Notes #1: Introduction **Phil. 4360**

Metaphysics: The branch of philosophy that deals with the general nature of reality: what are the most important things/sorts of things that exist, and (on a very abstract level) what are they like?

Ontology: Branch of metaphysics that deals with what exists.

Examples of things that might exist:

universals, physical things, minds/souls, God, numbers, sets, imaginary/fictional objects

The problem of universals:

• *Universal:* "That which may be predicated of many." Note 2 aspects: (1) a universal is predicable; (2) a universal can be shared. Universals are what multiple things can "have in common."

• *Particulars:* Things that aren't universals. The ultimate subjects of judgements/propositions. Particulars cannot be 'shared'; they can't be present in more than one place at a time.

• *The metaphysical questions:*

1. Do universals exist?
2. If not, why does it seem like they do? (E.g., why do we have words & concepts that apply to multiple things; why do things seem to 'have something in common'?)
3. If so, is their existence dependent on particulars, or independent of particulars?

• *Famous positions:*

  • *Nominalism:* There are no universals.

  • *Realism:* There are universals. Two versions of realism:

    a) *Transcendent (Platonic) realism:* And the universals are independent of particulars.

    b) *Immanent (Aristotelian) realism:* And the universals are dependent on particulars.

Plato's Theory of Forms:

• Extreme & bizarre form of Platonic realism. Incorporates the additional assumptions:

  • The 'self-predication' assumption: The form of F-ness is, itself, F.

  • The theory of recollection: Our knowledge of forms derives from acquaintance with them in a previous lifetime, before the soul was joined to the body. This knowledge was forgotten and needs to be recollected.

  • The inferiority of concrete reality: Plato thought concrete objects were mere pale imitations of, and somehow 'less real' than, concrete objects.

  • The form is *wholly present* in each particular instance.

The Parmenides:

• Plato's criticism of his own theory. Questions and arguments:

  • Is there a form of mud? Or hair? (No.)

  • If the form is in many separate particulars, then it is separate from itself.

  • Nor could it be in many particulars in the way a sail is spread over many particulars. For consider the form of 'largeness', 'equality', or 'smallness'.

• *The Third Man Argument:* (famous)

Suppose you have two men. They have something in common, so there is a form, the form of Man. But (by the self-predication assumption), this form is itself a third man. Hence, it has something in common with the first two men. Hence, there is another form . . . Leading to an

infinite series.

- Why Forms cannot be ideas/concepts:
  - "Is it an idea of something, or of nothing?" Something.
  - "And of something that exists?" Yes.
  - "And of some one thing that multiple things have in common?" Yes.
  - "And so you reintroduce the Form again."
  - Second argument: "Then you would be saying that everything has ideas in it."
- Why objects cannot resemble the Forms:
  If they do, then they are like the Forms, and the Forms are like them. If so, they and the Forms have something in common. And what will that be? The Form? (Another form? The same Form again?)
- Why we cannot know the Forms:
  Forms are only related to each other, not to things in our world; and vice versa. Ex.: A particular master is a master of a particular slave, not of Slavery (the form). Similarly, a particular person will have knowledge of particular things, not of forms.
- Why we can't reject universals either:
  If we do, our discourse loses significance. [This probably means that all the general terms in our language would have to be meaningless, since there would be nothing for them to refer to, since there is nothing that their instances have in common.]

Notes #2: Nominalism **Phil. 4360**

From last time: "universals". Nominalism vs. realism. Transcendent vs. immanent realism. Arguments for and against the "Forms".

Nominalism: Philosophical view which denies the existence of universals. Problem for nominalists:

- To explain how we group things together.
- Why there *appear* to be things in common between distinct particulars.

Forms of nominalism:

1. Predicate nominalism.
2. Concept nominalism.
3. Class nominalism.
4. Resemblance nominalism.
5. Tropes (fits with #4).
6. Ostrich nominalism (Quine).

Hume/Berkeley's nominalism:

- Hume's theory is concerned with 'abstract ideas': Ideas that (allegedly) represent *determinable* properties without representing any *determinates* falling under them.
- Hume held an imagistic conception of ideas: All 'perceptions' divide into (i) Impressions and (ii) Ideas. Ideas are faint 'copies' of impressions. Ideas are thus images in the mind.
- Why people believe in abstract ideas:
  The concept "man" applies to all men, regardless of size, shape, color, etc. Therefore, it must (a) represent all possible sizes, shapes, etc., or (b) represent no particular size, shape, etc. (a) is absurd, so it must be (b).
- Why you shouldn't believe in abstract ideas: Three arguments:

  1. (a) The mind cannot conceive separately of what is not in principle separable.
     (b) Determinates are not separable from their determinables.
     (c) So the mind cannot conceive of a determinable separately from its determinates.

  2. (a) No impression is abstract.
     (b) All ideas are just copies of impressions.
     (c) So no idea is abstract.

  3. (a) The mind cannot coherently conceive of the (logically) impossible.
     (b) An indeterminate object is (logically) impossible.
     (c) So the mind cannot conceive of such an object.
     (d) So the mind cannot form an abstract idea, for (?) such an idea would represent an indeterminate object.

- The alternative theory:
  - When we see a resemblance among objects, we apply the same name to them.
  - When we hear this word, it reminds us of some of the particular ideas of particular things the word applies to.
  - When we reason 'abstractly,' we are reasoning using a particular (non-abstract) idea, but we assert only things that also apply to the other ideas falling under the same category. We have a disposition to think of other cases when relevant.
  - A general term refers to multiple different particulars; it does not refer to an abstract item that the particulars share. "Ideas are particular in their nature, but general in their representation."
- Hume's theory seems to be a combination of:
  Resemblance nominalism
  Predicate nominalism
  Concept nominalism

Phil. 4360
Notes #3: Infinite Regress Arguments

The problem of universals according to Armstrong:

To explain what it is for something to have a property.

> [A better formulation would be that the *questions* about universals are: (1) Do universals exist? and (2) If not, why does it seem like they do, e.g., why do we say multiple different objects have the same property? —mh]

• Realists claim that there is some universal that multiple objects can all share.
• Nominalists give some other sort of analysis—attempting to explain away the appearance of universals, while appealing only to particulars.

Predicate Nominalism:

a is F = a falls under the predicate 'F'.

Problems:

1. The *object regress*:
   - The predicate-type 'F' appears to be a universal.
   - To analyze this away, the pred. nominalist will have to say that what it is for something to instantiate the type *predicate 'F'* would be for it to fall under another predicate (presumably the predicate "the predicate 'F'").
   - This leads to an infinite regress.
2. The *relation regress:*
   - What is *falling under*? It appears to be a universal.
   - To analyze this away, the pred. nominalist will have to say:
     x *falls under* y = <x,y> falls under the predicate 'falls under.'
   - This is either circular or leads to another regress.

