

men are dogs (and women too)

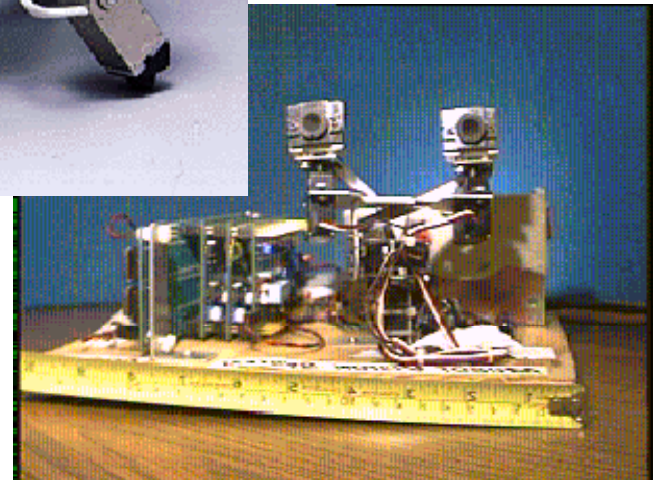
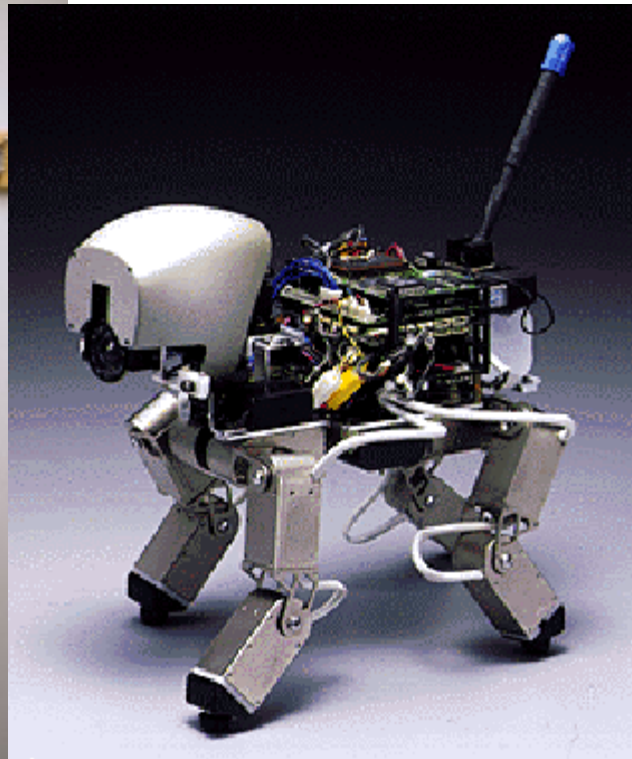
ian horswill

departments of eecs and radio/television/film

northwestern university

ian@northwestern.edu

I used to work on robots



and probably will again

but I find human behavior vexing



and I'd sure like to understand it better

interactive characters

- Strangely, nobody wants a passive-aggressive robot with Oedipal conflicts
- But it's okay for dramatic characters to be screwed up
- So they're a nice domain for modeling personality



Mateas and Stern, *Façade* (2006)

toward human-level AI

dysfunction
toward human-level ~~AI~~

claim

- The human nervous system is a **refinement of** the mammalian nervous system
- So we should use mammalian neurophysiology and behavior as a starting point for **character architectures**

