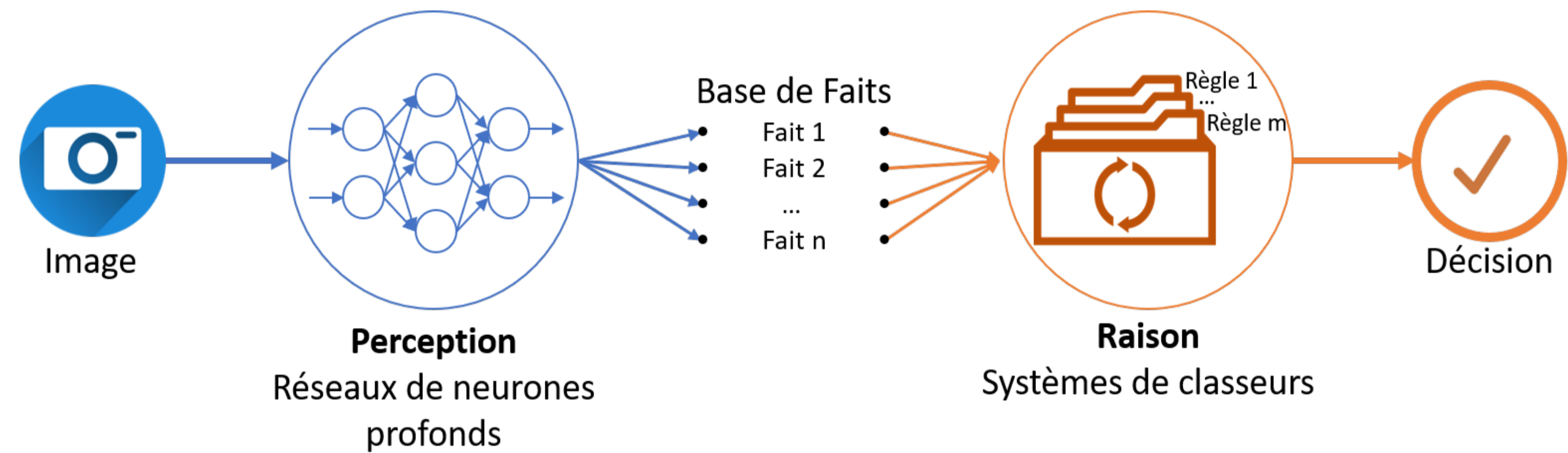


Figure 4 – Illustration de DeepExpert : un couplage entre réseaux de neurones profonds et systèmes de classeurs

Légende :

