Conceptual systems align to aid concept learning

Article (Accepted Version)

Horst, Jessica S and Bird, Chris M (2020) Conceptual systems align to aid concept learning. Nature Machine Intelligence, 2. pp. 92-93. ISSN 2522-5839

Conceptual systems align to aid concept learning

Jessica S. Horst, Chris M. Bird

University of Sussex

Both as children and adults, we can learn about the world in different ways, such as via direct instruction, reading books or watching media. Although several studies explore whether these different input types lead to different success rates[1], questions remain about whether the underlying conceptual understanding is the same. For example, is the conceptual understanding of wizarding the same for a child who has read *Harry Potter and The Philosopher's Stone* as it is for a child who has listened to the audiobook or one who has seen the movie adaptation? Moreover, can a rich conceptual understanding in one domain facilitate learning related information in another domain? For example, does reading the *Harry Potter* series help one understand a movie about fantastic beasts? Recently, Roads and Love[2] used computational modelling to examine how well our conceptual understanding aligns across modalities – even when learning within each modality occurs independently and at a different time.

By putting the focus on alignability and using that to update learning across modalities we can have less reliance on temporal dynamics. As far back as Quine's[3] famous articulation of the referential ambiguity problem (a new word could hypothetically have many meanings), it has been accepted that what else is happening when a new word is first encountered plays a key role in learning, in real world situations. Moreover, Hebbian learning[4], the cornerstone of connectionist modelling, relies on input occurring at the same time. Roads and Love's reliance on the unique structural profiles of concepts now presents an innovative and exciting method for exploring multi-modal learning when the input is not encountered simultaneously.

To explore how conceptual systems might align, Roads and Love first created three separate conceptual systems for later comparison based on concept co-occurrences in written text, images and audio clips (see Figure 1). Next, they used the "ground truth" of the correct mappings between concepts (e.g. the concept "sheep" might be present in all three systems) to experimentally manipulate the proportion of correct mappings between the systems. For example, they might constrain the proportion of correct mappings between the text system and the image system to be 0.25 or 0.75, while all other mappings are randomly allocated. Then, within each system, the authors constructed distance matrices between every concept, similar to the distance look-up tables between cities that are found at the back of road atlases. These distance matrices were then correlated between systems, to calculate what the authors termed the “alignment correction”. If the systems are indeed aligned, then the alignment correlations should be higher when 0.75 of the mappings are correct, compared to when 0.25 of the mappings are correct. This is what the authors found. The authors sampled 10,000 uniquely generated pairings for every level of mapping accuracy and demonstrated a clear relationship between mapping accuracy and alignment correlation.

The finding that the different conceptual systems are aligned might be exploited by the brain in the service of new learning. A key property of *supervised learning* is that it enables direct mapping between conceptual systems. For example, although unsupervised pattern recognition enables recognition of the same individual when they are encountered again, learning the name of that individual usually requires someone to explicitly provide this information. Even in the case where associations are learnt across multiple situations, the information is still provided by someone and therefore learning is still supervised to some extent. Roads and Love’s study suggests that information learnt in one (unsupervised) system, such as nouns, could be directly applied to information learnt in another (unsupervised) system, such as visual object categorisation by comparing the corresponding object occupying the same space in another conceptual system. That is, the alignment of similar conceptual mappings in both systems allows transfer of information from one modality to another even when input is presented at different points in time in an unsupervised manner.

By demonstrating that unsupervised learning can occur in the absence of temporal relationships, Roads and Love’s approach may contribute to our understanding of how we do learn words for concepts with low imageability or the names of absent objects[5]. For example, when 17-18-month-old children hear an adult use a new word while tapping a spot on an empty table where a novel object was previously, they will associate the word to the object although they never encountered the word and object at the same time (though see also [6]).

Follow-up analyses revealed that the number of concepts in the system dramatically impacts how well concepts map on to each other across systems with richer (larger) semantic systems corresponding more closely. The more concepts there are, the more unique the structural relationships between the different concepts. Given that semantic systems comprising fewer concepts align less well, the authors tested whether this was still the case when selecting concepts that are typically acquired early. They found that early learnt concepts tend to have relatively distinct and distinguishable conceptual signatures within semantic space, which may aid in successful mapping between systems. This finding may have implications for understanding the so-called 'vocabulary spurt'[7] (a period of sudden, rapid vocabulary growth in the second year). McMurray[8] has demonstrated this phenomenon is a product of how words are distributed across levels of difficulty (i.e., there are easy words and difficult words). Roads and Love's simulations suggest it may be the current structure of the child's conceptual system that determines whether a word is difficult. As such, this has additional implications for children with other learning profiles (e.g., late-talkers). Future research should explore whether unsupervised conceptual alignability can capture these individual differences.