folk psychology

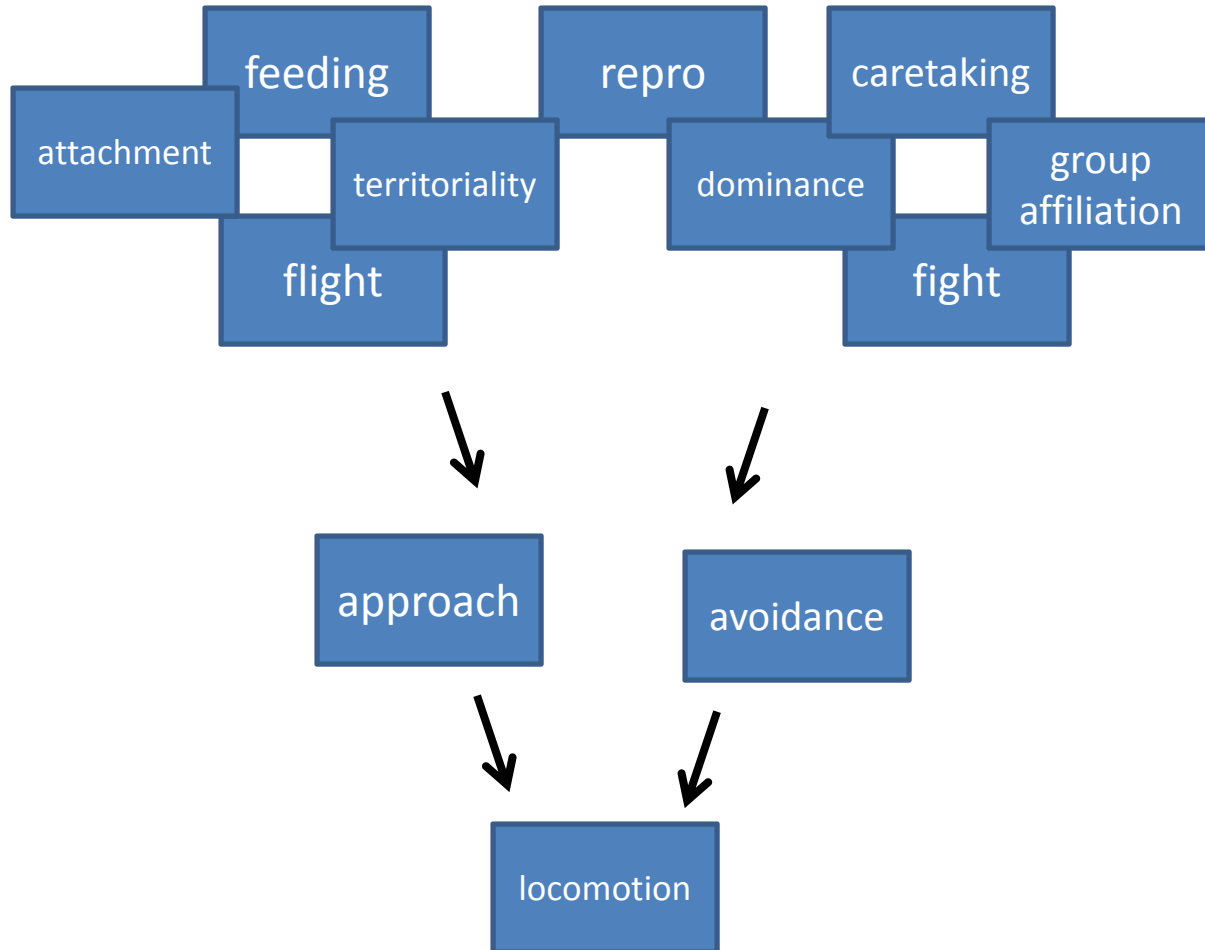
human = animal + x

$x \in \{ \text{rationality, language, thought, cognition,} \\ \text{tools, soul, culture, ...} \}$

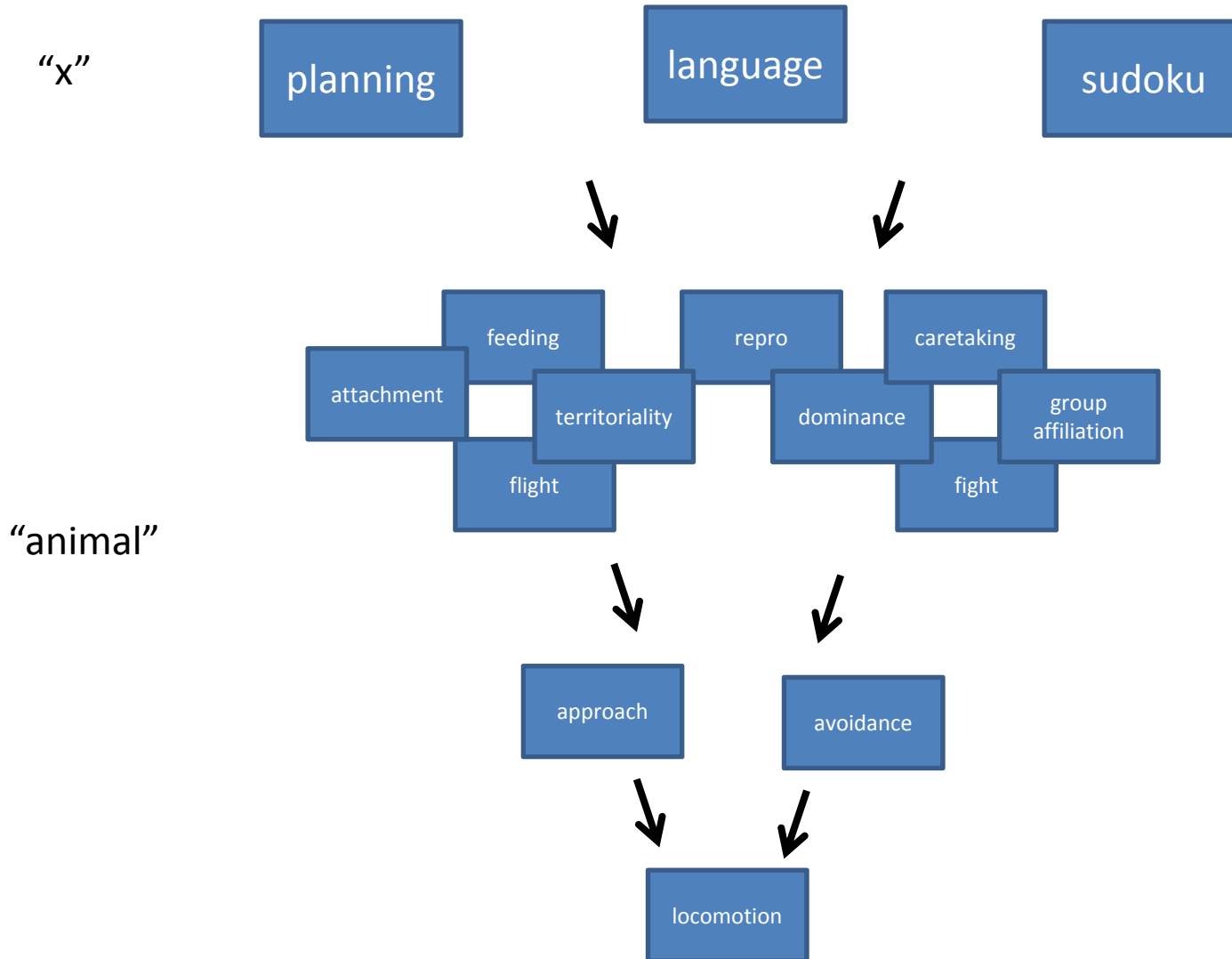
x is where the action is

“Man [sic] is the rational animal.”

