

Neural information and computational formats

PaC, 17-19 XII 2024, Paris

Wojtek Mamak,
Section for Logic and
Cognitive Science, PAS

Overview

1. Intro – why bother with formats? Some received views
2. New idea – computational profiles (Coelho Mollo & Vernazzani 2023)
3. Empirical examples! Evidence from spatial nav, games and neural recording.
4. Exploitability and decodability of brain information
5. How to combine it? (Kulvicki 2007; 2009; Larkin & Simon 1987)
6. Conclusions

Why do we want or need formats?

- ❖ Computational theories of mind (CTM) – how to explain representation through information. Three ingredients: contents (*~what's stored*), vehicles (*~what stores?*) and formats (*~how stores?*). *Conflated all the time. Maybe deflate then?*
- ❖ Main intuition: same information can be formatted differently – ramifications for how it can be processed *downstream (legibility + exploitability)*
- ❖ However, not only lazy computer metaphor (files), but actually employed widely in literature and does important explanatory work (Heinen 2024; Pacheco-Estefan 2024)
- ❖ Cox (2014: 189) notes that decoding research on the visual system [...] implicitly recognizes that the problem of vision **is not one of information content, but of format**.
 - Hunt for ‘real’ information – **mental currency**
 - **Decodability**
 - **Formats help explain (explanatory purchase)**

Historically speaking

- ❖ Existing literature on formats – mostly philosophical but dated and scant
- ❖ In CompSci: routine and understandable + intuitive. Robust in CogSci, but connection with conceptual work in philosophy is wobbly. **This needs to be ameliorated.**
- ❖ Philosophical debates
 - Based on intuitions about public representations (maps – paradigm) to judge against
 - Main ambition – to demarcate distinctions and look for counterexamples
 - Pictorial vs linguistic Imagery Debate (Tye 1991); also: digital/analog, symbolic/distributed, continuous/discrete. What's common – not many types of formats for these philosophers!
 - Specialized formats for modalities (Raftopoulos 2020); seems explanatorily vacuous
- ❖ Existing theories of representational formats can be grouped into **three broad categories**, depending on what conceptual component of the notion of representation they take to be **central to individuating formats: contents, vehicles, or the function from the former to the latter** (Lee 2022)
- ❖ Brain information cannot plausibly fully rely on our common sense ideas about icons, signs and cartography

Coelho Mollo's & Vernazzani (2023) - computational profiles

- ❖ Thesis 0: Let's get rid of all the existing misleading talk about maps and whatnot it's not helping. We should start from scratch to achieve something realistic that accounts for empirical endeavours in the attempts at explaining brains
- ❖ Formats are not icons, maps, sentences, but something more abstract – computational profiles.
- ❖ **T1: Representational formats are the computational profiles of representational (sub)systems in cognitive systems, be they biological or artificial.**
- ❖ **T2: Computational profiles, in turn, are individuated by the inner and outer constraints of computational vehicles—that is, the values they can take and their mutual constraints.**

So, like... action figures?!

The parts that can be moved are what we may call, with quite some stretch, the “vehicles.” The number of relative positions the moving parts can occupy and the relations between variations over them (e.g., one leg also making the other leg move) are their **inner and outer constraints, respectively**.



We can categorize different types of action figures in terms of their moving parts, the values that those moving parts can take (which positions they can occupy relative to the body and to each other), and the relations between variations over those parts (e.g., whether moving a leg also leads to moving the other leg or, rather, an arm), and we can do the same with computational vehicles. **How many parts of the vehicle can be computationally wiggled, and what values can they take? How does wiggling one part affect the possible values of another part?** And how does wiggling the values of a vehicle affect (or not) other computational vehicles—that is, does changing the values taken by one vehicle affect the values of the others? (2023:10–11)

Explanatory desiderata

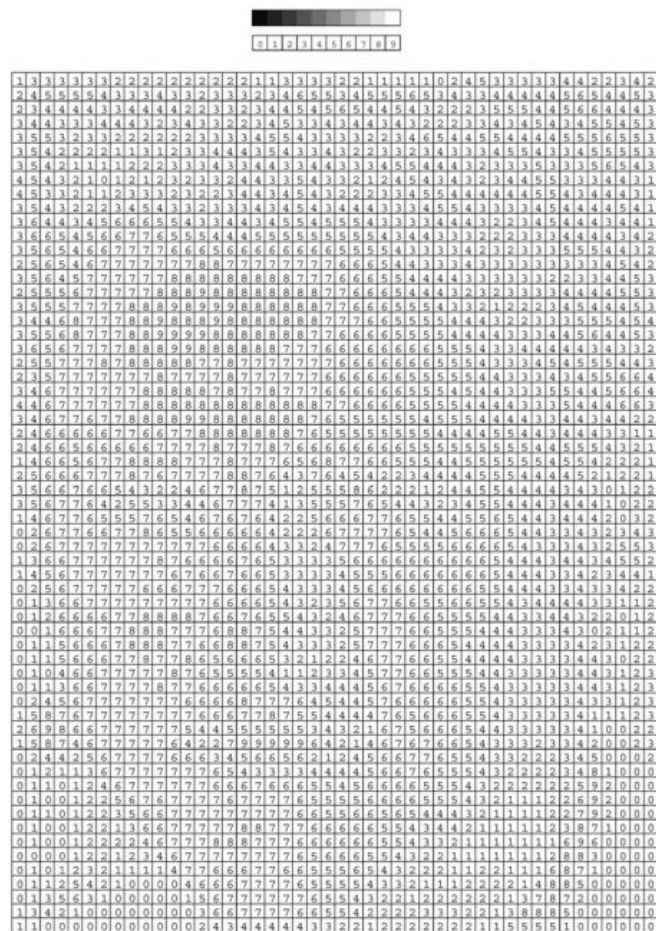
- ❖ **Transformation-based explanations:** help explain the workings and behavioral effects of cognitive states and processes in terms of the specific kinds of transformation or manipulation available and performed over such states and processes
- ❖ **Efficiency-based explanations:** help explain why certain cognitive states and processes are more (or less) adequate for a specific task in terms of certain sets of transformations/manipulations being more efficient, powerful, less taxing, and/or temporally advantageous
- ❖ In addition, a theory of representational formats should be epistemically fruitful (**epistemic fruitfulness**) [sort of normative philosophy of science req]