Concept Nominalism:

a is F = a falls under the concept 'F'.

Problems: Same as above.

Class Nominalism:

a is F = a is a member of the set of F's.

Problems:

- No object regress, since the set of F's does not appear to be a universal.
- But still a relation regress, since 'being a member of' appears to be a universal.

Resemblance Nominalism:

a is F = a sufficiently resembles the paradigm F-things. For example:
a is red = a sufficiently resembles this fire engine, this apple, and this pool of blood.

Problems:

- Again no object regress, since the paradigm F things are particulars.
- But still a relation regress, since 'resemblance' appears to be a universal.

Generalization:

• No general analysis of 'what it is for a thing/things to have a property/relation' can succeed. For in order to say anything, the right hand side of the analysis must ascribe some properties and/or relations to something. The right hand side of the analysis cannot be just a list of particulars; then it would not even be a coherent statement. Thus, suppose:

a is F = $aR\Phi_F$

• A relation regress ensues with the predicate 'R'.
• Thus, *if the nominalist must accept Armstrong's project* (analyzing talk about properties in terms that only mention particulars), then the nominalist clearly loses.

Platonism:

a is F = a partakes in the form of F-ness.

Problem:

- Object regress: The form of F-ness is a form, so it participates in the form of Form-hood. Which is itself a form, so it either participates in itself or participates in a higher-order form of Second-Order-Formhood. If it participates in itself, you get a version of Russell's Paradox.
- Relation regress: 'partaking' seems to be a universal. So the analysis must hold:
  x partakes in y = <x,y> partakes in the form of Partaking.
  • This is either circular or introduces a higher-level partaking relation, leading to an infinite regress.

Phil. 4360
Notes #4: Armstrong's Realism

Armstrong's General View:

- Rejects nominalism & Platonism (see previous notes).
- The alternative is Aristotelian realism.
- Properties and particulars are interdependent: neither can exist without the other.
  - Corollary: the basic constituents of reality are 'thick particulars': objects-with-all-their-properties.
- Particulars are not *related to* their properties.

Objects are not bundles of properties:

- Assume properties are universals:
  - Then universals will be 'bundled' together when they stand in some relation. Call it the 'co-instantiation' relation.
  - If an object has P and Q, then P and Q (the universals) stand in this relation.
  - If an object has P but lacks Q, then P and Q (the universals) do not stand in this relation.
  - Thus, if there are two objects, and one has P and Q, while the other has P but not Q, then the universals P and Q both do and do not stand in the 'co-instantiation' relation.
- The trope view of properties doesn't work because it is a form of nominalism, which doesn't work. (See previous notes.)

Armstrong's Scientific Realism:

- What properties exist is an empirical, not a priori question. Leave it to the scientists.
- There are no uninstantiated properties.
- Rejects arguments from meaning:
  - "In order for a predicate to be meaningful, there must be a property for it to stand for." (The Empiricist "will regret that at present we seem to lack any satisfactory theory of meaning." —p. 149)
  - Even instantiated predicates need not correspond to single properties. (Could be, e.g., family resemblances.)
  - Non-synonymous expressions may denote the same property.

Other Alleged Kinds of Properties:

- There are no disjunctive properties.
  For suppose *a* has F but lacks G, while *b* has G but lacks F. It seems laughable to conclude from these premises that *a* and *b* have something in common. Yet the disjunctive predicate "F or G" applies to both.
- There are no negative properties.
  - We don't need negative properties to explain how negative predicates apply to things. "~P" applies to a thing when it isn't P.
  - Use Occam's Razor.
- But conjunctive properties are ok.
  - Because there might be nothing but conjunctive properties.

Phil. 4360
Notes #5: Quine on Ontological Commitment

From last time: Main points of Armstrong's scientific realism. Does he have good arguments against nominalism & Platonism? How much of his 'scientific realism' is supported by these or other arguments, and how much is unsupported?

I. Problems of ontology: Arguments for the existence of Pegasus

- Question: What is there?
- McX's theory:
  - Every meaningful noun corresponds to something (that exists).
  - It cannot be that Pegasus doesn't exist, for: (a) if so, "Pegasus" is meaningless, (b) and if so, then "Pegasus doesn't exist" is meaningless. Hence, Pegasus exists.
  - Quine cannot even say how he disagrees with me (McX), for then he would have to make mention of the entities he wants to deny exist.
  - Since there isn't a physical, flesh-and-blood winged horse, uh, Pegasus "exists in my mind," i.e., Pegasus is an "idea". Problem: confuses an idea of an X with an X.
  - McX is a fictional character (so he doesn't exist, right?) However, compare Meinong's famous doctrine that "There are some things of which it is true to say, that there are no such things."
- Wyman's theory:
  - Some things "exist" while others only "subsist." Existence = actuality. Mere subsistence = mere possibility.
  - *Existing* things just have an additional property ("actuality") that the merely subsistent entities lack.
  - *Problem*: How many possible fat men in the doorway are there? How many of them are bald?
  - Note: Is Wyman David Lewis?
  - *Problem*: What of the Round Square Cupola on Berkeley College (an impossible object)? *Answer*: "round square cupola" is meaningless. *Further problem:* This leads to the rejection of the method of reductio ad absurdum.

II. A Russellian solution:

- Definite description: A noun phrase beginning with "the" that purports to refer to a unique thing.
  - Ex.: "the author of the *Waverly* novels", "the present King of France."
  - Notice that they function like nouns (or like names) in a sentence: "The King of France is bald" just like "George is bald."
- Russell's analysis of definite descriptions:
  "The King of France is bald" = $(\exists x)\,[Kxf \,\&\, Bx \,\&\, (y)(Kyf \supset y=x)]$
  or: $(\exists x)\,[Kxf \,\&\, (y)(Kyf \supset x=y)] \,\&\, (x)(Kxf \supset Bx)$
  Read: There is something such that: it is a king of France, it is bald, and all kings of France are identical to it. Or: "There is at least one king of France, there is at most one king of France, and every king of France is bald."
  - *Note:* "the king of France" *is not a singular term*. It therefore is meaningful even if there is

no King of France. It only requires there to be a *predicate*, "being the king of". Thus:
"The King of France doesn't exist" = $\sim(\exists x)\,[Kxf \,\&\, (y)(Kyf \supset y=x)]$

- How to deal with "Pegasus": Introduce a predicate "pegasizing" (perhaps defined as being a horse w/ wings?). Thus:
"Pegasus does not exist" = $\sim(\exists x)\,Px$

III. On the problem of universals:

- For universals: the one over many argument:
  1. *a* is red.
  2. *b* is red.
  3. Therefore, there is something that *a* and *b* have in common, viz. 'redness'. Thus, there is at least one universal.