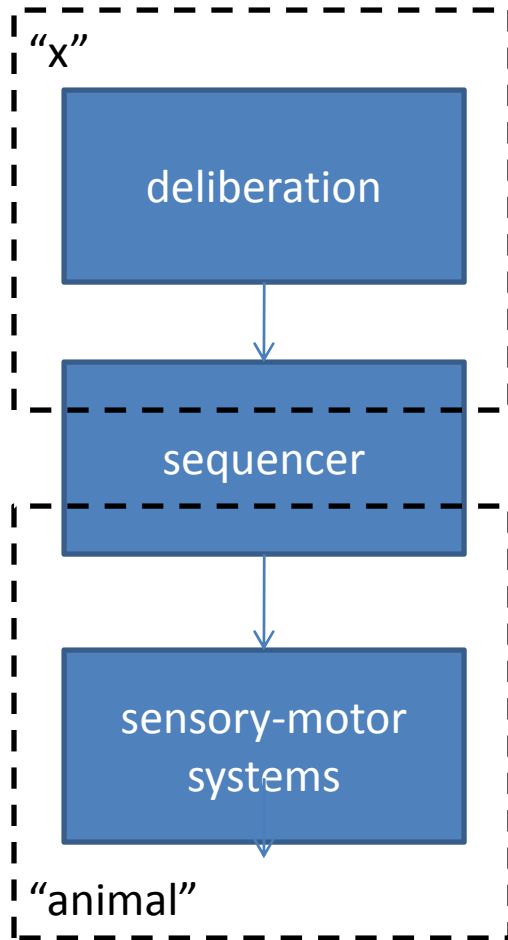
so if animals are something like ...



... then humans are something vaguely like



folk agent architecture



- Most agent architectures in use today are tiered
 - Details vary
 - Something AI-complete on top
 - Network of parallel sensory-motor systems on bottom
- “X-centric”
 - Most behavior starts with goals in a centralized cognitive system
 - Sensory-motor systems mostly do what they’re told to do by higher levels

