From: Jeffrey Epstein <jeevacation@gmail.com>
To: roger schank <██████████████████>
Subject: They agreed with your ai winter
Date: Wed, 13 Feb 2013 06:44:30 +0000

---------- Forwarded message ----------
From: **S.M. Kosslyn**
Date: Tuesday, February 12, 2013
Subject: Today's discussion
To: Joscha Bach <████████████████████████>
Cc: Jeffrey Epstein <jeevacation@gmail.com>

Hi J(s)…

Please see >>>> below for one more iteration…

_/

Stephen M. Kosslyn
Director, Center for Advanced Study in the Behavioral Sciences
Stanford University
75 Alta Road
Stanford, CA 94305
Voice: ████████████

On 12 Feb 2013, at 5:45 PM, Joscha Bach wrote:

>
>>>>>> I've generated various predictions (some of which I later found out have already been tested in other contexts, such as the "pratfall effect").
>
> I am curious: how does the pratfall effect relate to SPS?
>
>>>>High-status people are liked more if they have a visible flaw (e.g., they trip and spill coffee when entering the room). In my view, this implies that they are not "perfect" and hence they can profit from being helped. And this in turn implies that they might be willing to incur a debt by letting you be their SP, leading them perhaps to serve as one's own SP in the future.

>>>>>> Yes. I spent a lot of time with Kevin when I lived in Paris. He's got an interesting perspective.
>
> Also, he seems to be very kind. I like his ideas, but I have no idea how he intends to turn them into a research program.
>
>>>>I don't much like most of his ideas; they are quite "behaviorist" in flavor, and anti-mechanistic… But his intuitions are often interesting, and make me think

>>
>>>) A whole bunch of people have taken off on the extremist tangent of 'enactivism', lured by the creation of

meaning through environmental affordances, and want to construe the mind not as what the brain does, but as an phenomenon that emerges over the interaction of body and world.
>>
>>>>>> If we really understand human behavior we should be able to build something that behaves the same way as humans. In my view, that's the "gold standard" metric for genuine understanding. I don't find the above approach helps toward reaching this goal.
>
> I guess that Enactivism is not meant to be a research agenda, but an anti AI position that mainly serves a spiritual need. Rolf Pfeifer and Thomas Christaller, for instance, participated in the failure of expert systems and then had spiritual awakenings; for them, the expert system problems has shown that minds cannot be explained by computation, and must therefore be the result of non-computational properties conveyed by the touch of the mystical substance of physical reality.

>>>>That strikes me as a non-useful response

> Both have turned towards robotics, Rolf Pfeifer advocates passive control (i.e. robots without representations), and Thomas shaped the German research landscape towards robots controlled by dynamical systems (i.e. unstructured ANNs). I think that both have been very influential into turning many AI departments into 'cognitive robotics', whereby 'cognitive' is a special German synonym for 'cool'. We have 'cognitive' systems with less than a handful of neurons.

>>>>They could model effects of extreme advanced age…
>
> I still think that minds, not bodies, are primary.

>>>>Both are important, but the mind seems like the rate-limiting factor for what the body can do, not vice versa...

> At some point in the not too far future, the returns from working with robots might exceed the costs of getting them to work, but for the time being, robots do not have interesting affordances, social embedding and all those other neat things the Nouvelle AI guys clamored for. Robots mainly locomote and push against things (and if you are working for the military, they might also blow a few things up). But if we want to do intellectually interesting stuff, even at toddler level, virtual environments may be better suited to compare human and AI performance.

>>>>Agreed
>
> By the way, have you looked at Minecraft?

>>>>Nope. Never even heard of it. JEE?

> It provides an open ended virtual Lego-like world from a first person perspective, is complex enough to build working computers in it, allows for continuous movement and interaction, but can also be represented and acted upon with discretized models. Objects in MineCraft have proper semantics, for instance, trees are composed of wooden stems that gradually grow from the ground and end in foliage (every tree looks different). Trees may be burned or chopped down, and the individual wooden blocks of the stem can be crafted into tools. Once chopped down, trees cannot be rebuilt (unlike walls or mountainsides, for instance). That is unlike On the other hand, every object in Minecraft is either atomic or compositional (Fodor might like that). There is not much fuzzyness and vagueness, except in the changing shades of colors and the features of terrain and vegetation. For an AI living in that world, low level perception could cheat when it comes to edge perception, volumetric geometry modeling etc., but classifiers for trees, water bodies, hills, regions, buildings, tunnels etc. would need to handle fuzzy and vague categories.
> In my view, that could turn MineCraft into an intermediate paradigm for comparative problem solving tasks.
>
>>>>I still like the idea of building a mind that can take IQ tests (the WAIS, for example…). This does have an

effector piece, requiring fitting pieces together into objects, arranging cards into narrative arcs, etc.

