I. Setting the Stage

A. So far we are worried about how linguistic expressions such as sentences obtain their meaning. However, it is plausible to think that the semantical value of linguistics expressions is parasitic on our mental states.

B. Although the problem is harder since it involves minds, we must tackle the question of what accounts for the intentionality or aboutness of mental states.

C. To keep the problem manageable, consider the propositional attitudes: believing that snow is white, desiring that I drink a cold beer. The content of these states is the proposition: that snow is white and that I drink a cold beer (respectively). What explains how content is "attached" to mental states (or brain states presuming you are a naturalist about the mind)?

D. Note that Davidson and Soames both come to the conclusion that a theory of meaning need not explain the relation between (say) atomic sentence and its truth conditions. But the problem posed here requires we do that for mental states.

II. Buzzwords: Ideas Cummins Relies on

A. Folk Psychology (FP) presumes the existence of propositional attitudes with contents, and holds that they are genuinely involved in an explanation (perhaps a causal explanation) of ones behavior. For example, why did Joe go to the fridge? He had a desire: that Joe drink beer, and a belief: that there is a beer in the fridge. Were he to have had a different desire and/or a different belief, his actions would likely have been different. So contents matter to explanation of behavior.

B. Computational Theory of Cognition (CTC) is the thesis that our best chance of understanding mental content is to think of the mind/brain as being like a computer, which operates over **data structures**, that is, explicit language-like symbolic representations. Since we (presumably) know a lot about how data structures manage to have content in computers, the lessons learned may be imported to a theory of mental content.

C. Causal or Inferential Role theories of content are inspired by CTC. What makes the symbol '00000010' mean the number two is that the program of the computer puts that symbol in a causal role that satisfies all the arithmetic properties of two with respect to its interactions with all the other symbols. Such theories are therefore holistic.

D. The Language of Thought (LOT) is the hypothesis that human cognition depends on a symbolic system that is language-like.

E. Implicit Content is mental content that is not explicitly represented in a symbolic system. An example is that the visual system "knows" that objects are rigid, without representing that idea in any symbolic way. Some suggest that the LOT is false because all mental content is implicit rather than explicit. Cummins thinks not.

F. The Frame Problem describes a set of related difficulties faced by artificial intelligence in trying to duplicate human intelligence. If we load the computer with a symbolic representation of the world *knowledge of a typical four-year old, then processing of even rudimentary reasoning about how things work is bogged down by the sheer size of the task. Human reasoning seems to avoid these computational difficulties in a way that artificial intelligence researchers have yet to duplicate.

G. Eliminativism with respect to mental content is the thesis that propositional attitudes do not exist, in that these concepts do not have any useful role to play in the best science of human cognition. In the context of providing an understanding of how meaning plays a role in human cognition, the eliminativist has a duty to propose an alternative to the CTC that explains human cognition without appealing to propositional attitudes and symbolic data structures like a LOT. Some connectionists who model the neural net structures in the brain have attempted to argue an eliminativist position.

III. Covariance Theories

A. Why are we interested in Covariance Theories of Content?

1. It is plausible to think that some words have meanings defined by their causal/inferential roles with respect to other words. But this picture doesn't seem plausible for certain words, notably proper names and natural kind terms ('water').

2. Imagine that Jim lives on Earth and Twim lives on **Twin Earth** where a different chemical XYZ plays the role of water, but Jim and Twim are otherwise molecular duplicates. A widely held intuition is that by 'water' Jim means H2O, while by 'water' Twim means XYZ. Furthermore, Jim's mental state for "Lo water" is about H2O and Twim's is about XYZ. What could explain this? Perhaps covariance theories can fill in this gap. The rough idea is that Jim's perceptual or recognitional mental state meaning |Lo water| can be identified by what (typically) causes it: (H2O), and similarly for Twim: (XYZ). Once the meaning of perceptual mental states are fixed in that way, a causal/inferential role theory will help fix the meanings of other mental states.

B. The Causal Theory of Reference: Another Motive.

1. In the case of language, what is called the causal theory of reference (not to be confused with causal role theories of meaning) is a popular account of how reference gets fixed (cf. S. Kripke, "Naming and Necessity"). It says (roughly) that 'Jim' said by me is about Jim because of an act of "baptism" where Jim was given the name 'Jim'.

2. My use of 'Jim' successfully picks out Jim even if I never attended the baptism, because of a chain of people who use the term who hand the referent from one to another down to me in something like a causal chain. So my use of 'Jim' means Jim because there is a complicated causal chain leading from Jim to me.

C. Crude Covariance Theory.

1. Of course it would be crazy to say that a perceptual brain state H gets the meaning |Lo a horse| due to an act of baptism that gets carried down to me. But there is still room to try a variation of this idea, namely that H gets that meaning because it is caused by horses in my environment

2. The idea was that Jim's mental state W means |Lo H2O| because it is H2O that causes it, while Twim's state W means |Lo XYZ| because XYZ does the job on Twin Earth.

3. More accurately, brain state H in me means |Lo a horse| (or |horse| for short) because there is a perceptual causal mechanism in my brain that causally insures that H covaries with the presence of horses. That is, H occurs in the presence of horses, and fails to occur in their absence.