  • *Reply:* Quine rejects the inference from 1+2 to 3. He agrees with *predicative* uses of "red" but refuses to accept the *noun* "redness".
- The argument from meaning:
  In order for "red" to be meaningful, it must have a meaning. Its meaning appears to be a nonphysical entity which is very much like a universal; perhaps it *is* a universal.
  • *Reply:* Quine rejects the existence of 'meanings': he accepts only the *predicative* uses of "is meaningful" or "is synonymous with", but not the *noun* "meaning".
- McX: "What? Does *nothing* you can say commit you to the existence of universals?!"
  Answer:
  Quine's Criterion of Ontological Commitment (important):

  > "To be assumed as an entity is ... to be reckoned as the value of a variable." (p. 13)
  >
  > Or: S is committed to the existence of F's iff S makes a statement in which the bound variables must range over at least one F in order for the statement to be true.

  • Examples:
  "Some dogs are white"
  "There are infinitely many prime numbers"
  "There is no Pegasus"
  • *Note:* This is not a criterion of *what exists* but of what a statement *says* exists.
- How to evaluate an ontology:
  • We should choose the simplest conceptual scheme that allows experience to be described and organized.
  • Q. hints that physical objects are a convenient "myth", just like mathematical objects, just useful for predicting sense data.

Phil. 4360
Notes #6: Ostrich Nominalism vs. Mirage Realism

I. Devitt:

The One-over-Many Problem:

- How can many things have a common nature? For instance, what makes (1) true?
  (1) *a* and *b* have the same property, F-ness.

This is a pseudo-problem:

- Answer: (1) is explained by:
  (3) *a* is F. = Fa
  (4) *b* is F. = Fb
- There is no further need to explain (3) or (4).

Realism is a pseudo-explanation:

- Suppose (1) is explained by
  (5) *a* has the universal, F-ness. = Raf
- (5) should need to be explained as well, leading to an infinite regress.
- Armstrong's claim that *a* and the universal F-ness are not related but some sort of organic unity is just obscure. How can we understand (5) unless it is of the form Raf?

Some real problems:

- The following statements may imply the existence of universals:
  (6) Red resembles orange more than it resembles blue.
  (7) Red is a color.
  (9) The dresses were of the same color.
- Quine proposes to treat them in terms of sets or open sentences.

II. Armstrong:

The real one-over-many problem:

- We commonly say distinct particulars are 'of the same type' or 'have something in common.' What does this mean?
- The resemblance-nominalist answer: It is for them to have resembling properties.
- The realist approach: Takes the idea literally: there is something literally identical between distinct particulars. E.g., the color of a = the color of b.
- This is the question the Ostrich nominalist refuses to answer. (The question is not what (3) means.)

Applying Quine's criterion of ontological commitment:

- Quine fails to explain how he would account for sentences (6), (7), and (9) above.
- He considers:
  (10) Humility is a virtue.
  (10') Humble persons are virtuous.

- Problem: (10) is neither necessary nor sufficient for (10'). For:
  - Humility could be a virtue, but there could still be plenty of unvirtuous humble people.
  - If (10') implies (10), then (11') implies (11):
    11'. Tall persons are virtuous.
    11. Tallness is a virtue.
- Anyway, there is no reason to accept Q's criterion of ontological commitment.
  - Why not say use of predicates also involves ontological commitment?
  - Unclear how Quine should explain what "applies to" means without referring to properties.
    The semantic theory proposed by Devitt says:
    'Fa' is true iff there is some thing that 'a' refers to, and 'F' applies to that thing.
    To explain the meaning of 'applies to,' fill in the ellipsis:
    $(x)(y)\ (Axy \leftrightarrow ...)$
    For the realist:
    $(x)(y)\ [Axy \leftrightarrow (\exists z)(z \text{ is a property, x refers to z, and y has z})]$

Problems for the Realist:

- "How are the two components [the substance and its properties] of a particular to be put together?" Two views:
  - Relational view — leads to infinite regress.
    Fa = I(a,F)
    = I(<a,f>, I)
    = I(<<a,f>,I>,I)
    ⋮
  - Non-relational view — obscure but seems required by failure of the relational theory.
  - Alternative: To hold that the regress exists but "isn't vicious." Compare:

The Instantiation Regress	The Truth Regress
Fa = I(a, F) = I(<a,f>, I) = I(<<a,f>,I>, I) ⋮	p = p is true = 'p is true' is true = "'p is true' is true" is true ⋮
Armstrong says we should reject the first equivalence.	*Should we reject the first equivalence here?*

  - When is a regress 'vicious' (bad)? When it implies an *impossible-to-satisfy* condition for a *true* statement, as:
    a. When succeeding steps are preconditions for the earlier steps, and the series can't be completed.
    b. When succeeding steps lead to more complex conditions, and there is a limit to the complexity of the condition that can be satisfied.

Phil. 4360
Notes #7: Criticism of Quine's Criterion

What sort of dispute is that between the nominalist and the realist?

- It is not a factual dispute, for:
  - Both agree on the underlying facts. Both agree that Sam is bald.
  - The realist deduces from this that baldness exists, but the nominalist does not.
  - The realist says baldness would exist even if Sam were not bald:

Bs.	~Bs
$\therefore (\exists\phi)\ \phi s$	$\therefore (\exists\phi)\ \sim\phi s$

  - Hence, the existence of baldness follows from the mere existence of a meaningful predicate, "is bald."
  - Hence, the existence of baldness is nothing more than the meaningfulness of "bald." So a nominalist should not object to it.
- The dispute is semantic. If two philosophers agree on sentence "P", and one infers "Q" but the other does not, then either:
  a. They mean something different by "P", or
  b. They mean something different by "Q".
  - "There are no other possibilities." (106) [Why not the possibility that one of them makes a mistaken inference or fails to make a correct inference?]
- *Note:* If P is a tautology, then if Q follows from P, Q is also a tautology.

Criteria of Ontological Commitment:

Criterion 1: S is ontologically committed to F's iff S 'quantifies over' F's.
Criterion 3: S is ontologically committed to F's iff S says or implies that F's exist.

