

# Minds and Machines

spring 2003

Content:  
intentionality and externalism



from the philosophical toolkit:

**extrinsic** and **intrinsic** properties

perfect duplicates:

not-so-perfect duplicates:

# extrinsic and intrinsic properties

the all-purpose  
duplicating machine



# extrinsic and intrinsic properties

- property  $P$  is *intrinsic* iff it is necessarily shared between perfect duplicates



- otherwise,  $P$  is *extrinsic*
- (there are many complications here, but we will ignore them; see the [handout](#) on properties and particulars)

# extrinsic and intrinsic properties

- being cylindrical
- having mass 200g
- being red

(intrinsic)



- being a Campbell's soup can
- having weight 200g
- being in the cupboard

(extrinsic)



# are mental properties intrinsic?

yes, according to:

- Descartes (well, arguably)
- the identity theory (taken as theory of all mental states, not just properties like being in pain)
- functionalism and behaviorism (on one natural way of spelling these theories out)
- commonsense(?)

“thoughts are in the head!”

# are mental properties (of kind K) intrinsic?

- yes, according to **internalism** (about mental properties of kind K)
- no, according to **externalism** (about mental properties of kind K)
- we are about to look at some famous arguments for externalism about “propositional attitude” properties like wanting a glass of water, believing that Cambridge is pretty, etc.

# a “twin earth” thought experiment



# Cambridge and twin-Cambridge

Harvard Square



Hilary

twin-Harvard Square



twin-Hilary

# Cambridge and twin-Cambridge

“Cambridge is pretty”



Hilary

“Cambridge is pretty”



twin-Hilary

# different utterances

- Hilary's utterance is about *Cambridge* (not Twin-Cambridge, of which he has never heard)
- his utterance is true iff Cambridge is pretty
- the aesthetics of twin-Cambridge are totally irrelevant—if we imagine that twin-Cambridge is an imperfect duplicate of Cambridge (a twin Harvard Square, but exceptionally attractive elsewhere), then Hilary's utterance remains false, although twin-Hilary's utterance is true

# Cambridge and twin- Cambridge

“Cambridge is pretty”

“Cambridge is pretty”



Hilary



twin-Hilary

# different thoughts

- Hilary's thought is about *Cambridge* (not Twin-Cambridge, of which he has never heard)
- his thought is true iff Cambridge is pretty
- the aesthetics of twin-Cambridge are totally irrelevant—if we imagine that twin-Cambridge is an imperfect duplicate of Cambridge (a twin Harvard Square, but exceptionally attractive elsewhere), then Hilary's thought remains false, although twin-Hilary's thought is true

# “The meaning of ‘meaning’”

- this contains the original presentation of the “twin earth” thought experiment
- Putnam drew the conclusion that *meanings* aren’t “in the head” (i.e. aren’t intrinsic), not that thoughts aren’t in the head
- this paper is quite difficult, and we won’t discuss everything in it

# “The meaning of ‘meaning’”

- **intension** and **extension**
  - the *extension* of a term (e.g. ‘rabbit’, ‘creature with a kidney’) is “the set of things the term is true of”
  - so ‘creature with a kidney’ and ‘creature with a heart’ have the same extension
  - the *intension* of a term is its “meaning”, in the intuitive sense
  - so ‘creature with a kidney’ and ‘creature with a heart’ have different intensions

# “two unchallenged assumptions”

- 1) knowing the meaning of a term is just a matter of being in a certain psychological state
- 2) the intension of a term determines its extension (in the sense that sameness of intension entails sameness of extension)



# “two unchallenged assumptions”

“I shall argue that these two assumptions are not jointly satisfied by any notion, let alone any notion of meaning. The traditional concept of meaning is a concept which rests on a false theory.”

# “methodological solipsism”

“no psychological state, properly so-called, presupposes the existence of any individual other than the subject to whom that state is ascribed.”

if this is right, then:

- being jealous of Smith
- seeing Jones
- knowing that the cat is on the mat
- remembering last night’s party

are not psychological states, properly so-called

- psychological states, according to ms, are “narrow”

the first “unchallenged assumption”,  
stated more precisely

- 1) knowing the meaning of a term  
(knowing that intension I is the  
meaning of term A) is just a matter of  
being in a certain *narrow*  
psychological state

# consequences of (1) and (2)

- suppose Oscar<sub>1</sub> and Oscar<sub>2</sub> are in the same narrow psychological states, and that Oscar<sub>1</sub> knows that intension I is the meaning of term A
- by (1), if Oscar<sub>1</sub> knows that I is the meaning of A, then Oscar<sub>2</sub> knows that I is the meaning of A
- if Oscar<sub>1</sub> and Oscar<sub>2</sub> both know that I is the meaning of A then Oscar<sub>1</sub> and Oscar<sub>2</sub> both use A with the same meaning or intension
- by (2), A as used by Oscar<sub>1</sub> has the same extension as A as used by Oscar<sub>2</sub>

“It is this last consequence...that we claim to be false. We claim that it is possible for two speakers to be in exactly the *same* psychological state (in the narrow sense), even though the extension of the term A in the idiolect of the one is different from the extension of the term A in the idiolect of the other. Extension is not determined by psychological state.”

# Putnam's twin earth



earth



twin earth

a perfect duplicate of earth,  
except...