Can they really account for ‘epistemic fruitfulness’ desideratum?

- ❖ We contend that a purely computational approach to formats can fulfill the explanatory roles (...)—transformation-based explanation, efficiency-based explanation, and *epistemic fruitfulness* (2023:16; emphasis mine)
- ❖ ‘In light of the ambiguity of many of the appeals to representational formats in the literature, a theory of representational formats **should identify a clear, motivated domain of questions that can or should be tackled** by such an appeal.’
- ❖ I think it’s the ‘same information–different format’ idea.
- ❖ Their account struggles to account for Yousif and President examples.
- ❖ But this has to be reconceptualized as ‘formatted for the receiver to be exploitable (to guide behaviour or at least pass further to another node)

The invisible president - format's important

(Hyvaerinen et al., 2009:4-5)



***Legibility* -> utilizability**

- ❖ More fundamental than *usefulness* of representation – looser condition but more general: renders the information accessible
- ❖ Condition states the correct formatting licences *utilizability*
- ❖ Information could be present in the data (e.g. neural signal), but not useable (in principle or practice)
- ❖ *Getting information from a table is like extracting sunlight from a cucumber.* (Farquhar & Farquhar 1891:55)

Tic-tac-toe to 15 (Yousif 2022)

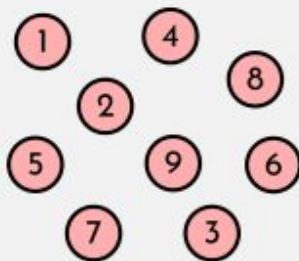
A

Tic-tac-toe

X		
	X	O
		O

Fifteen game

Nine disks marked with the digits 1 through 9 are placed faceup on a table. Two players take turns picking one disk at a time. The winner is the first player to obtain the sum of exactly 15 from among any three of her disks.



The same game!

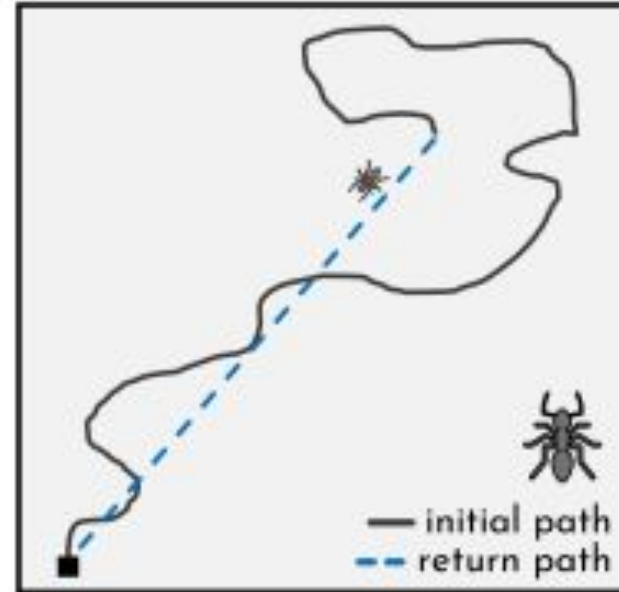
8	1	6
3	5	7
4	9	2

Both games have the same initial position, the same allowable moves, and the same win conditions, but the **format** of the game is very different. The game of tic-tac-toe traffics in x's and o's, whereas the fifteen game traffics in numbers. When the two are superimposed, you can see they are functionally equivalent.

How do desert ants (*cataglyphis fortis*) navigate? (Wittlinger et al. 2006)



C



- ❖ Apparently + built-in pedometer + sense of direction
- ❖ 2-place vector effectively implements polar system **to guide behaviour**
- ❖ Way to investigate - look for patterns in errors
 - If errors in both variables (dimensions) correlated, then unlikely as real values stored (real **formatting**)
 - Manipulations: ants on stills
- ❖ Holy grail - **direct access into 'mental' format**
 - Gold-standard for testing and explaining, but limited for humans of course (pluralistic beasts). Evidence for co-existing nav. systems.

