Comments on
"The Emulating Interview... with Rick Grush"

Aaron Sloman
School of Computer Science, University of Birmingham, UK
http://www.cs.bham.ac.uk/~axs

(To appear in:
Avant. The Journal of the Philosophical-Interdisciplinary Vanguard
http://www.avant.umk.pl/en/
An Open Access Online Polish-English Journal)

May 3, 2011

This is a response to some parts of the interview by Przemyslaw Nowakowski published here:
http://www.avant.umk.pl/en/2010/11/the-emulating-interview...-with-rick-grush/

My views are very close to those expressed by Rick Grush, but I think some of the things he says are misleading. I don't know whether that is because he simply has not expressed himself clearly, or because what he meant to say really is different from what I think he should have said. I hope these comments will turn out to be clarifications rather than criticisms. I have italicised quoted extracts from the interview.
The sources are indicated as follows: [Q] Interviewer, [RG] Rick Grush, [AS] Aaron Sloman.

[RG]*My conception of emulation is fairly simple: emulation is representing something by using a model to stand in for it.*

[AS]
The common notion of "standing in for" is much too specific to cover all the cases I think Grush is trying to cover. Typically standing in involves taking the place of in some role. E.g. someone standing in for a teacher needs to be able to teach. That is usually not what mental information structures do in most cases. E.g. you may have information about a banana, where it is, how ripe it is etc. and that can lead you to choose and carry out a variety of possible actions, including going to the banana, picking it up, throwing it in a bin, eating it, etc. But you can't pick up, throw, or eat your information about the banana nor the bearer of the information (physical brain mechanisms, or virtual machine symbols implemented in brain mechanisms).

Use of information content requires use of an information bearer: something that, for purposes of reasoning, planning, or interpreting perceptual data, can be constructed, analysed, manipulated (not necessarily physically), decomposed, combined with other things, rearranged, and inspected, for various purposes, including both *semantic* purposes, such as: drawing conclusions about what is the case, or would be the case in some situation, and *syntactic* purposes, e.g. working out which further manipulations, combinations, decompositions, mappings, are possible.

Uses of information structures are normally so different from uses of the things they are information about, that it is normally seriously misleading to regard any sort of representation as standing in for what it represents. Manipulating the materials referred to in a recipe for making a cake can produce a cake, whereas no amount of manipulating the recipe will do that. Grush may think these are minor exceptions to his claim. Later I'll show that even the central cases he describes don't involve standing in for.

[RG]
This happens all the time: we use flight simulators as models for airplanes, we use chess boards to try out moves before we commit to making our official move. What these cases have in common is that some active agent is interacting with one thing, a model or emulator, in the same way that it would interact with something else, the represented entity. You interact with a flight simulator the same way you would interact with a real aircraft, you interact with the unofficial chess board (the one you use for trying out moves) in the same way you would interact with the official chess board.

[AS]
It is important to notice that in some cases the simulation leaves nothing essential out, e.g. when you play a game of chess with a chess simulator, it can actually function as an opponent.

However, it is impossible to use a flight simulator for an airbus as an airbus. E.g. you cannot get into it, fly it, then get out on another continent. The flight simulator works for teaching purposes because it is being used for teaching, not for flying. The teaching uses are limited because the simulator cannot replicate all the physical experiences a pilot might have in extreme emergencies, for instance, with part of the cockpit on fire and dangerously full of poisonous fumes, etc. More importantly, when you think about what you can do or should do in case of a fire or some other emergency you normally need to work out reasons, not merely try things to see what happens. I'll return to the differences below.

Using a chess program as an opponent is certainly possible, though (at present) it will not be an adequate simulation for all purposes. When playing a chess program you cannot try to distract it by means of your facial expressions, though you can with many human chess players.

Similarly, when thinking about possible moves you may use an internal representation of an opponent but that representation will be different from a real opponent in many ways, including our inability to fool or distract the internal representation.

More importantly, there is a major difference between actually playing a game against a chess program and using the chess program as a sort of "scratch-pad" while trying to select a move against another opponent. This is typical of the difference between using an internal model or representation to reason with and acting on the external entity that is represented by the model.

A chess program with more than the mechanisms required for playing chess may allow you to investigate possible moves and their consequences without the sorts of mistake that could occur if you did it using chess pieces and a real board (which would allow you to make illegal moves).

Likewise the use of the chess program as a "scratch pad" for exploring moves could allow you to avoid mistakes that would occur if you tried to do all the move-planning in your head. Unlike your internal mechanisms, the program may remember all the options you have tried and allow you to re-inspect them to assess their advantages and disadvantages. All of that functionality, provided when reasoning about chess. is different from the functionality of an opponent who is trying to win against you.