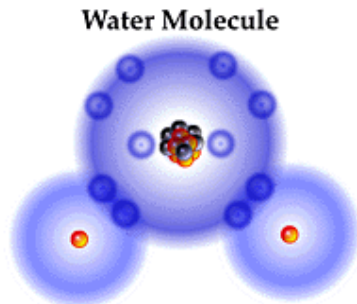
# Putnam's twin earth



earth



twin earth



...the oceans and lakes contain “XYZ”, which is a very different chemical kind from H<sub>2</sub>O, although superficially like it at normal temperatures and pressures



- let us ignore the complication that our bodies contain lots of  $H_2O$
- further, let's pretend that no one (on earth or twin earth) knows any chemistry (accomplished in Putnam's example by "rolling the time back to about 1750")

twin-Gene singing in XYZ  
on twin earth





“water is wet”



true just in  
case  $H_2O$  is  
wet

Oscar<sub>1</sub> (on earth)

“water is wet”



true just in  
case  $XYZ$  is  
wet

Oscar<sub>2</sub> (on twin earth)



So, the extension of 'water' as used by Oscar<sub>1</sub> is different from the extension of 'water' as used by Oscar<sub>2</sub>, despite the fact that they are "*in the same psychological state*"

"Cut the pie any way you like, "meanings" just ain't in the head!"





But doesn't Putnam's example also show that *thoughts* aren't in the head? Oscar<sub>1</sub> has beliefs about *water*, Oscar<sub>2</sub> has beliefs about *twater*



- it was soon realized that Putnam's example, if it shows anything at all, shows that some mental properties (like the property of believing that water is wet) are not intrinsic

# “Individualism and the mental”

- Putnam’s example arguably shows that differences in the subject’s environment (e.g. H<sub>2</sub>O vs. XYZ) can by themselves make a mental difference
- Burge’s examples purport to show that differences in the subject’s *linguistic community* can by themselves make a mental difference

# Burge's thought experiment



- **stage 1**
- Alfred has various beliefs about arthritis: that he has had arthritis for years, that stiffening joints is [are] a symptom of arthritis... (all true)  
and:
- that he has arthritis in his thigh (false, because arthritis is an inflammation of the joints)

# Burge's thought experiment



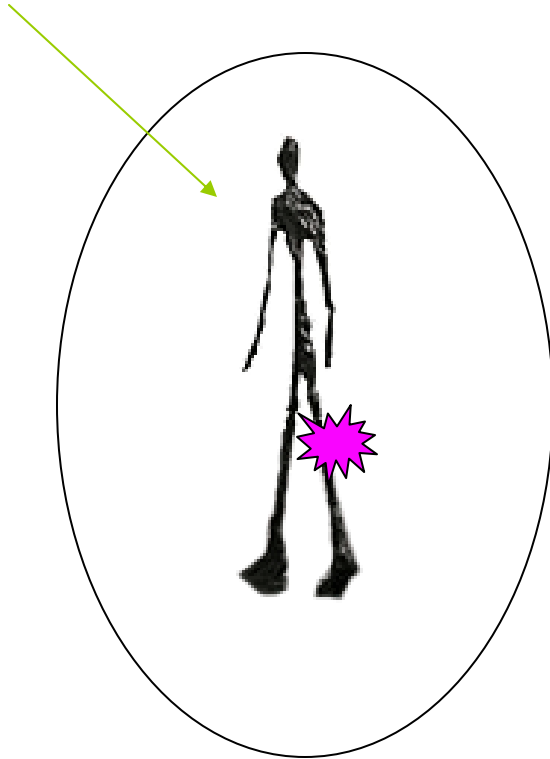
- **stage II**
- a “counterfactual situation” (a non-actual possible world) in which Alfred is exactly the same in all intrinsic respects, but lives in a slightly different linguistic community
- in this community, ‘arthritis’ applies “not only to arthritis, but to various other rheumatoid ailments”
- in the language of this community, ‘Alfred has arthritis in his thigh’ is true

# Burge's thought experiment



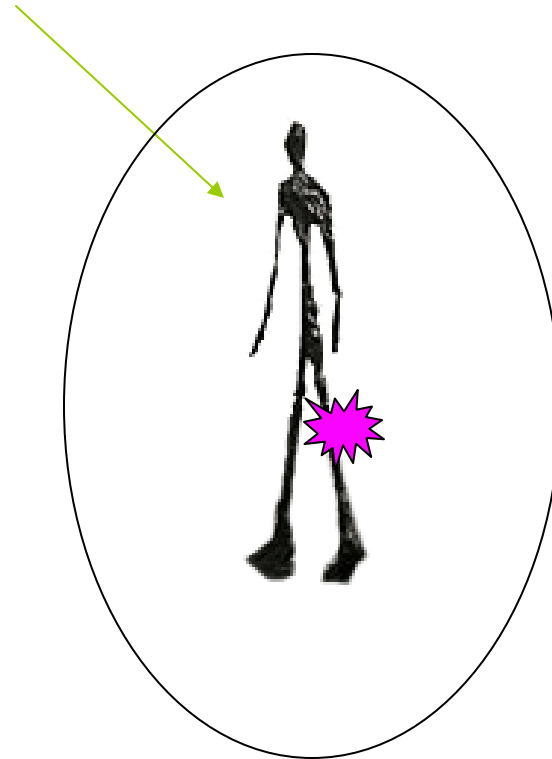
- **stage III**
  - an “interpretation of the counterfactual case”
  - Alfred has no beliefs about *arthritis* (in particular, he doesn't believe that he has arthritis in his thigh)
  - instead, he has beliefs about the sort of general rheumatoid ailment that is labelled in his community by the word ‘arthritis’

Alfred with arthritis beliefs



@ (the actual world)

Alfred (a duplicate of Alfred as he is in @) without arthritis beliefs



$w_1$  (the counterfactual situation)



# Minds and Machines

spring 2003

- read Brentano, Clark & Chalmers
- read Crane on internalism and externalism
- first paper draft due in recitation