- Q: We can eliminate ontological commitments by paraphrasing statements into a notation which does not quantify over the unwanted entities. Ex.: "There are 4 miles between N and T" can be rewritten, "Distance in miles between N and T = 4," showing that we need not accept the existence of 'miles.'
- A: There are two possibilities:
  1. The paraphrase does not preserve the meaning of the original statement. In this case, it is an invalid paraphrase and does not indicate the true ontological commitments of the original sentence.
  2. The paraphrase has the same meaning as the original statement. In this case, the first cannot entail anything that the second does not. Thus, if the first entails the existence of miles, the second must do so also.

  Either way, the paraphrasing procedure accomplishes nothing.

- Four interpretations of Quine's criterion: Sentence S implies the existence of F's iff:
  1a. S (as written) quantifies over F's.
  1b. There is *some* paraphrase of S which quantifies over F's.
  1c. *Every* paraphrase of S quantifies over F's.
  1d. "The correct" paraphrase of S quantifies over F's.

- 1a is wrong. For it implies that synonymous sentences can have different ontological commitments.
- 1b is wrong. For it implies that ordinary statements contain commitments to universals.
- 1c is wrong. For it implies that no sentence has any ontological commitments.
- 1d seems to be wrong. For what is the "correct" paraphrase? How would that be determined?

- An illustration of the problem with 1c.
  - Let K = the conjunction into a single proposition of all human knowledge.
  - Let Px mean, "K, and x = this pen."
  - Consider the statement: $(\exists x)$ Px. This is a paraphrase of all existing human knowledge, and it quantifies over only this pen. (Only this pen need be in the domain of the variable "x" in order for the sentence to be true.)
  - Thus, 1c implies that all existing human knowledge is committed only to the existence of this pen.

Phil. 4360
Notes #8: Numbers

Q: What are numbers, and do they exist?

I. A proof that 2 exists.

1. 2 > 1. (premise)
2. "2 > 1" is a sentence of the form 'Rab.' (premise)
3. The truth-conditions for a sentence of the form 'Rab' are that there be something that 'a' refers to, something that 'b' refers to, and that those things should stand in the relation denoted by "R". (premise)
4. If "2" refers to something, it refers to 2. (premise)
5. "2 > 1" is true. (from 1)
6. The truth-conditions for "2 > 1" are that there be something that "2" refers to, something that "1" refers to, and that those things stand in the relation denoted by ">". (from 2, 3)
7. There is something that "2" refers to. (from 5, 6)
8. There is (there exists) 2. (from 4, 7)

• Premises 1-4 are all intuitively obvious. #1 is the least subject to doubt; even 3-year-old children know that.
• The argument does not establish the nature of numbers, apart from the fact that they exist and can be greater than one another.

II. Are numbers properties of external things?

Objections:

- What is the number that belongs to this deck of cards? It is 1 (deck), 52 (cards), or 10^{23} (molecules). If numbers are properties, how can one thing have these different, seemingly contrary properties?
- What is 0 a property of?
- What is the object that the number belongs to? (The 'agglomeration.') What about in the case of "there are 4 ways to prove this theorem"?
- It's surprising that one property could apply to so many different kinds of things.

III. Are numbers subjective?

- The number of petals of a flower is just as factual as its color.
- "The objectivity of the North Sea is not affected by the fact that it is a matter of our arbitrary choice which part of all the water on the earth's surface we mark off and elect to call the 'North Sea.'"
- Arithmetic isn't psychology.
- If numbers were ideas, there could be conscious and unconscious numbers.
- There wouldn't be infinitely many numbers.

IV. Frege's theory of numbers

1. A number is a property of a *concept* (universal).
   - "Venus has 0 moons": 0 belongs to the concept 'moon of Venus.'
   - *Objection*: The number of inhabitants of Germany changes from one year to the next, but the concept 'inhabitant of Germany' suffers no alteration.
     Reply: The following are two different concepts: 'inhabitant of Germany in 1883' and 'inhabitant of Germany in 1884'.
   - Concepts are objective. (***Fregean 'concepts' are not ideas.***) Compare the sentence: "All whales are mammals."
     • This is an objective fact.
     • It is an assertion about concepts. 'Whale' is not the name of any particular.
2. The identity conditions for numbers: We must define "the number which belongs to F is the same as that which belongs to G."
   - *Equinumerate:*
     The concept F is equinumerate with the concept G $=_{df}$ It is possible to correlate the objects which fall under F one-to-one with the objects which fall under G. [*Or*: there is a one-one function from the extension of F onto the extension of G.]
   - *Extension of a concept:* The set of all objects that fall under it.
   - The number that belongs to the concept F $=_{df}$ The extension of the concept, 'equinumerate to the concept F.' [*Or*: is the set of all concepts that are equinumerate with F.]
3. Definitions of individual numbers:
   0 = the number which belongs to the concept 'not identical with itself.' [*Note:* this is the set of all uninstantiated universals.]
   1 = the number which belongs to the concept 'identical with 0.' [*Note:* this amounts to the set of all singly-instantiated universals.]
   and so on.

Historical note: Less Platonic philosophers (Russell) were later to remove the talk of concepts in favor of sets only.

Phil. 4360
Notes #9: Review, unit 1

<u>Know what these things are:</u>
Universals vs. particulars
Realism
  Platonic, Aristotelian
Nominalism
  Predicate, concept, class, resemblance
  Tropes
  Ostrich
the self-predication assumption
the Third Man argument

<u>Know what these people thought, esp. about these ideas:</u>
Plato
Hume. Including:
  'abstract ideas'
  ideas & impressions
Armstrong.
  The 'problem' of universals
  The object regress
  The relation regress. Incl.: the general argument against relational theories. What this applies to.
  Points in his 'scientific realism': empiricism; disjunctive, conjunctive, negative properties; uninstantiated properties.
Quine.
  McX's argument for Pegasus
  Russell's analysis of "the K of F is bald"
  Quine's answer to McX re: Pegasus
  Criterion of ontological commitment & how it applies to sentences
Devitt:
  On the one-over-many problem
  Armstrong's response.
    The "Humility is a virtue." example.
Searle:
  • On semantic disputes & tautologies.
  • The Alston argument for why the paraphrase does not succeed in removing ontological commitments
  The example of the commitment to "this pen."
Frege
  Concepts.
  The argument from the deck of cards & what it shows.
  Problems with 0.
  Frege's definitions:
    equinumerate
    extension of a concept
    the number which belongs to a concept
    0, 1, etc.

Phil. 4360
Notes #10: Paradoxes of Infinity

Important philosophical points about infinity:

- Infinity exists in nature:
  - Space, time are infinite *in extent* and infinitely *divisible*.
  - The infinite series of numbers (if numbers exist).
- Old (Aristotelian) views of infinity:
  - 'Infinity' is not a number/determinate quantity. 'Infinite' just means 'without limit.'
  - There are only *potential* infinities, not *actual* infinities.