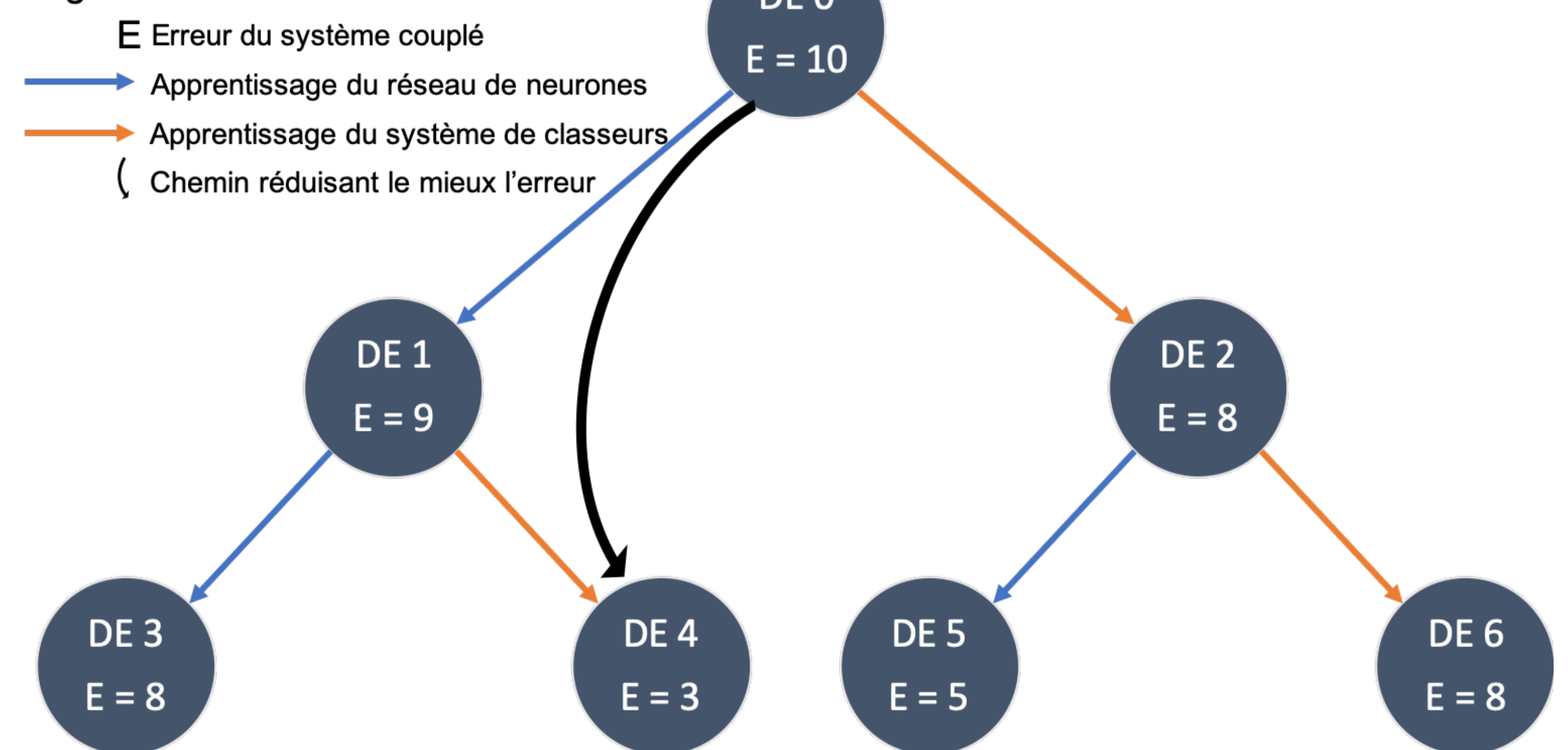


Figure 5 – Arbre d'apprentissage de DeepExpert

Approches existantes

- S'appuyant sur la théorie neurobiologique, les *Neural Classifier Systems* (NCS) sont des systèmes de classeurs où chaque classer est représenté par un réseau de neurones modulable, afin de permettre un apprentissage plus flexible aux modifications de l'environnement [7].
- Un couplage fondé sur un réseau de neurones hybride (réseau de classification et auto-encodeur, nommé le *Deep Classification Autoencoder*) et un système de classeurs basé sur la précision (XCSR) a été proposé récemment et permet d'accroître les performances obtenues comparativement à un XCSR seul [8].
- L'idée de mémoire se retrouve sous différentes formes : par exemple, dans la représentation des classeurs [3], ou bien, dans les *Neural Maps* [9]. Ces derniers sont des systèmes de mémoire externes, opaques et conçus par l'Homme dans le but de réaliser une tâche en particulier.