In short, there are different sorts of emulation, and not all justify claims about "standing in" for what is being simulated.

[RG]
The basic idea is that this phenomenon applies also to the brain itself – it constructs models of the body or the environment that it can then interact with in order to represent the body or environment.

[AS]
But, as I tried to indicate above, interactions with the models when reasoning, or planning, are normally very different from interactions with the things they are models of. For example, the model may be chunked into discrete components with discrete states whereas the reality is not. That chunking may be essential for exploring multi-step plans, a process whose computational complexity is tractable in discrete spaces, but explodes in continuous spaces.

Grush appears to be suggesting here that using a stand-in is a characterisation of all uses of information about the environment, for instance cases of trying to predict what will happen, or trying to plan a sequence of actions, or trying to find an adequate explanation of some observed state of affairs or event.

That extreme claim (if he is making it) is false, and false in very important ways. That's because processes like planning, predicting, and explaining essentially need to use not *things* but *information about the things*, and most things in the environment are very different sorts of entity from items of information. E.g. you can sit on a chair, but not on information about a chair. You can use information about a cake to bake a cake, but you cannot use the end product in the same way: it does not provide information about what to do to produce it (except for someone who already has information about many types of cake-making and can select the appropriate recipe after inspecting the cake. This is a case where the cake provides information about which recipe to use, which is different from the information in the recipe itself – the recipe does not tell you which recipe to use (though it may refer to sub-recipes).

[RG]
There are many complications beyond this simple idea, of course, such as what things are modeled, how models are built and what it means to use something .in the same way. as something else.

[AS]
Yes: those complications and the variations are important. In particular some of them involve not using the model (or more generally an information structure) "in the same way" as whatever is being modelled, represented or referred to.

In his next comment Grush makes some of the same points I have been making, but does not draw the same conclusions.

[RG]
As for the second part of the question, yes, I think there is an application for conception and inference. Obviously one common use of a model is precisely to make inferences – the reason I try out possible moves on a "model" board is it to draw inferences about what moves might be good or bad if I were to try them out for real. Of course the model by itself doesn't make the inference, I have to know how to use the model correctly, and I have to interpret the results correctly – I can't take the result of a possible move I try to be a memory of some past state of the chess board, or a perception of its current state, or a guess as to the state of a different board. I have to know, in some sense, that in this situation the state of the model is representing a hypothetical state of affairs, and this is not something that the model itself makes apparent.

[AS]
This answer begins to spell out the reasons why the use of internal information is in many cases totally unlike the use of whatever the information is about. I'll say more about this below.

[RG]
As for conception, I don't know. Addressing this topic would take a long time. Let me just say that I believe that articulated emulators can be usefully taken to be conceptual, in the sense that the articulants have many features of concepts. But I won't go into this any further. The issue of what concepts are is a tricky one, and to be honest, one that I am not that interested in getting mired in.

[AS]
Without arguing about definitions (e.g. of "concept"), we can see, as hinted above, that the alleged similarity between internal emulators used for reasoning and the things they emulate is illusory. A useful internal model of some external structure, or process, or mechanism, will differ in many important ways from the mechanism itself. In particular the user of the model needs to be able to do things with it, that can't usually be done with the with mechanism in its normal use. However, physical mechanisms can sometimes be used to reason with, and that is very different from using them as physical mechanisms, as I'll try to show.

For example, to answer a question about generic features of the mechanism (e.g. about the relationship between directions of rotation of a pair of meshed gear wheels with fixed axes) you can't answer just by modelling the process of rotating one of wheels, and inspecting the process.

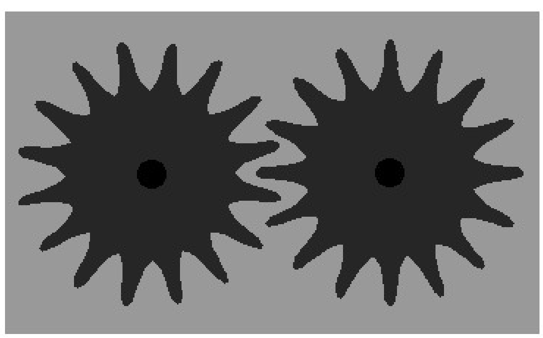
You also need to make use (consciously or unconsciously) of the fact that the wheels are made of materials which are impenetrable and rigid, and you need to think about how the rotation of one wheel causes surfaces of its cogs to interact with surfaces of the cogs on the other wheel, and how the cogs in contact keep changing as a result of the rotation, though without any time interval when there is no contact (if the gear wheels are well made).