> On the other hand, most obvious tasks in the game would be about construction, resource finding etc. Since it is a multi agent environment, social interaction is possible, but currently there is nothing that would make this strictly necessary.
>
> One of the things that make be curious is how we could explain personality properties with variances in the fundamental variables of the motivational system. For instance, there are problem solving tasks where performance is not related IQ, but where we can find correlation between strategy choices and conscientiousness, agreeableness, neuroticism. I wonder how to construct a set of tasks that
> - could be taken both by humans and computational agents
> - allows to predict personality properties for the humans
> - can evaluate a theory of emergence of personality traits by equipping the agent with a motivational model that will put the agent strategy choices into the same categories as the human strategies.

>>>>I think this may be biting off a bit too much for one task. Why not put the personality piece aside for the moment, and focus just on smarts?
>
> Explorative games are good candidates; I am still looking for an idea to tackle the traits that are related to social behavior.
>
>>>>If you want to go this way, how about World of Warcraft or something more social like that?

>>> But despite a lot of people writing about extended (and even de-brained, anti-representational) minds, I think that your idea of a socially extended mind is a very underexposed area, and will probably create a lot of resonance.
>>
>>>>>> I have seen no evidence of that; I think the idea is too weird for many people. (This is a problem I've had all along, even when I was just studying mental imagery.)
>
> I think that you should present your idea as a programmatic argument that shows why computers cannot be intelligent, instead of an actual research position. Some philosophers seem to love this, and you will invited to share panels with Searle and Penrose.

>>>>Ah, but I were wiser and more cynical -- I'm sure you are right, and i would be more popular. The problem is that I don't see any reason why computers can't evolve into SPs.. I don't believe that computers "cannot be intelligent" (I know you know this..).
>
> But seriously: Castelfranchi and a bunch of people concerned with multi agent modeling of sociality should be very interested.

>>>>I'm touched by your optimistic perspective (really)
>
> BTW: what was the opposition wrt mental imagery? Did people really follow Fodor's reasoning?

>>>>Some did

> And since everybody is talking about symbol grounding and Barsalou's perceptual symbol systems now: is the fight not won?

>>>>The younger generations seem more sympathetic, that's true

> Or do I underestimate the powers of the dark side?

>>>>Most of the old guys will die rather than change their views
>
> What happened to Jerry Feldman's group at Berkeley when he retired?

>>>>He's still active. I spent some time with him a couple of months ago.

> I found his sweeping and eclecticist approach to language very refreshing.

>>>>Not quite deep enough to do much with, I fear...

> I felt that he should have left the Hippocampus alone and make more progress on the functional modeling ('X-schemas'),

>>>>Yes!

> but that is probably just showing my personal preference.

>>>>No, I think you are right -- you can't just dabble in neurophysiology
>
>>> It seems that our social cognitive development has a few biological windows that cannot easily be compensated for when missed.
>>
>>>>>> Yes.. "sensitive periods" (not "critical periods")
>
> Yes. It seems easier to sell a straightforward, dramatic story ('use it or loose it"), because it creates a sense of urgency and thus relevance in the audience. 'Use it, or work much harder later on' simply does not have the same clout.

>>>>I think you are right
>
>>
>>> You even allude to these phases in the section immediately before it. You also remark that you bet that language takes an important role in shaping our brain, but had not seen this idea tested. Have you met Luc Steels?
>>
>>>>>> No, I have not even heard of him…
>
> See, for instance Steels & Belpaeme: Coordinating perceptually grounded categories through language: A case study for colour, in Behavior and Brain Sciences 2006

>>>>Ah
>
>>>>>> News to me. I recall a series of studies by Eleanor Heider and colleagues with the Dani (a tribe in New Guinea) where they showed that the perceptual boundaries of colors were the same, even when the language cut up the categories in different ways.
>
> As far as I know, the debate seems to be still going back and forth, with a tendency towards universalism for similarity and (mostly) saliency, but relativism for perceptual boundaries and categorization.

>>>>Who are the main players in this debate?
>
>>> By the way: I find the effect of language on mental processing very fascinating. Have you ever thought about deliberately creating another language that improves cognitive performance?
>>

>>>>>>> No… but that is a very cool idea!!
>
> Imagine adding more fine-grained spatial relations, a better decomposition of causal, taxonomic and partonomic modes, syntactic shortcuts for cross-references in the syntax tree that let our pronouns blush, tail-recursion and a built-in library of mnemonics, for a start!

>>>>>Very, very interesting!
>
>>>>>>> I'm wondering whether computers could allow expression of a language in ways not possible organically, such as by the conveyance of images..
>
> They already do! It is called 'movies', but the process of creating them is so damn tedious.

>>>>>Hmm.. I'm not sure that's quite when I mean, but perhaps close...
>
> So what we would want to have is a system that describes a visual and auditory narrative set in an arbitrary complex virtual world. This world/narrative is not constructed from an empty page, like a book, but from a vast set of dynamic defaults, like the mental representations that emerge upon reading a book. The act of creating the narrative would amount to giving constraints and pointers for disambiguation into the system.
>
>>>>>Yes, that's more along the lines of what I had in mind

> At the moment, we would probably want to use natural language for that. Roger Zelazny once wrote a short story ("He who shapes") about a psycho therapist who creates virtual worlds for his clients with a more holistic interface (lots of dials and levers), but unfortunately, he did not give us the details of that interface.

