4CAPS

February 14, 2002

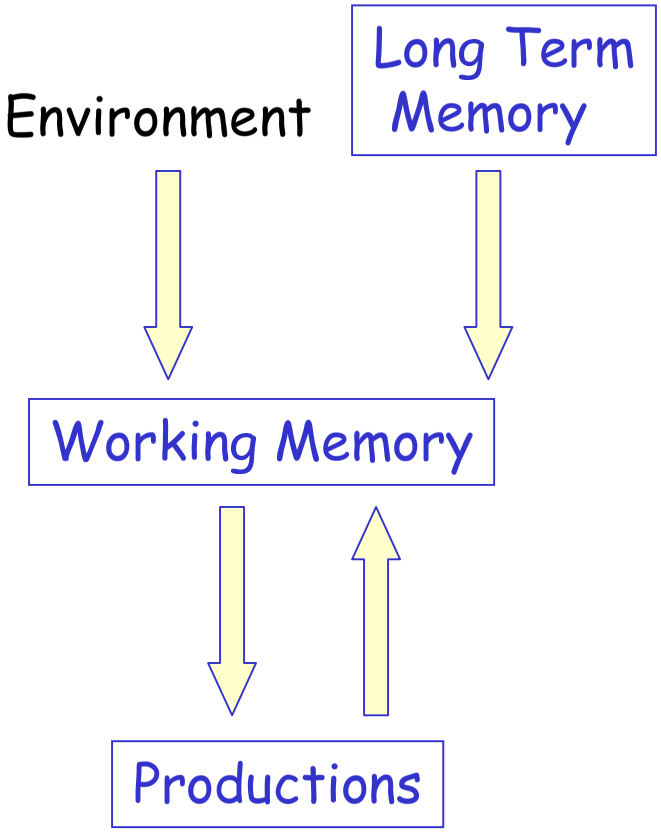
Neuroscience & Cognitive Theories

- New technologies such as fMRI call for integrative cognitive theories to account for their findings and to direct experimentation

- 4 CAPS:
  - successor of 3CAPS
  - can account for brain activation patterns
  - is inspired by brain organization and properties
  - is made of modules corresponding to brain areas

3CAPS Overview

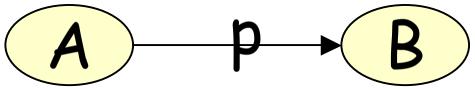
- Storage and process are both governed by activation
- Capacity = max activation available to support either storage or processing
- Productions can fire in parallel and they consume activation



Each item has a base-level activation

Each item is over a min threshold
Each item has a certain activation

If the activation of item1 > a threshold
Then increase the activation of item2



a+ d
a+2d…

Activation created in the system < Max Cap

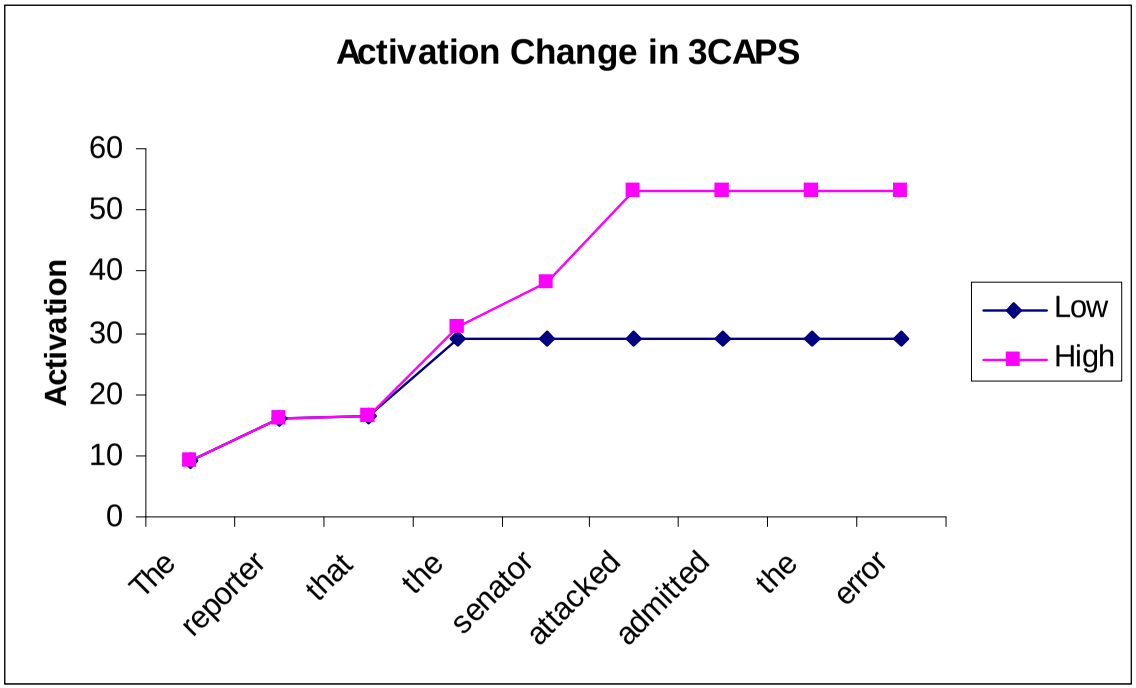
-activation propagated by productions
- activation of items in w.m.