Synonyms should intrinsically have nearly identical conceptual profiles. Future research on conceptual alignabilty may illuminate how children learn synonyms and nested categories (e.g., a sheep is a kind of mammal). The conceptual profiles for synonyms should be more similar to each other than to any other item in the system and profiles should also be highly similar for items that are nested.

Overall, this paper represents an innovative initial step in exploring alignment between conceptual systems and demonstrates how alignment can enable cross-modal mappings to be acquired without the requirement of either supervision or that information from multiple modalities is presented at the same time. The "algorithm-agnostic approach" paves the way for future work to explore the specific mechanisms.

References

[1] Valentini, A., Ricketts, J., Pye, R. E., & Houston-Price, C. *Journal of experimental child psychology*, **167**, 10-31, (2018)

[2] Roads, B. D. & Love, B. C. *Nature Machine Intelligence*, ????, ????? (2019).

[3] Quine, W. V. O. Word and Object, Cambridge, MA: MIT Press (1960).

[4] Hebb, D. O. The Organization of Behavior, John Wiley (1949)

[5] Samuelson, L. K., Smith, L. B., Perry, L. K., & Spencer, J. P. *PloS ONE*, **6**, e28095, (2011).

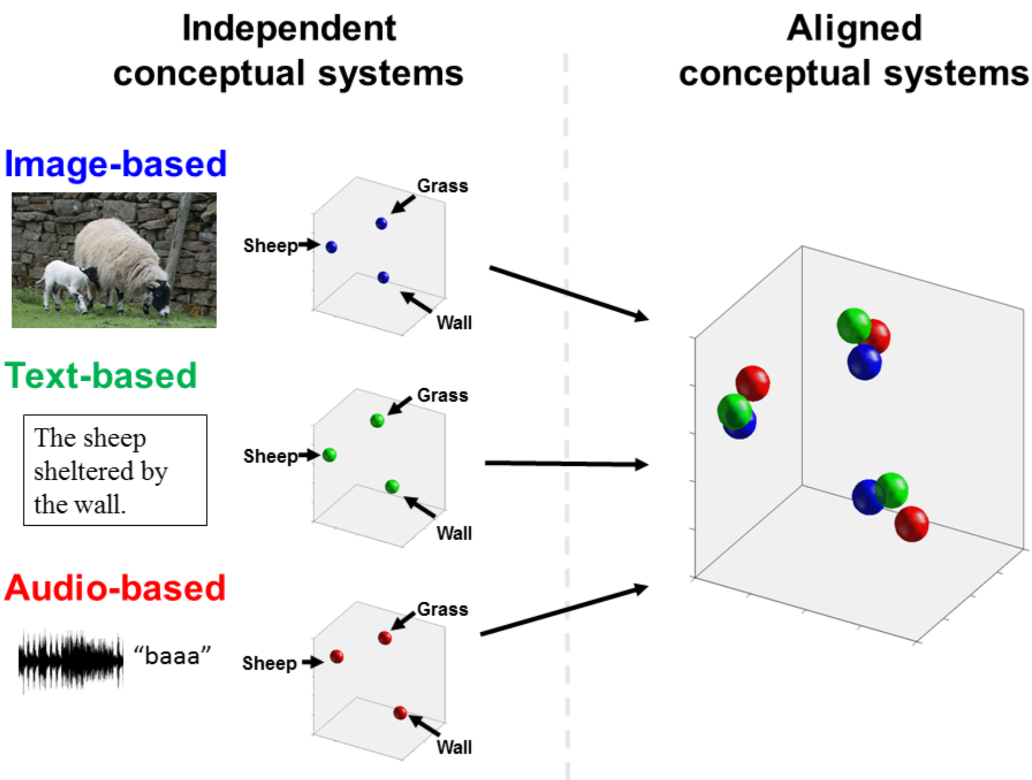
[6] Samuelson, L. K., Kucker, S. C., & Spencer, J. P. *Cognitive Science*, **41**, 52-72, (2017).

[7] Goldfield, B. A., & Reznick, J. S. *Journal of child language*, **17**, 171-183, (1990)

[8] McMurray, B. *Science*, **317**, 631-631, (2007).

Figure caption

Independent conceptual systems were assembled on the basis of unsupervised learning of co-occurrence of objects in images, written text and short audio files. For example the concepts, “sheep”, “grass” and “wall” may have locations within each conceptual system (although in reality the audio system contains much fewer concepts, since it is based on items that make noise). Roads and Love demonstrated that these independent conceptual systems are aligned such that the same concepts from one system occupy similar locations in the multidimensional space of another system.


Independent conceptual systems
Aligned conceptual systems
Image-based
Grass
Sheep
Wall
Text-based
The sheep sheltered by the wall.
Grass
Sheep
Wall
Audio-based
"baaa"
Grass
Sheep
Wall