folk agent architecture

human = animal + X

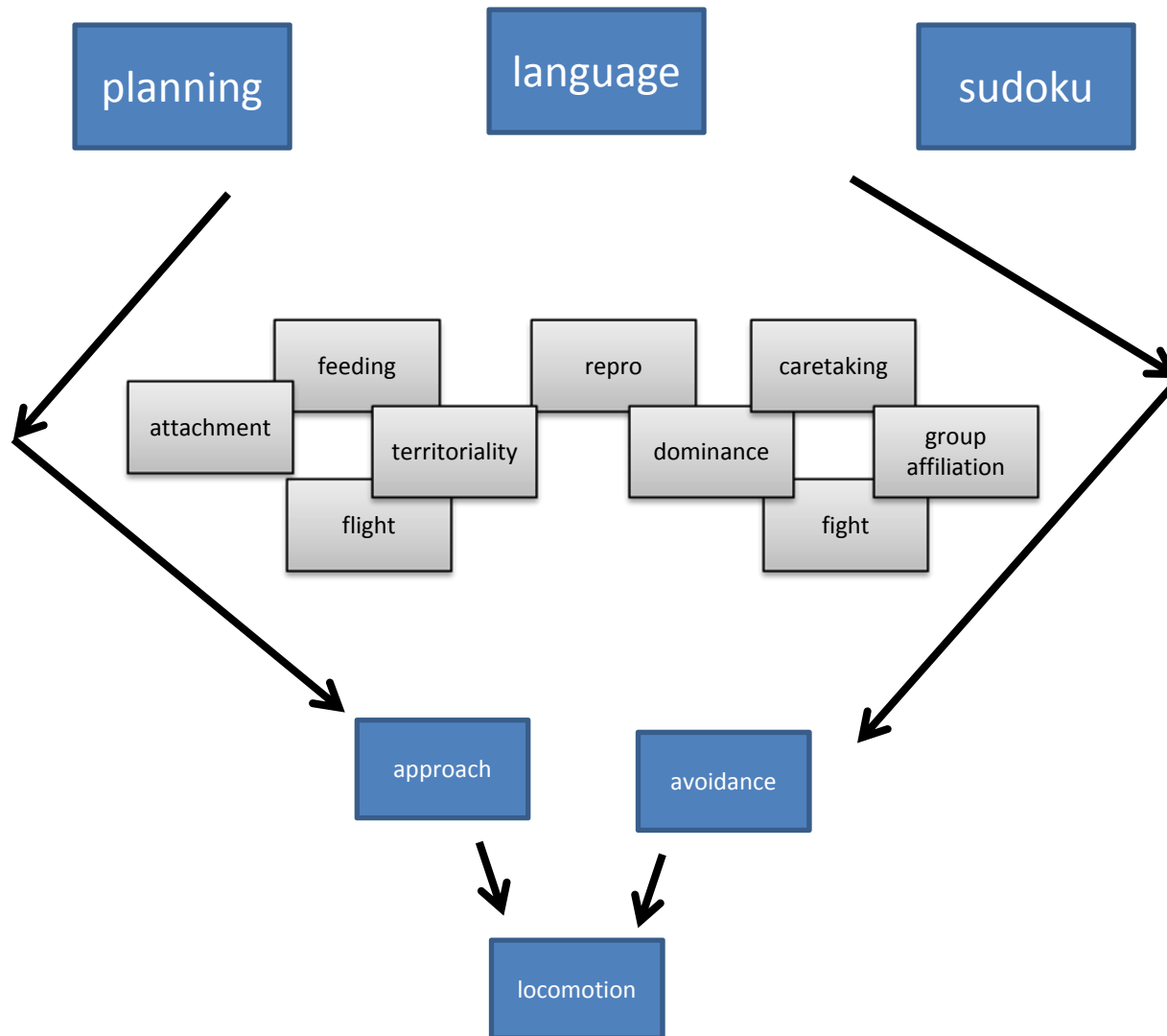
folk agent architecture

human = **subsumption** **cyc**
~~animal~~ + ~~X~~

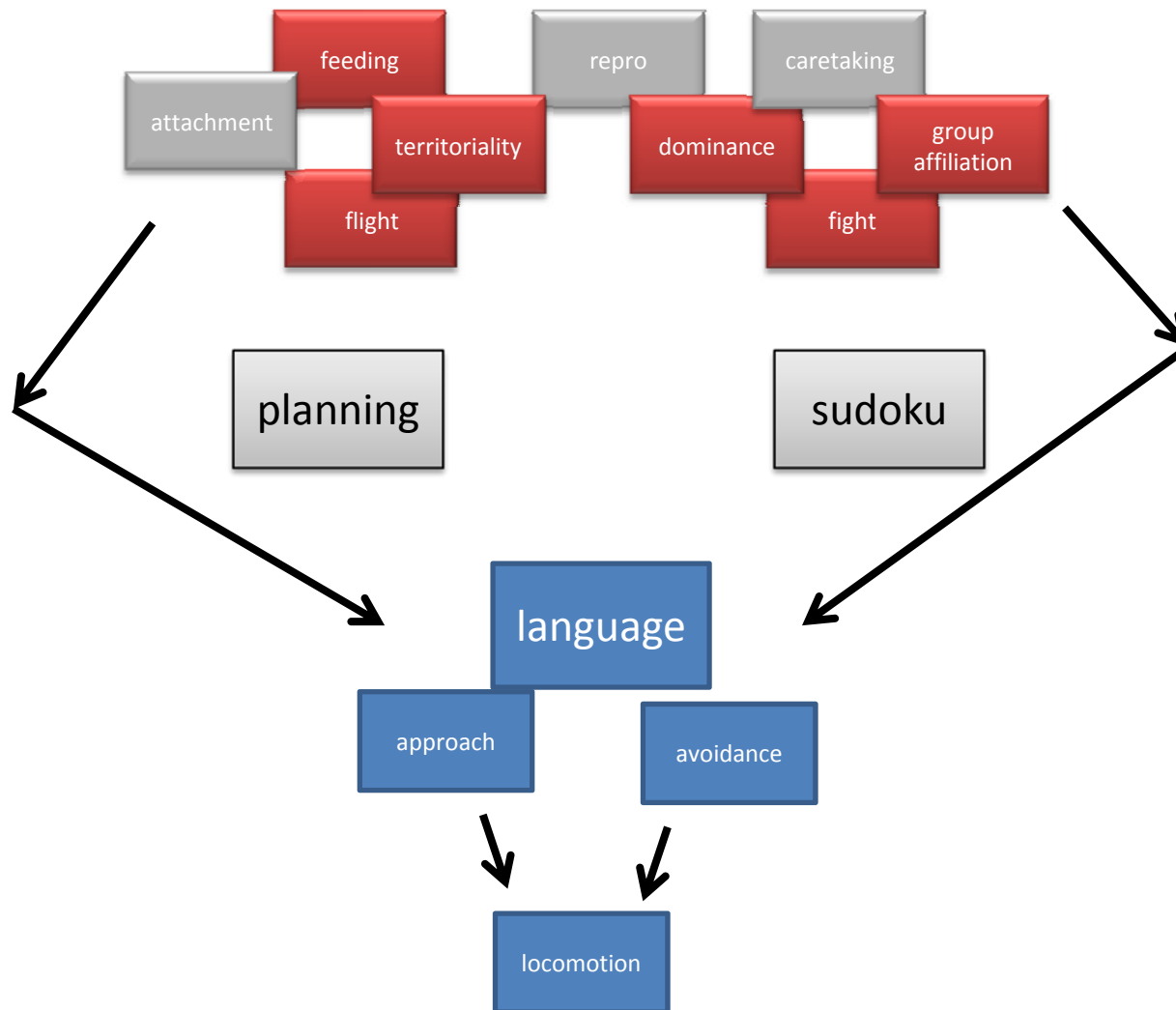
centralization

- High level systems like planners are generally Turing-complete programming languages
 - A lot of “animal” functionality gets implemented in the central system
 - Fight, flight, feeding, and reproduction
 - Emotion
 - Those functions no longer have special architectural status
- Good from an engineering perspective
- Arguably bad for character simulation
- The difference between McCoy and Spock
 - *Isn't* that Spock has self-control and McCoy doesn't
 - Even though that's the whole point of their characters
 - It's that they have different knowledge-bases