Exploitability, but also redundancy

- ❖ We mentioned utilizable contents. In neuro-jargon it is said that such information can be **‘exploited downstream’**
 - *Receiver* of any given information is crucial
 - *Consumer* in teleosemantic terms (Milikan, Neander; Cao 2019)
 - Important for eavesdroppers (neuroscientists). They are wildly different consumers/receivers than next neuron (or functional region) down the wire!
- ❖ Note, importantly: that existence of many formats and even their non-compatibility does not render them mutually exclusive or hostile (as explanations)
- ❖ Redundancy is almost certain and there are good pragmatic reasons for the brain to maintain pluralism (or even a surplus)

Neural decoding (*decoder's dictum*; Ritchie et al. 2019)

- ❖ In, multivariate pattern analysis (MVPA), or ‘neural decoding’, has transformed the field of cognitive neuroscience. Underlying its influence is a crucial inference, which we call the **decoder's dictum: if information can be decoded from patterns of neural activity, then this provides strong** evidence about what information those patterns represent (2019:1)
- ❖ In summary, one gets to the decoder's dictum by endorsing several claims:
 - (i) that MVPA reveals information latent in neural activity;
 - (ii) that an underlying neural population code implies decodability;
 - (iii) an informational view of neural representations and their contents; and
 - (iv) the hypothesis that biologically plausible linear classifiers are sufficiently similar in architecture to the decoding procedures employed by the brain.

The latter is what lets us infer that decodable information is appropriately formatted for use by the brain, even when we do not necessarily know what that format is. So,
 - (v) if we can decode information from patterns of activity using MVPA, this provides good evidence in favour of the hypothesis that the patterns represent the information

Decodability? Not so fast - formatting is in the way

- ❖ Williams et al. ([2007]) provided one of the earliest indications **that not all decodable information is ‘read-out’ in behaviour.** (retinotopy vs LOC)
- ❖ Cox (2014): We know that the **activity of retinal ganglion cells contains all of the information that the visual system can act ; firing of retinal ganglion cells is not in a format that can be easily read-out by a downstream neuron in order to guide action.**
- ❖ As DiCarlo and Cox ([2007], p. 335) put it, **‘the problem is typically not a lack of information or noisy information, but that the information is badly formatted’.** But even then: ‘badly formatted’ information might correlate with behaviour.
- ❖ Note that the problem **is not one of spurious correlation.** In an important sense, it is quite the **opposite problem. There is plenty of information, even in V₁,** which a clever decoding algorithm can often pick up on. More generally, a brain region might carry information that is **reliably correlated with the information that is actually used, but which is not itself used in behaviour.** This is because the information in a region might need to be transformed into a more appropriate format before it is read out.

Kulvicki (2007); Larkin & Simon (1987)

- ❖ *Immediate availability* cashed out in terms of the need for inferential mediation to provide ‘extractable form’ of the informational content
- ❖ Originally to defend the difference between interacting with numerals and images
- ❖ Kulvicki: ‘semantic and syntactic salience’ as necessary properties
- ❖ L&S: ***Informational equivalence*** of representations and their **computational equivalence (distinct properties)**. Representations can be alike in the information that they carry but differ in how they make that **information available to their consumers.**

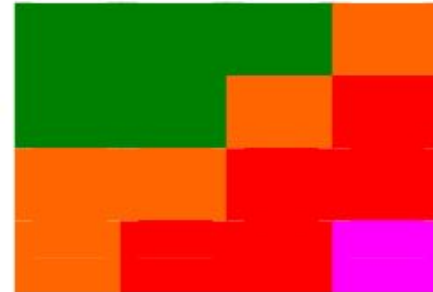
Structure for immediacy and extractability

(x-coordinate (1-4), y-coordinate (1-4), temperature (0-6))

1, 1, 0	3, 1, 3
1, 2, 2	3, 2, 3
1, 3, 2	3, 3, 4
1, 4, 3	3, 4, 5
2, 1, 2	4, 1, 3
2, 2, 2	4, 2, 4
2, 3, 3	4, 3, 4
2, 4, 4	4, 4, 6

0	2	2	3
2	2	3	4
3	3	4	5
3	4	4	6

0	2	2	3
2	2	3	4
3	3	4	5
3	4	4	6



Conclusions

- 1) Transition-based explanations are required and they carve up a niche for formats. CM&V are also right we should discard a lot of older distinctions about and the mode of judging only against intuitions.
- 2) However, types of formats do govern what types of operations are in the realm of possibilities for an agent (or a part of). They are required to guide behaviour
- 3) Kulvicki and Simon are also right in terms of structure providing action-guidance.
- 4) Receiver-relativity (Rathkopf) is a thing (teleosemantics)

But even more general problem lingers: who is the interpreter? Encodingism (Bickhard & Terveen)

Warszawa. Pomnik Kopernika
z widokiem na b. pałac Staszica.

Thank you for coming and for your attention.

Any questions welcome! If any of the references
grabbed your interest, please do ask or pop an
e-mail, I'll be glad to provide the bibliography!

wojciech.mamak@gmail.com