Only with all these details made explicit can you reason that if one wheel rotates the other *must* rotate in the opposite direction.

A toddler playing with meshed wheels may discover empirically that the rotations are always in opposite directions, without knowing why, and without understanding that it is *impossible* for them to rotate in the same direction.

The child's empirical investigation uses the gears as a physical machine, whereas the older person investigating the gears in order to explain why the wheels must rotate in opposite directions (under certain conditions) is using the physical machine as a representation encoding important information about the functioning of the machine. Those empirical and explanatory uses are very different – though justifying that claim would involve demonstrating how both uses could occur in a functioning machine, and how the two machines would differ.

Insofar as the knowledge required is empirical, the manipulation of the physical machine has the advantage: in general, experiments on imagined machines cannot be relied on to give the same results as experiments on physical machines. In contrast, when trying to understand *why* something must happen and *under what conditions*, one is not merely checking the truth of some generalisation. Rather, one is constructing an explanation, which, in part is like a proof, which makes use of a thought experiment as a basis for inferences.



It is also possible to use the physical gears to do a thought experiment! This involves considering the details of what happens when one wheel is rotated about a fixed axis, and how it applies a force to the other wheel (through the cogs that are currently meshed) and how the rigidity and impenetrability constrain the effects of motion of a cog, and also how the geometry ensures that there is always at least one cog of the wheel being turned that is in contact with a cog on the other wheel (which would not be true of all designs, e.g if the cogs varied in length). Using all that information to reason about a perceived process is very different from simply learning empirically what happens when a wheel is turned, or simulating the process mentally and observing the results.

So we can turn the argument about emulation on its head. Someone trying to *reason* about the consequences of a process or event in a piece of machinery can think about the implications by using a representation of all the relevant relationships and constraints, and propagating the changes through various parts of the representation to deduce consequences. However, in some cases the process is too complex to do mentally, and it can be useful instead to use an external diagram, or even the piece of physical apparatus itself, e.g. an assembly of gear wheels. Then you can use the perceived structure of the apparatus as an external memory store while you reason about the flow of causation from part to part.

This use of physical apparatus to reason about its capabilities is quite different from testing its behaviour empirically, e.g. by moving a part. It is more like the way a proof in Euclidean geometry can be constructed by modifying parts of a diagram, as explained in (Sloman, 1971).

When an external version of an internal model is used for making inferences, it does not matter if the physical parts actually have the properties (such as rigidity, or straightness) that would be required for the original mechanism to work as intended. All that is required is that it can be used to help with the reasoning about how the effects of a change must be propagated if the parts do have the required properties.

For the same reason, a mathematician using a diagram to prove a theorem in Euclidean geometry does not require the lines in the diagram to be infinitely thin or infinitely straight, as required in the statement of the theorem. That's because we can reason about consequences of such properties in their absence, because we have knowledge of those properties represented in a usable form.

All of this implies that both physical machines and internal models can sometimes be used for similar purposes. But each can be used for two very different sorts of purpose: namely (a) discovering empirically or recalling what the effects of some change would be, (b) reasoning from facts about the machines to a conclusion about what the effects *must* be. Many animals can do (a) but not (b) (as far as we can tell). Piaget (1983) reports experiments revealing the highly erratic development of such competences in humans during the first 12 years. The capabilities require architectural changes over and above the ability to learn empirically. The theory in (Karmiloff-Smith, 1992) makes some significant steps towards explaining such developments. I believe all this will eventually vindicate some of Kant's ideas about the nature of mathematical knowledge (Kant, 1781)[1].

Most robot designers ignore the differences between (a) and (b) and merely attempt to make robots that can do (a). So their robots cannot explain why they did not do something or reason

[1]Some speculations about the mechanisms and how they evolve and develop are presented in (Sloman, 2008b, 2008a, 2009b, 2009a, 2010), and other presentations and papers on the same web site.

about what would have happened if they had done it, and many of them have no ability to think about what they could do and what the consequences would be if they were to do it: they merely react immediately to each new situation as they have been trained to react. Machines that select actions by planning in advance are different.

[Q]
You seem to criticize enactivism quite often. How would you, then, place your conception of emulation among the ideas of enaction, embodiment or situated cognition?

