Part I: TERMS

Computationalism: The classical view that cognition can be reduced to rule-governed symbol manipulation. Encompasses both GOFAI (Good Old-Fashioned AI) and Computational Functionalism.

Relevance:
- Functionalism
- The classical view, as opposed to modern views
- Gofai and the Turing test
- Behaviorism, computationalism as a measurable, algorithmic process
- Mind as a machine

Connections:
- Artificial intelligence
- Alex Djedovic's paper re: consciousness being unexplainable by computationalism
- Dualism
- What it says about identity theory, multiple ways to get to the same answer vs. a computational single formula

Searle's Chinese Room: A thought experiment where a man without any access to the external world is given Chinese symbols and, using a rule book, outputs the correct response despite not speaking Chinese. Emphasizes that correct outputs don't imply understanding from the machine

Relevance:
- Computationalism and connectionism
- Embodiment
- Connects to emotion and language as a symbol grounding method (embodiment)
- Extended mind
- Conceptualization/creating concept
- LLMs

Connections:
- Language as a cognitive tool: a person in a Chinese room can have language without understanding; are they actually using it as part of the cognitive process?
- Language and other cognitive processes are somewhat rule-based (but this is somewhat incomplete)

Connectionism: The idea of realizing cognition through weighted nodes in neural nets. These weights are variable and easily adapted with new information based on Hebbian learning

[SKIPPED RELEVANCE AND CONNECTIONS FOR TIME]

Ghost in the Machine: Gilbert Ryle's critique of Cartesian dualism, which argues that treating the mind as a separate, non-physical 'ghost' inhabiting and controlling the body—a 'machine'—is a category error. Rather than being a distinct entity, the mind is understood through patterns of behavior, dispositions, and action.

> **Category Error**: a misunderstanding of the *type* of thing the mind is → ie, if someone asked you to point to "the university", you may point to a library building, or the student body, but both are inaccurate descriptions, as "the university" is an organization of all its parts
>
> Relevance:
> - Critique of dualism, but still somewhat plays into dualism
> - Can't get away from Cartesian dualist view
> - Subjective experience can't be realized physically, but people keep trying to do that
> - Chinese room
> - Mind-body problem
>
> Connections:
> - AI/consciousness
> - Criticizes connectionism
> - Can connect to almost anything in the course

Morphological Computation: The body is not a passive bystander to cognitive processes, but an intrinsic component. The body aids in processes by minimizing the computational labour requirements of the brain.

> Relevance:
> - Embodiment
> - Somewhat extended mind
> - Self-organization
> - Puppy and SEMO robots
> - Cost of transport
> - Passive dynamic walker
>
> Connections:
> - Relevance realization
> - Autopoiesis
> - Inherently evolutionary
> - Minimal minds and waggle dance
> - DST
> - Attractor basins
> - Bluefin tuna

- Horse gallop types/robot
- Free energy principle

Bayesian Brain Hypothesis: A mathematical framework where the cognitive system is a predictive machine that makes initial predictions (priors) which are augmented and corrected by incoming sensory/emotional/environmental error to make sense and interact.

Relevance:

- Active inference
- Predictive processing
- Binocular rivalry
- Precision weighting
- Error dynamics

Connections:

- Connectionism
- Horror movies
- Describe a process (ex. being a soccer player) and how it works through this hypothesis
- Babies seek out medium uncertainty
- Optical acceleration cancellation
- Free energy principle
- Constructed theories of emotion as a predictive system
- Schizophrenia dog neural net



Extended Mind Hypothesis: Proposes that cognitive processes can extend beyond the brain to include external tools and environmental structures when they are functionally and reliably integrated into mental activity.

Relevance:

- Otto and Inga
- Glue and trust
- Parity principle
- Embodiment
- Neuralink inside the brain vs. a computer with the same function outside of the brain

Connections:

- Socially extended
- Criticism: cognitive bloat
- Predictive processing
- Context-dependent memory
- Markov blanket
- Relevance realization

Binocular Rivalry Study: A study where, when each eye is presented with a unique image, we see the images swapping from one to the other rather than seeing a mashup of them. Supports the predictive model of the mind, pointing out that our predictions matched what we saw rather than actual incoming sensory information.

Relevance:

- Predictive processing
- Average person vs. monks → meditation as a form of increasing precision weighting
- Precision weighting and prediction

Connections:

- Predisposition to recognizing faces through a predictive lens
- Connect to other perceptive, predictive experiences (ie, phantom limb studies)

Cognitive Emotional Brain: Emotion can't be separated from cognition, cognition can't be separated from the body, therefore it's all connected, where a cognitive emotional brain acts as the first step towards an embodied theory of cognition.