4. Note the kinships between convariance theories and Quine's **stimulus meaning.** 'Gavagai' has the stimulus meaning |rabbit percept| because <u>assent</u> to 'Gavagai' covaries with the presence of the rabbit percept.

5. Terminological Point: Here we are not discussing causal role theory, nor the causal theory of reference. To stay on track, we will speak only of covariance theory.

IV. The Disjunction Problem

A. Why there is a Problem.

1. The main difficulty for covariance theories is to account for the possibility of error. If shrews (by mistake) cause a state S with meaning |mouse|, should not that very causal fact count as evidence that we have misidentified the content of S, and that it should instead be |mouse or shrew|? But if so, then S is not in error, and in principle no perceptual state would be in error under this policy.

2. To express this in highly abbreviated (and somewhat misleading) form, covariance theory says |A| iff A, but that directly rules out error: |A| but not A.

3. Here are some attempts to solve the problem.

B. Ideal Conditions

1. The tactic is to replace (exact) covariance with covariance under ideal, typical, or normal conditions C:

|A| iff A under conditions C.

2. Now error is possible because you can have |A| and not A when conditions are abnormal, i.e. when C does not occur.

3. The problem now is to come up with a C that does not appeal to intentionality. Obviously it won't do to say that normal conditions are ones where A occurs when |A| occurs, since this presumes we know the meaning of |A|.

a. Try: C = the perceptual apparatus is functioning properly.

Objection: Error often occurs when the perceptual apparatus is working fine (e.g. when viewing the Ames Room.)

b. C = what was in a. plus the conditions of observation are ideal.

Challenge: Give me an account of 'ideal' that doesn't appeal to intentionality. For example (Ideal): won't work since the definition clearly employs an intentional notion: about.

(Ideal) Conditions are ideal iff they are conditions where what the state is about is present.

Objection: Even if you can meet the challenge, there is another problem, namely that even for this new C, error occurs when C holds. You can have a horse present under ideal viewing conditions and perfectly functioning perceptual apparatus and fail to have (the state with meaning) |horse|. For example, the person may believe they are hallucinating. Unless the right inferential relations hold between percepts (horsey looks) and concepts (|horse|), horsey looks won't cause |horse|. What mediates the connection between horses and |horse| is inference based on learned *knowledge. (The * indicates that the knowledge need not be accurate. This is to head off the objection of the stickler philosopher who insists that X is not knowledge unless X is true.) For example, *knowledge might include that if you get horse visual percepts along with barnlike smell percepts and whinny aural percepts, then the visual percept can be relied on. So we need to add to C that inferential relationships of a certain kind are functioning correctly.

c. C = b. plus: *knowledge-based inferential relations between inputs and concepts are functioning correctly.

Challenge: Present an account of correctly functioning inferential mechanisms based on the right *knowledge without bringing in ideas dependent on the notion of mental content.

C. Semantic Reductionism (Verificationism)

1. Use the crude theory without C, and admit the theory fails for concepts such as |horse| or |proton| (where the theory is really implausible anyway.)

2. Claim, however, that the theory does work for raw percepts, that is for states with meaning (say) |Lo a red-sense-datum|. Presumably, there is no problem of error for raw percepts because we are never mistaken about those.

3. Then define the contents of all other states by showing how all contents are defined via the appropriate claims about raw percepts. (This should remind you of the old logical positivist verificationist epistemology: all knowledge is rooted in sense data, and all thoughts are logical constructions out of sense data.)

Objection: Verificationism has been a miserable failure. There may be no veridical raw percepts. (Percepts without concepts are blind, to paraphrase Kant.) Furthermore, no one has ever successfully provided a sense data account of even 'there is a chair in the room'. "If the lighting were good, and nothing were in the way and my eyes were pointing in the right direction, and I were not asleep and paying attention ..... , I would have a sense datum of a chairish shape", is of course

hopeless. There are lots more things to add where the '.....' occurs, and most of the antecedents mention objects not sense data, and so will need further elaboration.

D. Asymmetrical Dependence

1. Let ? be the state caused mice and (purportedly) incorrectly by shrews. What makes it be the case that |?|=|mouse| and error occurs, rather than |?|=|mouse or shrew| and no error occurs?

2. The asymmetrical dependence tactic is to show an asymmetry between cases where shrews cause |?| in error and mice cause |?| correctly. So the covariance theory is revised to read: |?| = |mouse| because:

mice cause ?, and any case of non-mice (shrews) causing ? asymmetrically depends on mice causing ?.

3. Asymmetric dependence amounts to (1) being true and (2) being false:

(1) If shrews are to cause ?, then mice must cause ?.

(2) If mice are to cause ?, then shrews must cause ?.

(Note: For simplicity, I have contraposed Fodor's: were mice not to cause |mouse|, shrews would not either. This is legal because ~B→~A entails A→B in any reasonable conditional logic. I also use '?' instead of '|mouse|' since the latter notation confuses the state with its meaning, and uses a notation that presumes what is in question, namely that the state has content |mouse|.)