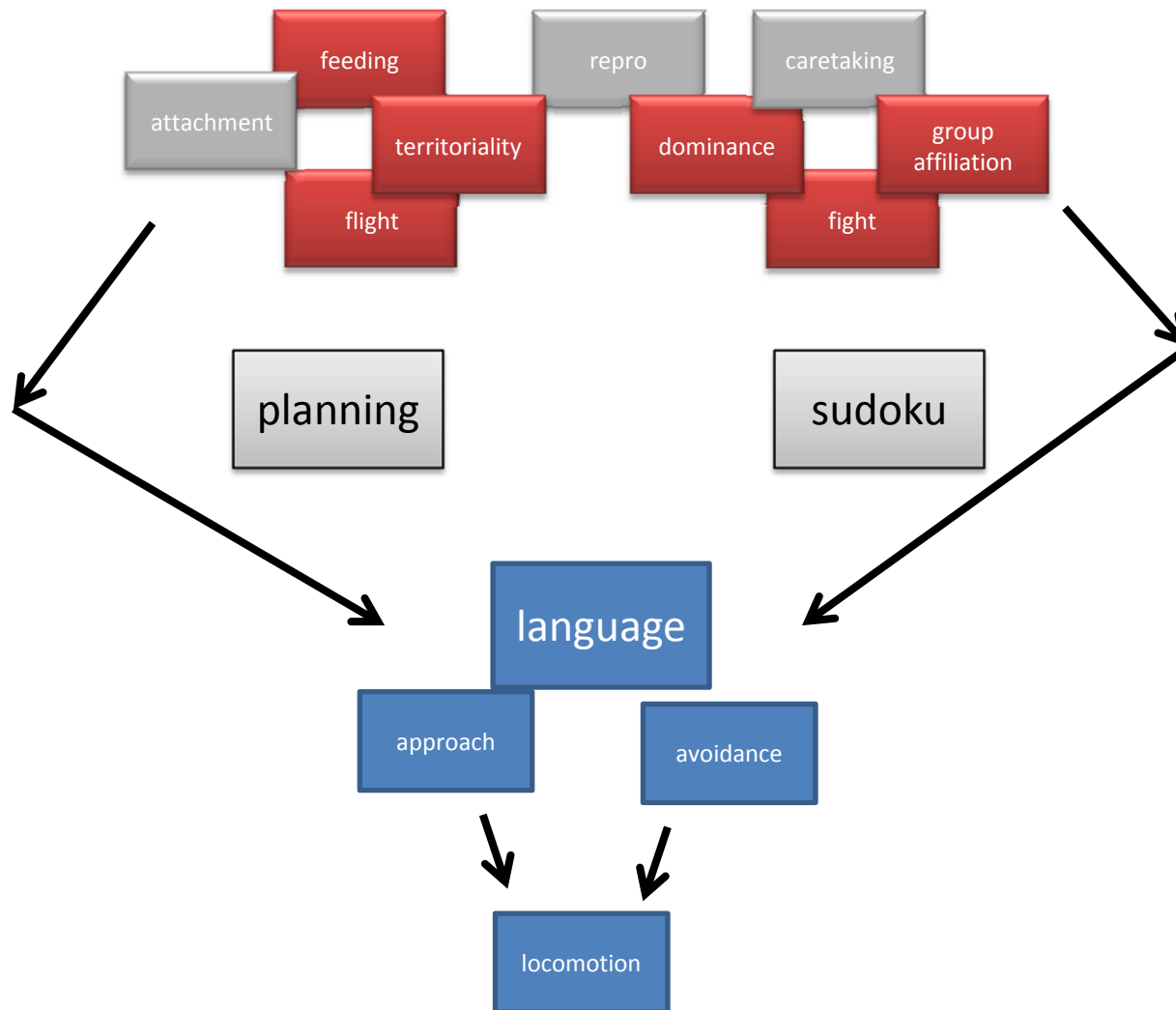
so now we have something vaguely like



and yet in faculty meetings
we sometimes seem more like ...



[in your heart, you know I'm right]



people are mammals

- Humans are social mammals
 - Affiliate into groups, tribes, etc.
 - Attachment and child rearing
 - Territoriality
 - Dominance hierarchies
- We have largely the same brain structure as other mammals
 - Just “better” somehow
 - But **all the old stuff is still running**
 - And (somehow) **influencing/being influenced by the new stuff**