Relevance:

- Facial paralysis = less emotion
- Driver having road rage (paragraph explaining it in the paper)

- Constructed theory of emotion
- Embodiment
- Action/attention driver

Connections:

- Against the localization of brain functionality
- Critique of classical views of cognition/computationalism
- Dynamic coupling
- Chess players: feeling of the board, not purely thinking
- Relevance realization
- Heart beating faster is not a result of emotion; it is emotion

Error Dynamics: refers to the idea that predictive systems are sensitive not only to the magnitude of prediction error, but also to how quickly that error is reduced over time relative to expectations. Rather than simply minimizing error at any given moment, organisms aim to optimize the *trajectory* of error reduction by either updating their internal models or acting on the environment to reduce uncertainty

Relevance:

- Reduce error in two ways: (1) changing our expectations or predictive model, or (2) changing the environment to match our model
- Central mechanisms within predictive processing
- Explains short-term spikes in error/uncertainty
- Morbid curiosity, surprising behavior, etc

Connections:

- Horror movies
- Any type of short-term uncertainty and the resolution of errors
- Reward systems, dopamine as a reward for error management
- Reinforcement schedules, ie, gambling (psychology)

Predictive Processing: a framework which proposes that the brain functions as a generative model that continuously predicts incoming sensory input and updates those predictions by minimizing the difference between expected and actual input (prediction error)

Relevance:

- Error dynamics
    - The act of reducing/minimizing error is what the predictive system seeks as a reward state
    - Example: cold outside = error, can be done by changing physical reality or changing the model, but changing the model is more cognitively costly
- Free energy principle
    - Any system always wants to be as energy-efficient as possible
- Morbid curiosity

        - Horror movies give a good chance to minimize error in a safe environment (good)
        - Vs. people who have a predictive model that believes the world is morbid, seeking out things like horror movies to confirm their morbid model (bad)
    - Top down vs bottom up processing
    - The majority of our cognitive processes are top-down (neurologically)
    - Why we see faces/common things in clouds, etc

Connections:

- Neural nets
- Neuroplasticity
- Neuronal pruning
- OAC
- Tetris visual saccades

Relevance Realization: Taking a large world with infinite things to attend to and creating a small world that is easier to attend to. Three pillars are autopoiesis, anticipation, and adaptation.

Relevance:

- Predictive processing
- Meaning creation

Connections:

- Argues that predictive processing is too computational, says that cognitive systems can't be modelled mathematically
- Searle's Chinese room
- Bayesian brain: priors and posteriors

Hard Problem of Consciousness: Similar to the ghost in the machine: how does consciousness emerge from our physical body/brain?

Relevance:

- **Qualia**
    - Subjective human experience
    - What it's like to be a bat - Nagel 1974
    - Connections:
        - minimal minds
        - situated conceptualization
        - Language as a cognitive tool
- **Philosophical Zombies**
    - Thought experiment of a machine without a ghost

    - If you perfectly recreate a human with everything but consciousness what is it?
    - Connections:
        - AI
        - Functionalism
        - Chinese room
        - Criticizes the Turning Test
        - Criticizes computationalism
        - Virtual fly

Connections:

- Ghost in the machine
- Embodiment
- Error dynamics = qualia (?)
- Active inference
- Levels of consciousness

Terms to know:

- Computationalism
- Turing Test
- Searle's Chinese Room
- Ghost in the Machine
- Morphological Computation
- Bayesian Brain Hypothesis
- Binocular Rivalry Study
- Cognitive Emotional Brain
- Error Dynamics
- Predictive Processing
- Free Energy Principle
- Morbid Curiosity
- Relevance Realization
- Hard Problem of Consciousness
- Qualia
- Dynamical Systems Theory

Part II: ESSAY

Q1: Is the Mind a Computer? Critically evaluate the computational theory of mind. In what ways does computationalism succeed or fail in explaining human cognition?

Cover: GOFAI & Connectionism, the Paradox of Mechanical Reasoning, original meaning, Ghost in the Machine, and at least one alternative framework (Embodied Cognition, Dynamical Systems, or PP). Novel angle: use LLMs or the AI alignment problem as a contemporary test case for computationalism's limits.

Q2: Predictive Processing Framework. What is Predictive Processing, and how does it challenge traditional models of perception and cognition?

Cover: Bayesian Brain Hypothesis, Free Energy Principle, top-down vs. bottom-up processing, precision weighting, Error Dynamics, and at least one connection to consciousness or morbid curiosity. Novel angle: placebo/nocebo effects as an empirical case, or addiction as a failure of error dynamics.

Q3: Future of Cognitive Science. Which model of cognition provides the best framework for understanding the mind? Defend your position with examples from the readings.

Cover: Computationalism, Connectionism, Dynamical Systems, Embodied Cognition, Predictive Processing. Pick one to defend and engage seriously with its rivals. Novel angle: ground your argument in a specific empirical case— the hollow mask illusion (PP), passive dynamic walkers (Embodied), or LLM failures (Computationalism).