Galileo on infinity:

- Lines & points:
  - A line is made up of points. Points are indivisible bits of space. Lines are divisible.
  - But: A divisible quantity cannot be made up of indivisible ones. For
    - If 2 points together made up a divisible quantity (a line?), then 3 points, 5 points, etc. would also be divisible.
    - In that case, we could divide the line in two. But then a point would be divided in two (since there are an odd number), which is impossible.
    - Answer: a 'divisible' quantity [I think he means a line, which is *infinitely* divisible] requires an *infinite* number of indivisibles to make it up.

- Can one infinity be greater than another?
  - Yes, for a line can be longer than another line.
  - No, for consider:

    Let S = the number of perfect squares (1, 4, 9, 16, ...)

    N = the number of natural numbers (1, 2, 3, 4, ...)

    $S > N$, $S = N$, or $S < N$?
    - $S < N$, for: N includes S, plus many other numbers. As *n* increases, the proportion of the first *n* natural numbers that are perfect squares decreases, approaching 0. Thus, there are *infinitely* many more N's than S's.
    - But $S \geq N$, for: every natural number has a (distinct) square; hence, there are at least as many squares as there are natural numbers.
    - Galileo's conclusion: infinities are incommensurable (cannot be compared in terms of greater, lesser, equal). Further implication: 'infinity' is not a determinate quantity; there are no such numbers as S and N.
    - Answer to above argument: Neither line contains 'more', 'less', nor 'equally many' points, since both contain infinitely many.
    - Important lesson: We cannot understand infinity by applying to it the same properties as apply to finite numbers.

- How many finite parts does a line segment (of finite length) contain?
  - Infinitely many? No, for then it would be infinitely long.
  - Finitely many? No, for then there would be a limit to how many times you could divide the line.
  - Answer: neither; rather, the parts "correspond to every assigned number." [I.e., for every number

n, the line segment contains exactly *n* parts of some length *x*.]

Zeno's Paradox

• You drop a ball. The ball cannot reach the ground, for:

1. In order to reach the ground, it must complete the series, (1/2, 3/4, 7/8, ...)
2. The series, (1/2, 3/4, 7/8, ...), is a series with no end. (It is infinite.)
3. A series with no end cannot be completed.
4. The series, (1/2, 3/4, 7/8, ...), cannot be completed. (from 2,3)
5. The ball cannot reach the ground. (from 1,4)

A math paradox:

• What is the sum of the following sequence of numbers:
1, -1, 2, -2, 3, -3, ...?

• It is 0, for:

$$\begin{aligned} &1 + -1 + 2 + -2 + 3 + -3 + ... \\ &= (1 + -1) + (2 + -2) + (3 + -3) + ... \\ &= 0 + 0 + 0 + ... \\ &= 0. \end{aligned}$$

• It is +∞, for:

$$\begin{aligned} &1 + -1 + 2 + -2 + 3 + -3 + 4 + ... \\ &= 1 + (-1 + 2) + (-2 + 3) + (-3 + 4) + ... \\ &= 1 + 1 + 1 + 1 + ... \\ &= +\infty \end{aligned}$$

• It is -∞, for:

$$\begin{aligned} &1 + -1 + 2 + -2 + 3 + -3 + ... \\ &= -1 + 1 + -2 + 2 + -3 + 3 + -4 + ... \\ &= -1 + (1 + -2) + (2 + -3) + (3 + -4) + ... \\ &= -1 + -1 + -1 + -1 + ... \\ &= -\infty \end{aligned}$$

• Conclusion: There is no such sum. (Or: commutativity and associativity do not apply to 'infinitely long' sums.)

The Thomson's lamp paradox:

• Thomson's lamp starts out on. After ½ second, it switches off. After another 1/4 second, it switches back on. Etc. For every *n*, after $(1/2)^n$ seconds, the lamp switches again. No other changes occur.

• At the end of 1 second, is the lamp on or off?

Phil. 4360
Notes #11: Modern conception of infinity

Modern conception of numbers:

- This is based on Frege's conception, but with 'concepts' replaced with sets.
- Understand the notion of a one-one function.
- Important technical notions: These are all 2-place relations between sets (see text):
  - 'has the same cardinality as'
  - 'has a higher cardinality than'
  - 'has a lower cardinality than'
- The noun usage of numerals:
  0 = The set of all sets with the same cardinality as {}
  1 = The set of all sets with the same cardinality as {0}
  2 = The set of all sets with the same cardinality as {0,1}
  etc.
- The adjectival use of number terms: There are *n* F's = the set of F's is an element of *n*.
- Every set has a number that 'belongs to' it.

The first 'transfinite' number: ω

- ω is the number which belongs to the set of all natural numbers, {0, 1, 2, ...}, the set of all sets with the same cardinality as {0, 1, 2, ...}.
- Other sets that are members of ω:
  {1, 2, 3, ...}
  {2, 4, 6, ...}
  {1, 2, 3, 5, ...} (set of prime numbers)
  {1, 4, 9, 16, ...} (set of perfect squares)
- An infinite set can be mapped one-one onto a *proper subset* of itself. This is not true of any finite set.

The next (?) transfinite number: c

- *c* is the cardinal number that belongs to the set of all real numbers, or (what is the same) the set of all real numbers between 0 and 1.
- *c* has a higher cardinality than ω. This is shown by the diagonalization argument (see text).
- By the way, there is an infinite hierarchy of ever larger transfinite cardinals.

Philosophically:

- There is a clash between at least two conceptions of 'infinity':
  1. The Aristotelian/Galilean conception. Infinity is not a number; there are 'potential infinities' but not 'actual infinities.' "Greater than" doesn't apply to infinite sets.
  2. The modern, Cantor conception. There are many infinite numbers (as described above). Numbers are equivalence classes of sets with the same cardinality. "Greater than" means "has a higher cardinality than" as described above.
- Cantor did not 'prove' his conception of infinity; nor did anyone else. He assumed it. (This is the usual method in mathematics.)
- Advantages of Aristotelian/Galilean conception:

  (a) The 'transfinite numbers' are odd anyway--normal rules don't apply to them. This supports that they aren't genuine numbers. E.g., the rule "$x + 1 > x$", "$x - x = 0$", etc.