4. The idea behind (1) being true is that shrews "poach" on the mouse-to-?, connection. Shrews being able to cause ? depends on the mouse-to-? connection already being there.

5. The idea behind (2) being false is that it is possible for mice to cause ? and shrews not to cause ?. For example, we could come to be able to discriminate shrews from mice, and so shrews would fail to set off ?.

6. The disjunction problem is claimed to be solved because though shrews cause ?, that causation is asymmetrically dependent on mice causing ?. Therefore shrews causing ? is irrelevant to the content of ?. So |?|=|mouse|, and not |mouse or shrew|.

7. Objection. (Simplification of Cummins) (1) is false not true. Take a world where our mental equipment is the same but where all mice have a disfiguring disease, so that only shrews cause "mousey looks" (the ~percept" caused by mice or shrews in our world). Here shrews will cause ?, but mice won't. (So we wont beg any questions here, let's refer to the percepts as "small-rodent-looks".)

8. Reply. Granted, shrews causing ? does not depend on mice causing ?, because mice might stop looking like shrews, but it does depend on small-rodent-looks causing ?. So if we concentrate on the connection between the **percept** and ?, we get to rule out the disfiguring disease counterexample.

9. Objection. But now (2') is no longer false! If small-rodent-looks don't cause ? then mice won't either.

(1') If shrews are to cause ?, then small-rodent-looks must cause ?.

(2') If mice are to cause ? then small-rodent-looks must cause ?.

10. Reply. OK replace (2') with the old (2) which <u>is</u> false:

(1') If shrews are to cause ?, then small-rodent-looks must cause ?.

(2) If mice are to cause ?, then shrews must cause ?.

11. Objection. This test generates a bogus asymmetry by being asymmetric itself. By this test, shrews causing ? is asymmetrically dependent on mice causing ?. But now let's check whether mice causing ? is asymmetrically dependent on shrews causing ?. Using the recipe for asymmetric dependence given in 10, we obtain:

(1") If mice are to cause ?, then small-rodent-looks must cause ?.

(2") If shrews are to cause ?, then mice must cause ?.

Note that (1") is true and (2") is false. (2") is false, remember, because of the disfigured mice case. It follows that both the shrew-to-? and the mouse-to-?

connections are asymmetrically dependent on each other. So the asymmetry has vanished, and we have no reason to choose |?|=|mouse| over |?|=|mouse or shrew|.

12. Moral of the story. You can't get something for nothing. Imagine that you have ? on your hands, and it is caused by both mice and shrews. You want to assign content to ?. The asymmetric dependence strategy will ask you to evaluate some counterfactuals like (1) and (2). But how you evaluate those counterfactuals will depend on your intuitions about what ? is about. If we are to explain error, we need more to go on than the fact that our intuitions rule that there is error. Whatever tactic we use to solve the problem of error, we must point to some empirically testable method that does not rest on our intuitions that error exists. So a good account must point to some cognitive difference between mice causing ? and shrews doing so. No matter how that story goes, citing certain counterfactuals won't help. We would still want to know the scientific basis for the asymmetry described in the counterfactuals.

V. The Problem of Omniscience

A. What is the Problem?

1. The disjunction problem is to explain how covariance (|A| iff A) is compatible with error: |A| and not-A. The omniscience problem is to explain how covariance is compatible with the fact that for most A, A does not cause |A|. (If As always cause |A|, the I would be omniscient, always knowing everything that occurs, even things outside my perceptual range.)

2. The problem reformulated.

A iff |A| is false since lots of As don't cause |A|.

Furthermore, "A and observed by me iff |A|" is closer to the truth. Therefore the state |?| caused by A is more aptly described as |A and observed by me| or better: |A and A causes |A||. But it's crazy to say that my mental states are about (say) horses-that-cause-my-mental-states. Clearly, they are about horses, even horses that don't cause states in anybody.

B. Outline of a Solution.

1. Revise covariance to say:

If A occurs, then |A| occurs under conditions C.

2. C = given a fair chance (i.e. the perceptual apparatus is OK, the conditions of observation are good, you are attending, your inference mechanisms are good, your *knowledge is good, etc.)

3. Challenge: Give this story without appealing to intentionality.

C. Fodor's Verificationist Answer. Let C= the perceptual apparatus is OK, and conditions for observation are good. Then you can get a good account for the content of raw percepts (since these, by Fodor's lights, don't involve inferential mechanisms). Then try to build the rest of content out of this covariationist theory of percepts.

Objection: See the troubles with the verificationist strategy mentioned above.

VI. The Problem of Indeterminacy of Reference

1. Suppose you solve the preceding problems. Another remains. Where there is covariance with ? at all, many things covary with ?, so how do we determine which of these ? refers to?

2. Why pick horse for |H| over other candidates that covary such as: a) horsey irradiation nearby b) horsey image in the visual field, c) horsey image on area V1 of cortex, d) presence of a temporal stage of a horse or undetached horse parts ...?

3. To narrow down the content to the "right" item, we will have to bring in more than mere covariance. Can you do it without relying on intentionality?