Deallocation

- When too much activation (> max cap) would be propagated and maintain in the system:
  - deallocate activation from elements in working memory
  - scale back activation to be propagated
- Deallocation is done based on current use
- Effects of deallocation:
  - More cycles need to be spent to get an item at the needed activation level → speed decreases
  - Elements in working memory may be purged out → accuracy decreases

Outputs

- Number of cycles between the encodings of two consecutive words
- Activation in the system at various moments of time



Implications of a Capacity Theory

- Individual differences

- Modularity: there may be not enough resources for a syntax-semantics interaction

  The evidence/defendant examined by the lawyer shocked the jury

- Effects of complex embeddings (error, RTs, brain activation)

  The reporter that attacked the senator admitted the error

  The reporter that the senator attacked admitted the error

4CAPS

- Several modules, each functioning like a 3CAPS production system
- Each module corresponds to a brain area
- Each module has a capacity (brain lesion, individual differences)
- Different modules collaborate to perform the same function:
  - they can deal with different aspects of the same function (as in the Broca-Wernicke collaboration)
  - when activation quota/module is reached, other modules are automatically invoked (left DLPFC)
  - one module may be more specialized than others in performing a certain function; therefore it will be preferred if possible

4CAPS

- One module may perform multiple functions
- "Processes and representations are graded"
  - A production is executed more than once before completely fulfilling its action
  - Representation's activation changes over time

Questions

Well, this 4-Caps model that is discussed in the reading seems interesting, but everything is so very vague. I looked for more detailed reports on Marcel and Patricia's websites, but was unable to find anything specific. The website looks like it has been a year or two since its last update. I also doubt that Marcel or Patricia will just e-mail me a copy of 4CAPS or give it to me on a disk, and even if they would I don't know the status of documentation.

Since there appears to be no published recent work with 4-CAPS besides what you gave us, how is one supposed to carefully review the claims for their system and evaluate the specific methods they used to get their results?

I realize that these are questions for Patricia and Marcel, but I hesitate to ask them directly. It seems they would be distributing this information if they intended it to be widely available. Is there system proprietary, or will polite enquiries be likely to get me more detailed 4CAPS information? (Phil Pavlik)

I have to admit I don't walk away from this paper with a really good sense of how 4CAPS works in some important areas. In particular, it's not clear to me how the different modules/brain areas interact in practice -- the high-level description seems to contradict the example with Broca's, Wernicke's, and the DLPFC.

In the high-level description (p.130), it's suggested that one area recruits another area and that there's emergent collaboration between areas, but in the example, it very much seems that the DLPFC is playing an "executive processing" function, monitoring and dispatching assistance to processes in need rather than being recruited by those areas. To what extent is the processing and recruitment in 4CAPS centralized, and to what extent is it emergent?