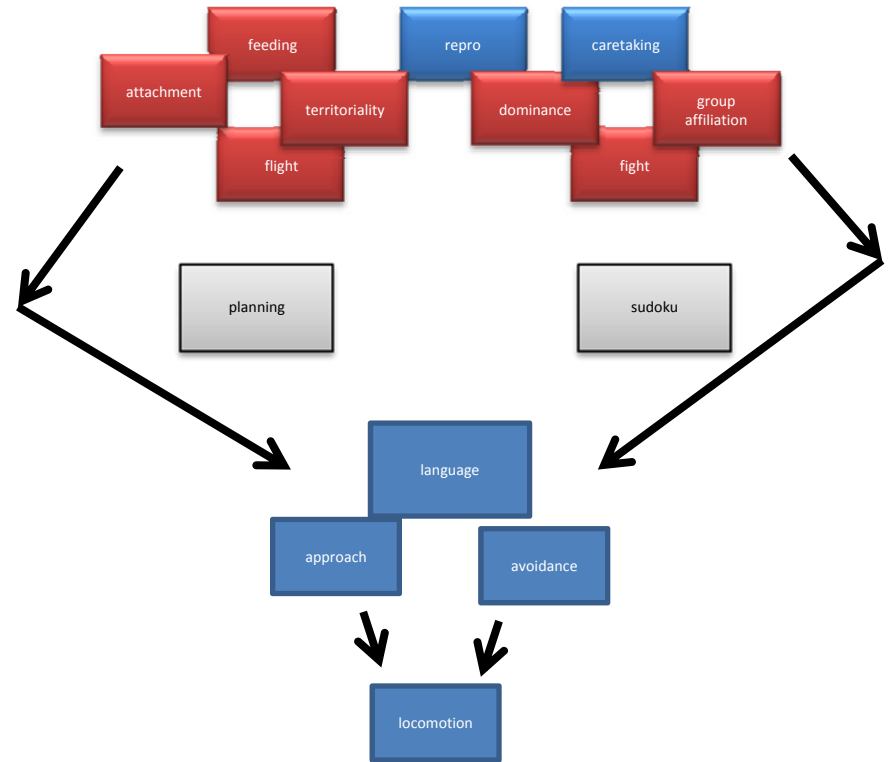
men are dogs (women too)

- **Claim:** humans are effectively **dogs with large forebrains**
- Dogs have much of the same bonding, affiliation, and dominance behaviors humans have
 - That's what matters most in characters anyway
- So we don't want to just understand how the forebrain part works
 - We also want to understand how the dog part works
 - And how it interoperates with higher-level cognition



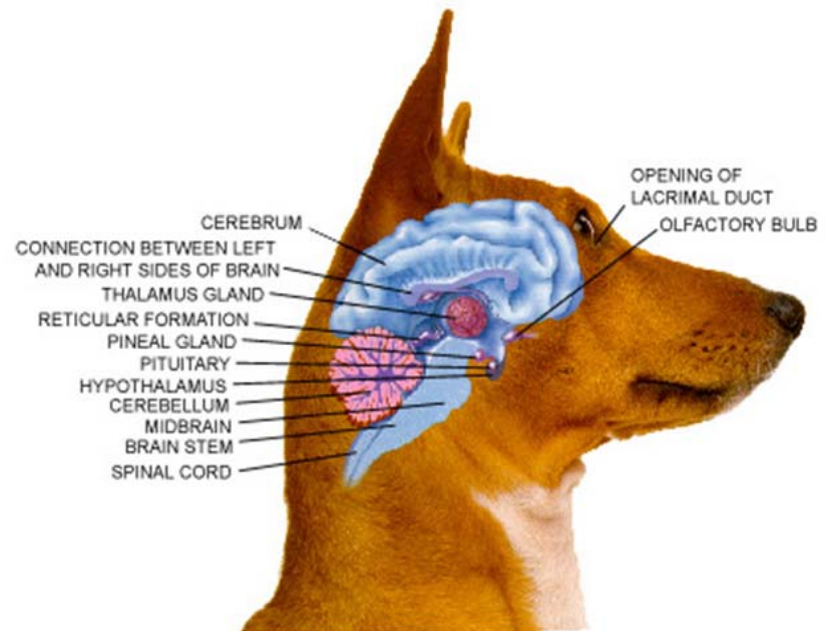
project

- Implement mammalian social behaviors
 - Including simple communication
- Use them to create interesting characters
- See how far you can take it



what do we need to add to get human-level AI?

- Probably something
 - Humans aren't literally dogs with large forebrains
 - Probably some architectural changes
- But maybe not a lot
 - There's no sign of a LISP machine having been added between chimps and humans
- The mammalian brain has
 - A largish memory
 - A finite-state controller
- That's already most of what you need to be Turing-complete
(if not AI-complete)



attachment

- Attachment is the drive to maintain proximity (accessibility) to a caregiver
- Psychoanalysis and behaviorism: attachment as a secondary drive
 - Child wants food
 - Parent gives food
 - Child wants parent
- Bowlby showed that children
 - Attach to parents even when they're abusive
 - Even in preference to surrogate caregivers who treat them better
- So he went off and read ethology, cybernetics, and cognitive science



attachment behavior system

- Bowlby argued there's an innate *attachment behavior system*
- Up and running long before language and planning