  (b) Allegedly removes the paradoxes of infinity.
  1. Zeno's paradox.
  2. The paradox of the infinite sums.
  3. Galileo's 'paradox'.
  4. Thomson's lamp paradox.
  5. The banker paradox.
- Problems with the Aristotelian conception of infinity: Aren't all of these 'actual infinities':
  - Infinite divisibility of space & time.
  - Infinite extent of space & time.
  - The infinite series of natural numbers.
- Now, compare to the following possible (?) kinds of infinity; is there a difference?
  - An 'infinite force'. (What happens when the infinite force meets the infinite mass?)
  - An object with an 'infinite velocity', or an infinite rate of change.
  - An infinitely large (massive, etc.) universe. An infinite number of particles.
  - A particle of infinite mass / infinite density in a given region.

Phil. 4360
Notes #12: Lewis on Possible Worlds

Review from last time: 2 conceptions of infinity. Advantages/disadvantages of Aristotelian conception. Different kinds of (alleged) infinities: an infinite temporal series; an infinite extensive, physical magnitude; an infinite intensive, physical magnitude; infinite abstract objects.

Lewis' craziness:

- There are 'possible worlds.'
- Other possible worlds are worlds, in exactly the same sense that the actual world is one.
- They exist, in exactly the same sense that the actual world exists.
- They each have their own spacetime, disconnected from our own spacetime. Like parallel universes, with no possible way to travel between them & no spatiotemporal relations between worlds.
- 'Actual' just means "pertaining to the world that I'm in."

An initial argument for the existence of 'possible worlds':

1. Some modal statements are true. (e.g., "I could have had a V8.")
2. Modal statements are best interpreted as assertions about possible worlds, as indicated below:
   "It is possible that p" = "In some possible world, p."
   "It is necessary that p" = "In every possible world, p."
   "It is impossible that p" = "In no possible world does p hold."
3. Therefore, possible worlds exist.

Alternative views:

1. Modal expressions are unanalyzable.
   Objection: "This is not an alternative theory at all, but an abstinence from theorizing."
2. 'Possibly, P' = ' "P" is a consistent sentence.'
   Objection: What does "consistent" mean?
   a. "consistent" means "could be true." Then the theory is circular.
   b. "consistent" means "whose denial cannot be derived from some formal system." Problem: From Godel's theorem, for any (consistent) formal system, there are truths of arithmetic that cannot be derived from it. The negation of such a sentence is therefore 'consistent' according to (b). But the negation of a truth of arithmetic is not possible.
3. 'Ersatz possible worlds': there are 'possible worlds', but they're really just sets of sentences.
   Objection: this will run into the same problem as (2).

Objections to possible worlds:

1. Only our own world actually exists.
2. Realism about p.w.'s is unparsimonious: there are too many entities in your theory.
   Reply:
   • Distinguish (a) qualitative simplicity: reduction in the number of *kinds* of things in a theory, (b) quantitative simplicity: reduction in the number of *instances* of a given kind.
   • My (Lewis) theory has qualitative simplicity.
   • Qualitative simplicity is all that matters.

3. Quine says possible objects are hard to individuate. Not so on my (Lewis') theory, since they are just exactly the same sorts of objects as the objects in our world, except that they happen to be in other worlds. Each possible object occupies only its own world.
4. (Not really an objection.) Tell us more about p.w.'s. How many are there? Are there multiple qualitatively indistinguishable p.w.'s?
   Reply: I don't know, and I don't know any way to find out.

Methodological/epistemological points:

- We start out with pretheoretical 'opinions.'
- Philosophy should seek a *systematic theory* that *respects (or explains the truth of?) those opinions*.
- Also, we seek qualitatively simpler theories, ceteris paribus.

Phil. 4360
Notes #13: Kripke on Possible Worlds

• Review Lewis' theory, Lewis' methodology. Does Lewis make a correct application of that methodology?

• From Lewis theory: the problem of transworld identity:
What are the conditions for 'identifying' an object across possible worlds? What makes the 'me' in the other possible world count as (a counterpart of) me?

• Instead of "possible worlds": Alternative ways of speaking:
• talk of "possible states/histories of the world"
• talk of "counterfactual situations"
• Replace with modal talk: "It is possible that . . ."

• The dice example:
• You have two dice, die A and die B. Each has 6 faces, so 6 possible ways of landing.
• Throwing the dice yields 36 possible results. These are 'possible worlds.'
• The 'actual world' is the result that actually occurs.
• It should not be confused with the physical object, the pair of dice itself. The 'possible states' are all possible states of *that* thing. There is only one pair of dice (not 36, with the other 35 existing in another dimension).
• Nor does this imply the existence of a 'bare particular.'
• There are no counterparts, nor is there a problem of transworld identity. The 'me' in the other 'possible world' is me because I stipulate that I'm talking about me.

Essential vs. accidental properties:
- Accidental property: A property a thing has but could have failed to have; a property that a thing has in the actual world but lacks in some possible worlds.
- Essential property: A property a thing must have; a property a thing has in every possible world in which it exists.

• Quine's objection to the meaningfulness of this distinction: Consider the two sentences,
a. "The President of the United States might not have been President of the United States."
b. "George W. Bush might not have been President of the United States."
• (a) is false but (b) is true. But "The President of the United States" and "George Bush" are just two descriptions of the same person. Thus, the property of being President of the United States can appear 'accidental' or 'essential', depending on which way you describe that person.
• Response: the distinction between rigid and non-rigid designators.
• *Rigid designator:* A term that refers to the same entity in (every description of) every possible world (in which it refers to anything). "George W. Bush" is a rigid designator.
• *Nonrigid designator:* A term that designates different things in (descriptions of) different possible worlds. "The President of the United States" is a non-rigid designator.
• Notice that the question, "Could Nixon have not been a human being" is metaphysical, not epistemological: We know he *was* a human being, in actual fact.

Bundles of properties vs. 'bare particulars':

- A particular is not a bundle of qualities. If a quality is an abstract object, so is a bundle of qualities.
- Nor are particulars things *without* qualities that 'stand behind' qualities.
- Instead, particulars are things that have qualities.

Phil. 4360
Notes #14: Causation

Regularity theories of causation:

- Derive from Hume's first definition of causation, take the idea that causation is just constant conjunction, or the instantiation of regularities in nature.
- State something like this: c causes e when
  - (i) C and E are true, where C is a proposition stating the occurrence of c, E is a prop. stating that e occurs.
  - (ii) L & F imply (C ⊃ E), where L is a statement of the laws, and F is some statement of initial conditions (or 'particular facts').
  - (iii) L & F don't imply E, nor does F imply (C ⊃ E).
- Problems:
  1. Don't differentiate cause from effect; c might instead be an effect of e.
  2. c and e might instead have a common cause. (E.g., d causes both c and e.)
  3. c might be a preempted potential cause of e. (c would have caused e, but d caused it first.)