Perspectives

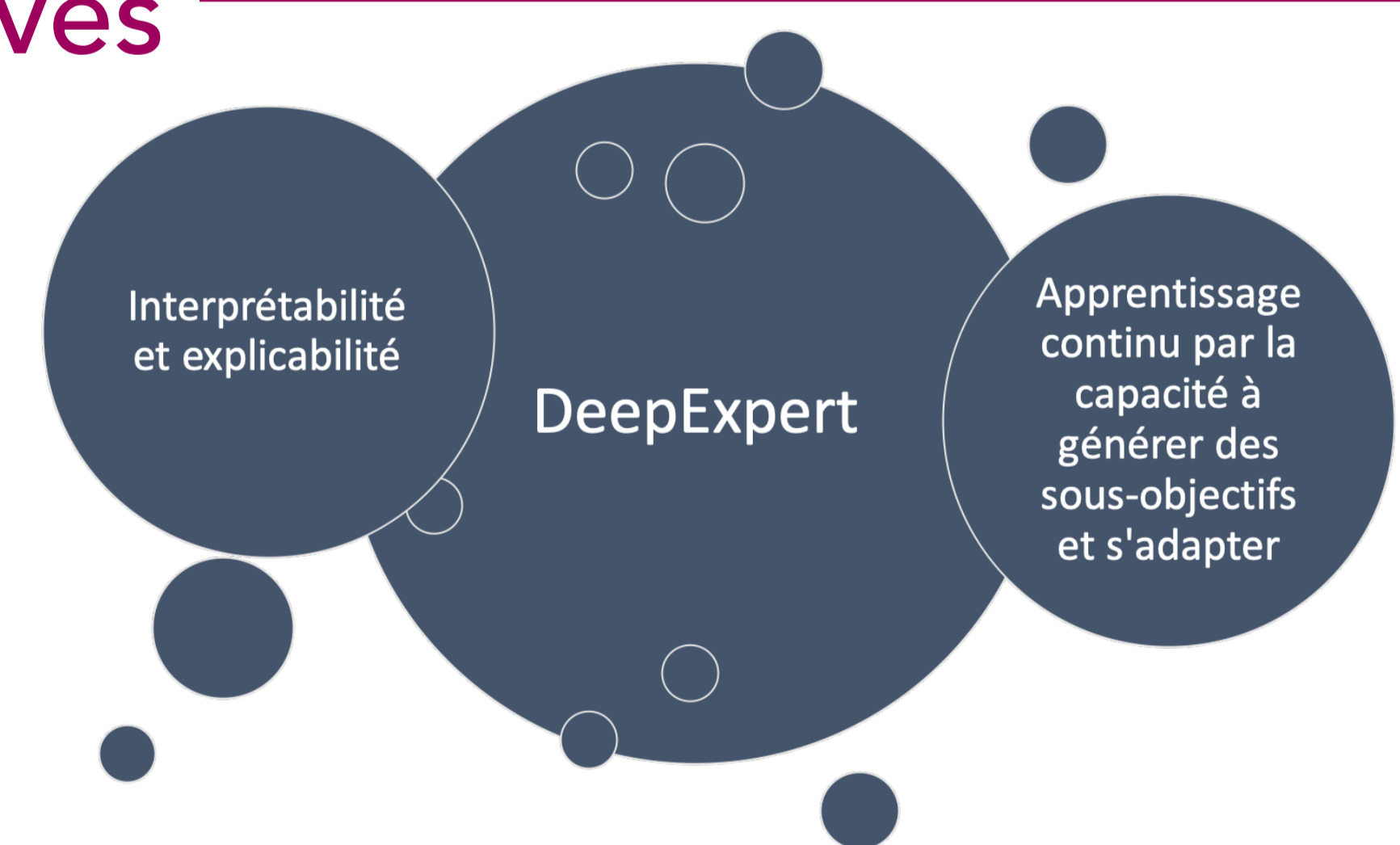


Figure 6 – Prochains axes de recherche associés à DeepExpert

Références

- [1] M. Serres, Les nouvelles technologies : révolution culturelle et cognitive, *Interstices*, INRIA, 2007.
- [2] Y. LeCun, Y. Bengio et G. Hinton, Deep learning, *Nature*, vol. 521, no. 7533, p.436-444, 2015.
- [3] C. Buche, C. Septseault et P. de Loor, Les systèmes de classeurs. Une présentation générale, *TSI*, vol.25, no. 8-9, p. 963-990, 2006
- [4] K. Manac'H, *Vers la notion d'agent énaïviste virtuel : Application à l'approche dynamique évolutionnaire*, Université Européenne de Bretagne, 2011.
- [5] J. Beer, A. Fisk et W. Rogers, Toward a framework for levels of robot autonomy in human-robot interaction, *Journal of human-robot interaction*, vol. 3, no.2, p.74-99, 2014.
- [6] P. de Loor, K. Manac'H et J. Tisseau, Enaction-based artificial intelligence : Toward co-evolution with humans in the loop, *Minds and Machines*, vol. 19, no. 3, p. 319-343, 2009.
- [7] L. Bull, On Using Constructivism in Neural Classifier Systems, *Parallel Problem Solving from Nature – PPSN VII*, p. 558-567, 2002.
- [8] K. Matsumoto, R. Takano, T. Tatsumi, H. Sato, T. Kovacs et K. Takadama, XCSR based on compressed input by deep neural network for high dimensional data, *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, p. 1418-1425, 2018.
- [9] E. Parisotto et R. Salakhutdinov, Neural map : Structured memory for deep reinforcement learning, arXiv : 1702.08360, 2017.