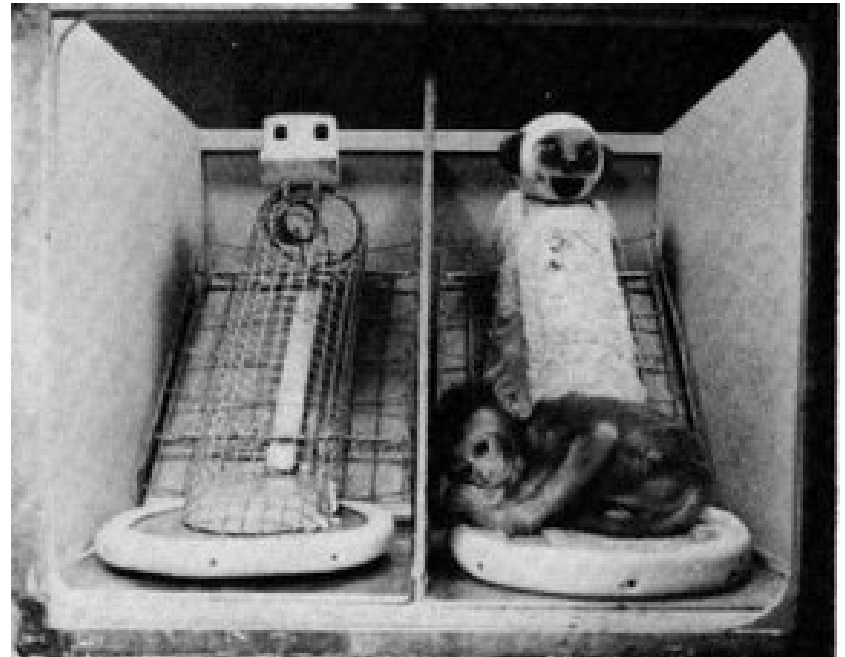


attachment is a very old system

- Most mammalian species show some kind of bond between caregivers and young
- Lorenz's work on imprinting was (presumably) one of the primary inspirations for Bobby's work



attachment is a very old system



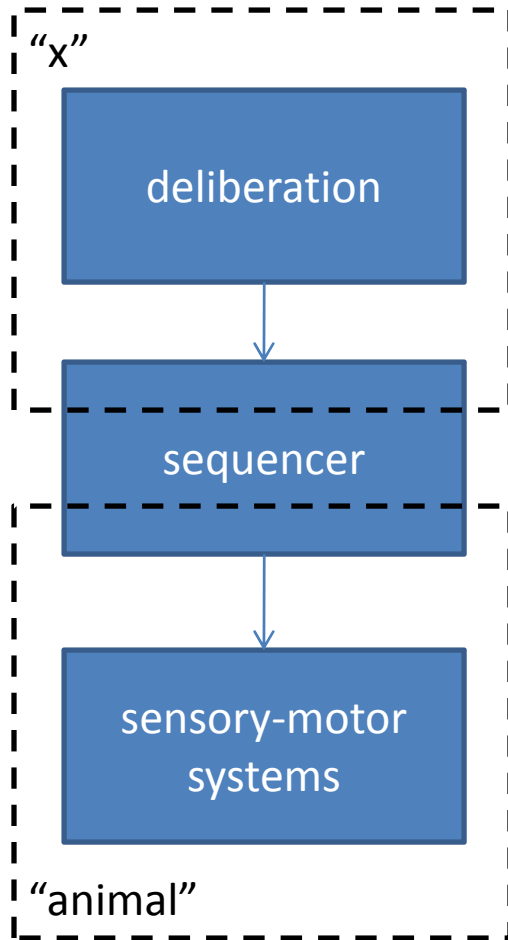
Non-human Primate infants behave almost identically to human infants in most attachment experiments

attachment and cognitive development

- Children need their caregivers to be **accessible**
- But accessibility becomes increasing abstract over time
 - Physical proximity
 - Line of sight eye contact
 - Negotiated reunions
 - Feelings talk



here's why attachment is so interesting

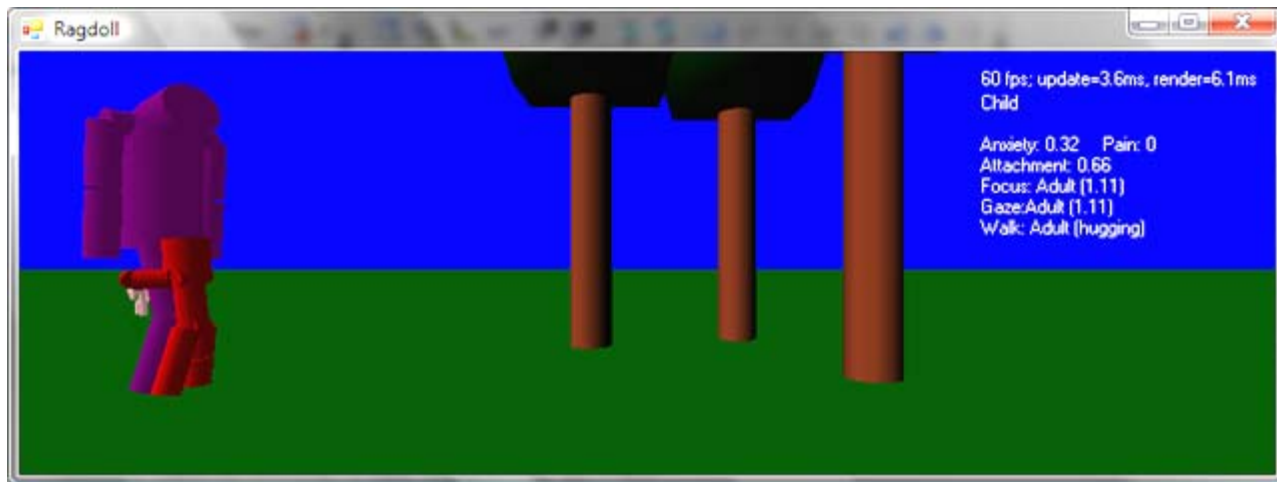


- It doesn't behave like a sensory-motor primitive
 - Acts semi-autonomously
 - Can task "x"
 - Can be influenced by "x"
- Doesn't behave like "x" either
 - Comes in much earlier than "x", both ontogenetically, and phylogenetically
 - And really does behave like an innate sensory-motor behavior during the first year of life
- **Argues for a (somewhat?) different kind of functional decomposition**
- **(Not that I know what that decomposition is)**

attachment persists into adulthood

- Attachment behavior system continues into adulthood
- People don't stop being attached to their parents
- ABS is thought to underlie adult romantic relationships
- Adult attachment style is a predictor of stalking behavior