Counter-factual analysis of causation:

- Similarity of worlds: w_1 is closer to *a* than w_2 is, when w_1 is more similar to *a*.
  - This similarity is a 'primitive' relation.
  - Similarities may be in two respects: (a) with respect to laws, (b) with respect to particular facts.
  - Generally, laws are more important than particular facts.
  - However, large regions of *exact* similarity of particular facts are specially important.
- Counterfactuals (Stalnaker-Lewis analysis):
  A □→ C = In the nearest world(s) in which A holds, C holds.
  Or: Some world in which A and C hold is closer to the actual world than any world in which A and ~C hold.
- Causal dependence:
  e depends causally on *c* = O(c) □→ O(e) and ~O(c) □→ ~O(e).
- Causation:
  - Casual dependence implies causation, but causation does not imply causal dependence. Because causation is transitive, but causal dependence is not.
  - Causation is *the ancestral* of causal dependence. *c* is a cause of *e* iff there is a causal chain from c to e.
    - Good concept to know: If R is a non-transitive or intransitive 2-place relation, the ancestral of R is the relation that holds between *a* and *b* whenever there exists a finite series of objects, beginning with *a* and ending with *b*, such that each successive pair is related by R.
- Nomic dependence: ignore this.
- The problems for the regularity theories:

1. The problem of effects
   - Assume that *c* causes *e* and *e* doesn't cause *c*. Assume *c* had to cause *e*, given the circumstances and the laws. Then if *e* hadn't occurred, it would have to have been because *c* hadn't occurred. Thus, if *e* didn't occur, *c* wouldn't have occurred.
   - Solution: No, if *e* didn't occur, *c* still would have occurred, but would have failed to cause

e. The circumstances or the laws would have been different.

• Why: to maximize the region of spatiotemporal match with the actual world, we choose a possible world in which everything is exactly the same up to the time of *e*, and then a 'divergence miracle' occurs (a violation of the actual world's laws of nature), whereby *e* fails to happen. Then we continue the laws forward from there.

[• Problem: we could also introduce a 'convergence miracle' right after the time of *e*, to bring the course of the world back to the course in the actual world. Lewis presupposes a spurious asymmetry between the past and the future.]

2. The problem of 'epiphenomena.' [Actually the problem of common causes.]
   - Assume *c* causes both *e* and *f*, *e* doesn't cause *f*, and *e* happens before *f*. Then if *e* didn't happen, *f* wouldn't have happened.
   - Solution: No, as above: if *e* didn't happen, then *c* and *f* would still have happened, but *c* would have failed to cause *e*.
3. The problem of preemption.
   - Suppose c_1 causes *e*, but another event, c_2, was standing by, ready to cause *e* if c_1 hadn't been there. Then $\sim O(c_1) \boxrightarrow \sim O(e)$ fails.
   - Solution: So there is no causal dependence of c_1 on *e*, but c_1 still causes *e*, because:
     • There is some chain from c_1 to *e*--e.g., $\sim O(c_1) \boxrightarrow \sim O(d) \boxrightarrow \sim O(e)$.

Phil. 4360
Notes #15: Events

Davidson on events:

- A distinction: *particular* events vs. event-*types*.
  - *Ex.:* A particular event: the War of 1812. An event-type: War.
  - Note that Davidson is talking (only) about *particular events*. Says we don't need event-types to explain recurrence. (Doing 'the same thing' again just means doing something similar.)
- Argument for the existence of events: The following inference is valid:
  1. John ran quickly.
  2. Therefore, John ran.

  What is the form of this inference?
  1. (Ex)(Rx & jBx & Qx) (There was a running, by John, that was quick.)
  2. (Ex)(Rx & jBx) (There was a running, by John.)

  Not:
  1. Qj
  2. Rj

  Notice that the form of the inference is the same as that of:
  1. There was a grey cat on the mat.
  2. Therefore, there was a cat on the mat.
- Ordinary language makes many apparent references to events. "There were many wars in the 20th century." Etc.
- How to individuate events:
  x is the same event as *y* iff: *x* and *y* have all the same causes and effects.

Events according to Kim:

- Kim uses "event" broadly: Includes both *changes* and *states* (or 'unchanges').
- Events are property-exemplifications. Each event has three things:
  - A constitutive object. (Roughly, the object that is doing something, or being some way.)
  - A constitutive property. (What the object is doing or being.)
  - A time of occurrence.
- Two conditions for defining the concept of an event:
  Existence condition: The event [x,P,t] exists (occurs) iff x has P at t.
  Identity condition: [x,P,t] = [y,Q,t'] iff: (x=y, P=Q, and t=t').

  That explains what an 'event' is. There can also be more complex events with ordered n-tuples for constitutive objects, and relational constitutive properties [and time intervals of occurrence?].
- Misc. issues:
  - What sort of constitutive properties are allowed? (Not just any, for then 2+2 equalling 4 would be an event.)
  - When John ran quickly, is the constitutive property "ran" (with quickness a property of the event), or "ran quickly"? Note: must distinguish the event's *constitutive property* from a property *of* the event.
  - Kim's account of events is compatible with Davidson's analysis of event-sentences--Kim events may be just what are quantified over in the example sentences (1) and (2) above.

• Suppose Mary kissed two admirers at once. How many kissings occurred? Answer: at least 4 (Kim says 3, but this is a mistake), 1 monadic kissing, 2 dyadic kissings, and 1 triadic kissing, namely:

1. [Mary, ① kisses an admirer, t]
2. [<Mary, Steve>, ① kisses ②, t]
3. [<Mary, Larry>, ① kisses ②, t]
4. [<Mary, Steve, Larry>, ① kisses ② and ③, t]

• The role of causation: Some events include other events plus their effects: ex.: Brutus' killing of Caesar includes Brutus' stabbing of Caesar and the effect, Caesar's death.

The Individuation of Events (Kim vs. Davidson):

- Famous example: Brutus stabbed Caesar. Let's say he stabbed Caesar at 12:00 noon, and Caesar died of the knife wounds 1 hour later. Q: Was Brutus' stabbing of Caesar = Brutus' killing of Caesar? (Is this one event, or two?)
  - Davidson: Yes. (One.)
  - Kim: No. (Two.)
  - Q: When did the killing of Caesar occur? (12:00? 1:00? Throughout the interval? Did it have a scattered time of occurrence?)
- How many stabbings occurred?
  - A: infinitely many (on Kim's account): a stabbing, a stabbing-with-a-knife, a stabbing-quickly, etc. (These descriptions all provide different *constitutive properties*: "stabs", "stabs with a knife", "stabs quickly", etc.)
  - But this is not so strange. Similarly, how many tables are here? Many billions. Illustration: let A be an electron on the edge of the table. Consider two objects:

    Table 1 = the table, including A.