(Ryan Baker)

How does the 4CAPS system limit the amount of productions that fire at once. It allows productions to run in parallel, so many can be running at the same time, but there has to be some kind of limit on how many. It talks about an activation, but it doesn't seem to explain much about how that activation effects the firing of the productions. The article says that the productions fire when their preconditions are met, and doesn't mention the activation. But there has to be some kind of limit on how many productions can be fired at once, or else you could do an infinite amount of computation at once. You could have the system solve multiple problems at once, seemingly without penalty. This is obviously not realistic, as people have a limit on how many things they can do at once. How does the 4CAPS system handle this? (Andy Herrman)

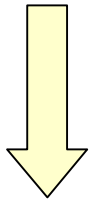
4CAPS claims there exists a number of collaborating component computational systems, each of which fires their own productions concurrently with the others, correct? ACT-R claims that there is only one central production system, correct? if each of these things are accurate about the claims of the respective architectures, then what are the implications of these two claims and how can they be resolved? (Eli Silk)

1. The productions in Table II (p.132) seem related to symbolic level events, the productions in Table I (p. 131) seem related to sub-symbolic level events. How can Table II becomes 'A specific example' of mechanism illustrated in Table I?

2. What is the operating mechanism of the multiple production systems (e.g., where and how the productions are stored? how to fire a production (e.g., check where to see if the condition part match knowing-information? How to check?))?

(Yulin Qin)

If an adjective follows a determiner
Then assume it will modify the head noun
of the noun phrase



If the proposition <Adj A follows a Det D> has
activation over a certain threshold
Then increment the activation of the prop
<Adj A modifies the head noun of the NP>

Back

1. Where are those production rules in our brain? In which form do they exist? The production rules in 4CAP architecture seem to be at a very low level of the process in our brain. In the sense of, every possible cause and effect in any single neuron could be described in an if-then form, human behaviors could be perfectly modeled by those (low-level) production rules. How should we see the fact that if one behavior could be modeled by two cognitive architectures with different levels?

2. In ACT-R, is there any mechanism that we could model the cognitive load? What are the relationships between cognitive loads, the activation in ACT-R, and the activation in an fMRI scan?

Off the wall question:

1. Since I took this course, I have been thinking how to model what I had just thought about quite frequently. However, sometimes, I find that I don' t have any particular goal in my mind. Although, later, I might start to think about what I should do next, it raises an interesting question. Do we people always have a goal in our minds? Isn't it questionable that we could just put a meta-goal, which will seek another goal when the current goal is fulfilled?

(Mon-Chu Chen)

1. Multiple Productions Firing

When multiple productions' conditions are met, ACT-R use conflict-resolution to select one over all the others. 4CAPS, on the other hand, let all fire and whichever is the most "efficient" produces the result the fastest. It seems to be that the end result is thesame: the best production gets the "say". Is this differentiation important in the functionality of the model? ... or is it only at an implementation level? What is the significance in the dichotomy of "parallel production firing" and "conflict resolution of parallel productions prior to firing"?

2. Grainsize of Data

FMRI data comes at a ~2second grain size. How does one use that kind of data to design interactions and processing at a grain size of productions (~50 ms)?

CAPS's most promising attribute is its ability to relate cognitive computations to brain activation. But the grain size of studying the brain at the level of Broca's Area, Wernicke's Area, and the grain size of individual productions involved in lexical, syntactical processing seems to be on different scales. How does the knowledge of brain region activations help in the task analysis of breaking down language processing to individual/parallel productions?

3. In parallel production, what happens when the less efficient (slower) productions fires, but produces a different response than the faster production? Are slower productions generally ignored, regardless of what they produce?

(Junlei Li)

If S is set too low in the fan effect model, then S_ji can become negative, pushing some chunks' activations negative. When this happens, ACT-R chooses the wrong answer in the experiment.

Although I have some idea of why this happens, I am still unclear of the exact way activation is spread around, and of the reasons why ACT-R chooses the wrong production in this case. Does ACT-R never pick a production with negative activation?

(Adam Goode)

I'm curious as to how 4CAPS stores declarative memory. Specifically, is there decay with time? I was also watching for any mention of whether or not declarative memory elements (chunks) could activate other memories while reading the paper. I don't recall any mention of it; an explanation of 4CAPS' handling of it would be helpful.

In general, what is the consensus in the field on chunks activating productions? For instance, if I think of "cat," I think of "it's time to change the litterbox," and the actions that ensue. How would such a structure be stored and activated?

(Peter Hu)

1. On page 130, it says the "4CAPS model dynamically recruits the appropriate team of components to perform a task". Since we don't know what the appropriate team of components is for every task, can 4CAPS only model those things which we know the necessary components, or can it still perform in new domains??

(Jonathan Giloni)