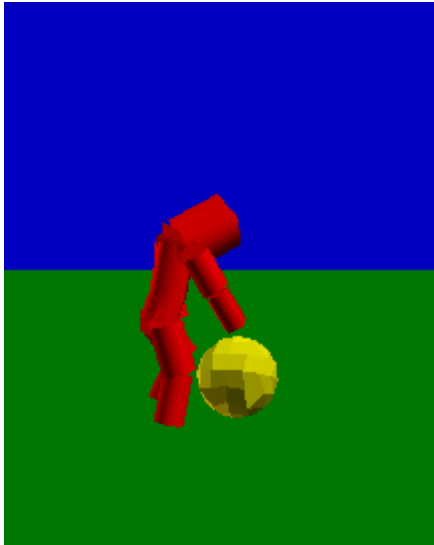


partial implementation

Simulates “safe home base” behavior (Ainsworth)

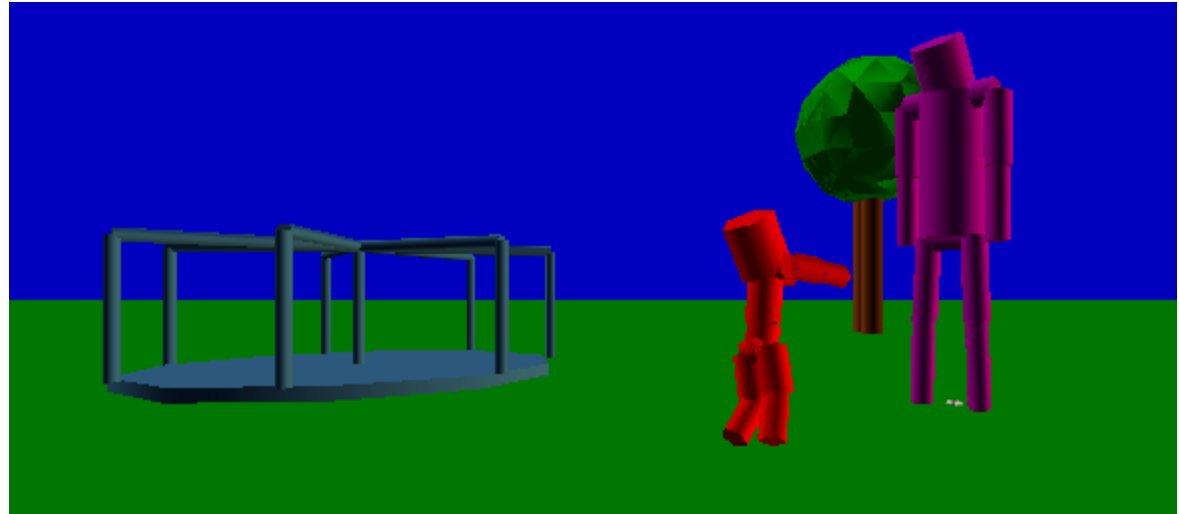
- Simple ragdoll physics simulation
- Straightforward behavior-based control
- **No higher-level cognitive component (yet)**

attention and appraisal



- Characters continually reappraise objects in view and in STM
 - Valence
 - Monitoring priority
- Valence modulated by anxiety
 - Anxious: accentuate negative appraisals
 - Secure: accentuates positive appraisals
- Focus of attention shifts to highest salience object

monitoring and gaze



- Gaze shifts regularly to monitor environment
- Mostly follows
 - Focus of attention
 - Target of current approach behavior
- But also periodically checks objects with high monitoring priority
 - Caregiver
 - Threats

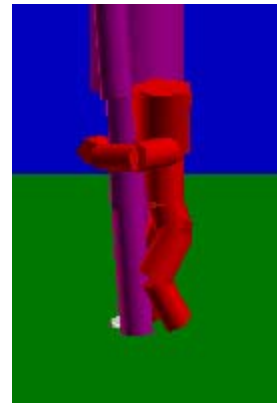
security and anxiety

- Anxiety is inverse security
(not a definitional claim;
that's just how the code
works now)
- Security increases with
 - Proximity to caregiver
 - Line of sight to caregiver
 - Eye contact with caregiver
 - Physical contact with caregiver



attachment

- Activated when security drops below threshold
- Remains active until security rises above another threshold
- Engages
 - Approach to caregiver
 - Reach
 - Hug



let us pray to the demo gods that they might smile kindly on us

demo + questions

related work

- EU Felix Growing project (Cañamero et al. 2007)
Wide range of work, including modeling on robots
(c.f. Lola's talk yesterday)
- Petters (2006)
Developed computational models that could explain
child attachment style in terms of parental caregiving
style
- Likhachev and Arkin (2000)
Use of safe-home-base phenomenon for controlling
robot mapping and exploration