>>>>>I vaguely recall that story
>
> I think what the problem of language boils down to is this: mental representations are (functionally speaking) directed, hierarchical, dynamic hyper-graphs with variable link weights. These graphs can be instantiated as activation patterns, on the canvas of our cortical circuitry. For conveying mental representations, we depend on discrete strings of symbols. The set of natural language families is probably almost co-extensional with the set of possible elegant solutions for the translation problem from discrete strings to mental representations, and vice versa, given cognitive resources with very particular limits ("stack size", "number of universal pointers", "number of special purpose pointers", timespans for setting, fading and resetting short term memories, etc.).

>>>>>The above sounds reasonable to me, but I'm no expert on language. Have you met Steve Pinker? I wonder what he would say to this…
>
> I would so love to work on this, but I think it is a major effort that will span more than a generation to get right (which of course means that we should get started).

>>>>>I think that's probably right

> Instead at the geographical regularities and peculiarities of language, we should be looking for its plate tectonics; the heart of this are universal mental representations.

>>>>>That's my sense, too
>
> ---
>
>>> Speaking of which: you mentioned the relationship to your wife in the SPS context, too, and it got me to reflect on my own, quite symbiotic relationship. I often suffer from the illusion
>>

\>\>\>\>\>\>\> Why do you think this is an illusion?

\>

\> Our mutual empathy concerns values, and the adoption of many value-related goals, but not skills/strategies, personality properties and often not beliefs. My intuition tells me that I can delegate mental and social tasks to her, and I think that this happens because of strong bonding and value-based goal adoption. But to get it to work, she would need to share the same set of beliefs and suitable strategies to act on them. Thus, my intuition is misguided.

\>\>\>\>\>I'm not so sure… a SP doesn't need to "share the same set of beliefs… etc"

\>

\>\>\> that my self partially includes her, and vice versa, even to the extent that I sometimes have the impression that I can leave interaction with my friends, reading a book or going to a movie to her (after all, it feels as if I was doing it by proxy through her). However, that's of course a delusion.

\>\>

\>\>\>\>\>\>\> Why are you so sure?

\>\>

\>\>\> We share similar values and have been getting along perfectly for the last fifteen years, but we do not complete and complement each other.

\>\>

\>\>\>\>\>\>\> Are you sure that's true in all domains?

\>

\> Well, I am told that I sometimes leave a trail of books, used teacups and more general mayhem in my wake, and it mysteriously rights itself. I get to do her tax returns and computer repairs, so there is a lot of domain-dependent completion going on. ;-)

\>\>\>\>Sounds good, so far

\> -- More importantly, we have a mutual understanding and acceptance of why we do what we do, and what we live for, which prevents us from feeling lonely, or getting into conflicts with each other. However, that does not require mutual completion in any interesting logical or cognitive sense. (We do not form one of the spherical hermaphrodites, like Aristophanes described them.)

\>

\>\>\>\>Still sounds good

\>\>\> On the other hand, I have co-workers that are very good at thinking and programming with me, but with them, I usually do not feel much in terms of symbiotic self-extension.

\>\>

\>\>\>\>\>\>\> Members of a team are not the same as SPSs

\>

\> Even if I can unthinkingly and intuitively delegate cognitive tasks to them?

\>

\>\>\>\>Yep. I've written a short piece distinguishing SPSs from various other sorts of relationships, and SPSs are definitely different from teams..

\>\>\> On p 549, you suggest that terrorists might be spreading their identity over many other people. Is that really the case? I suspect that here, people might be less important than ideas. That becomes even more apparent when the terrorists are not the exponents of the receiving end of colonial wars, but themselves part of western culture, like the German RAF. I think the RAF were mostly radical intellectuals identified much more with ideas (of a just society) than with the people they wanted to make a revolution for. The subjective side of SPS might have played a role, but probably not the functional one.

\>\>\>

\>\>\>\>\>\>\> I'll be happy to discuss the above further, but am still convinced that my account makes sense of many data points

\>

> That was not an objection to the theory itself, but merely an objection to a possible overextension. Or would you say that ideologies are always a stand-in for people/social agents?

>>>>No, ideologies are not always a stand-in… but in that particular case, I think the motivation lines up with what I would expect based on the SPS idea
>
>>> Wrt art and skill acquisition, which you also relate to SP seeking: in my intuition, artists and craftsmen are often obsessive "nerds", forced by a strong intrinsic motivation,
>>
>>>>>> Yes, but why are our brains wired to allow us to have such "intrinsic motivation"? That's my point…
>
> For obvious evolutionary benefits, of course. It helps if some guy starts obsessing over making a better stone axe.
> I wanted to get at the point that many artists are introverts or socially disinterested, whereas your theory might predict that higher social needs correlate with competence needs. (Or do you think that the underlying drive merely gets sublimated in one way or the other?)

>>>>I think the underlying drive gets manifested in different ways for different people -- but it's the same underlying drive (interacting with other variables)
>
>>> whereas I imagine that those good at supplicating or dominating others in SPS would be emphatic, social personalities ("geeks" or "jocks").
>>
>>>>>> Just different surface manifestations of the same underlying drive, in my view
>