[RG]
What these views and mine have in common is in their departure from certain ways of thinking about cognition and representation. We don't view cognition and representation as primarily a matter of logic, or language like strings of symbols, and we take cognition to be connected to and based upon embodied motor engagement. The main difference is that these other views are often anti-representational in nature, and they also often claim that cognition is really not in the head. My own position is that given an appropriate understanding of what representations are, (not sentences, but, more like models), we can make perfect sense of the idea that the brain represents. Furthermore, this notion of representation is as connected if not even more so to motor behavior than the other views. It also shows how representation and cognition can all take place entirely in the brain. I.m not saying it always does: we can and often do use external models. My point is weaker. Namely, that representation and cognition are often internal to the brain.

[AS]
And my point is that whether representation and cognition occur fully internally or make use of external representations to reason with, what is going on is very different from what goes on when an animal or robot merely acts on the environment or when a computer game engine simulates some portion of a physical world to produce consequences of a simulated action. Many animals, and apparently newborn humans cannot reason about possibilities and necessary connections: they appear to be able merely to sense and react.

Some other animals, including corvids, primates, elephants, and squirrels defeating allegedly squirrel-proof bird feeders, as well as older human infants and toddlers make the transition from merely being able to act, to being able to *reason about* possibilities and constraints, either internally or externally, possibly using external aids. But then even if that involves motor behaviour it involves much more than motor behaviour. Compare the processes of Representational Redescription hypothesised in (Karmiloff-Smith, 1992).

The trouble with many who theorise about embodied cognition is that they ignore the variety of types of functions for which physical action can be used, and variety of types of reflective cognition that can precede, accompany or succeed physical action. Their theories may be a good approximation to many sorts of insect cognition and microbe cognition.

I think Immanuel Kant's philosophy of mathematics depends on these distinctions, though he lacked some of the conceptual tools required for thinking about explanatory mechanisms. He would have recognized Artificial Intelligence as a profound new way to do philosophy.

Note:
Because many of these ideas will be unfamiliar to most philosophical readers, I have included several references to my recent attempts to explain them in more detail.

References

Kant, I. (1781). *Critique of pure reason*. London: Macmillan. (Translated (1929) by Norman Kemp Smith)

Karmiloff-Smith, A. (1992). *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge, MA: MIT Press.

Piaget, J., & *others*. (1983). *Possibility and Necessity Vol 2. The role of necessity in cognitive development*. Minneapolis: U. of Minnesota Press. (Tr. by Helga Feider from French in 1987)

Sloman, A. (1971). Interactions between philosophy and AI: The role of intuition and non-logical reasoning in intelligence. In *Proc 2nd ijcai* (pp. 209–226). London: William Kaufmann. (http://www.cs.bham.ac.uk/research/cogaff/04.html#200407)

Sloman, A. (2008a). *Evolution of minds and languages. What evolved first and develops first in children: Languages for communicating, or languages for thinking (Generalised Languages: GLs)?* (Research Note No. COSY-PR-0702). Birmingham, UK. (http://www.cs.bham.ac.uk/research/projects/cosy/papers/#pr0702)

Sloman, A. (2008b). *A New Approach to Philosophy of Mathematics: Design a young explorer, able to discover "toddler theorems"*. Available from `Onlinepresentationhttp://www.cs.bham.ac.uk/research/projects/cogaff/talks/#toddler`

Sloman, A. (2009a). From "Baby Stuff" to the World of Adult Science: Developmental AI from a Kantian viewpoint. In F. McNeill (Ed.), *Proceedings Workshop on Matching and Meaning, AISB 2009 Convention* (pp. 10–16). SSAISB. Available from `http://www.aisb.org.uk/convention/aisb09/Proceedings/MATCHING/FILES/Proceedings.pdf`

Sloman, A. (2009b). Some Requirements for Human-like Robots: Why the recent over-emphasis on embodiment has held up progress. In B. Sendhoff, E. Koerner, O. Sporns, H. Ritter, & K. Doya (Eds.), *Creating Brain-like Intelligence* (pp. 248–277). Berlin: Springer-Verlag. Available from `http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0804`

Sloman, A. (2010). If Learning Maths Requires a Teacher, Where did the First Teachers Come From? In A. Pease, M. Guhe, & A. Smaill (Eds.), *Proc. Int. Symp. on Mathematical Practice and Cognition, AISB 2010 Convention* (pp. 30–39). De Montfort University, Leicester. Available from `http://www.cs.bham.ac.uk/research/projects/cogaff/10.html#1001`

Additional tutorial presentations relevant to these topics are in my talks directory:
`http://www.cs.bham.ac.uk/research/projects/cogaff/talks/`