    Table 2 = the object consisting of all of table 1 *except* electron A. This is also a table, and it is not identical with Table 1. So it is a second table.

    Similarly, there are billions more tables in the offing.

Phil. 4360
Notes #16: Parthood & Identity across Time

Review from last time: Davidson's argument for events; Kim's identity & existence conditions for events; Davidson vs. Kim on individuation of events, the stabbing vs. the killing, &c.
Review from earlier: Lewis on causation.

Questions: What are material objects; under what conditions is a material object *x* identical with object *y*? Special case: When does a material object *persist* over time, and when does it cease to exist?

I. Important concepts:

x is part of y : you all know what this means. Note that everything is a 'part' of itself.
x overlaps y : x and y have a part in common.
x is discrete from y : x and y have no part in common (do not overlap).
the fusion of the S's (where S is some set of objects) : The thing that has as parts the members of S and nothing else (no parts that are discrete from the members of S).

The fusion principle: Given any set of physical objects, there is a unique thing that fuses them, called "the fusion" of the set. (Also called the 'mereological sum'.)

II. About the tinkertoy house:

Thomson has a tinkertoy house on a shelf at 1:15. Let

H = The tinkertoy house that is on the shelf at 1:15.
W = The fusion of the tinkertoys that are on the shelf at 1:15.
W' = The wood that is on the shelf at 1:15.

Plausible claim: H = W = W'

III. Problem:

I remove one stick (alpha) from the house at 1:30 and throw it on the floor, and replace it with another stick (beta). At 1:45 there is a tinkertoy house on the shelf. Is it H?

If Yes: Then there is a contradiction:

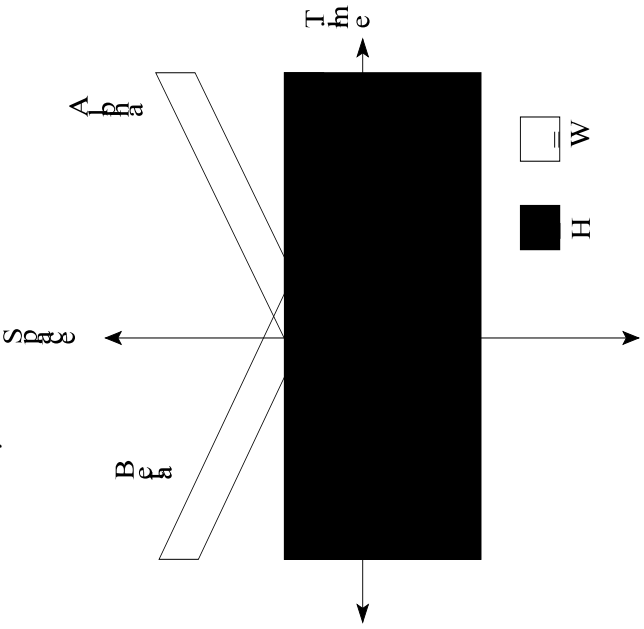
1. H = W. (and W')
2. H is on the shelf at 1:45.
3. W is not on the shelf at 1:45. (nor is W')

Which of these should we reject?

- #1? But H is made only of tinkertoys; therefore, H fuses the tinkertoys, so H is the fusion of the tinkertoys. Likewise, it is the fusion of the wood.
- #2? Then even removing a single particle from H gives you a new house.
- #3? No. The fusion of the S's can't be on the shelf when one of the S's is on the floor.

IV. The 4-dimensionalist solution

- Physical objects are *four-dimensional space-time worms*.
- *The idea of temporal parts:* The 'spatial parts' of a physical object are the parts that occupy different parts of space. Each spatial part exists only in its own region of space. Similarly, the *temporal* parts of an object are the parts occupying different times. Each temporal part exists only in its own *duration* of time.
  - Compare Alfred (the right half of a piece of chalk) with Bert (the later half of a piece of chalk).
- A physical object is the fusion of its temporal parts.
- 'Identity' across time is just *being stages of the same spacetime worm*.



- Solution: H and W are not identical, but they overlap (temporally) from 1:15 to 1:30. (See diagram.)

- Surprising consequence: The chair's seat is not a part of the chair.
  Why: the chair is a certain space-time worm. The seat existed prior to the chair, i.e., the seat-space-time-worm extends prior to the beginning of the chair. So the seat *overlaps* with the chair but is not *part of* the chair, since (a temporal) part of the seat is outside the chair.

V. Objections

a. It implies that change is impossible.
Reply: No, change occurs when an object has some feature at one time that it lacks at another time. Nothing about the theory implies that this doesn't happen.

b. It implies that two physical objects can be in the same place at the same time.
Reply: Yes, but they are *overlapping* objects. This is not so surprising.

c. The problem of temporal chalk extrusion.
- Suppose I hold a piece of chalk in my hand from 12:00 to 1:00. Let Bert = the temporal part of the piece of chalk that exists from 12:30 to 1:00. Then
- Bert is some chalk. Bert also is white, cylindrical, weighs about 3 ounces, etc.
- Bert didn't exist before 12:30. I.e., Bert came into existence at 12:30.
- So a new quantity of chalk popped into existence (while another one, of course, simultaneously went out of existence at the same place) at 12:30.
- The same applies to every other time. Thus:

"As I hold the bit of chalk in my hand, new stuff, new chalk keeps constantly coming into existence *ex nihilo*. That strikes me as obviously false." (p. 213)

VI. A Better Solution

• A thing can have parts at a time, and different parts at another time. I.e., the part-whole relation is actually a 3-place relation:

x is a part of *y* at *t*

not simply:

x is a part of *y*.

• Alpha is part of H at 1:15, and is not part of H at 1:45.

• Alpha is also part of W' at 1:15, and is still part of W' at 1:45.

• So H ≠ W'. ((1) above is false.) Because

x = *y* iff (*z*)(*t*)(z is a part of *x* at *t* iff *z* is a part of *y* at *t*).

I.e., x=y iff x and y have exactly the same parts at all times when they exist.

• W doesn't exist, because there are at least *two* fusions-at-1:15 of the tinkertoys that are on the shelf at 1:15.

[• Alternate view: Identity is also relative to a time. H=W' at 1:15, but H≠W' at 1:30.]

• *Objection:*

Suppose H and W' came into existence together (*ex nihilo*), and then went out of existence at the same time. Then, according to the above, H=W'. But H≠W' since they have different modal properties: W' *could have* existed in the shape of a ship, but H could not have.

• *Reply:*

Modify the above principle to:

x = *y* iff *x* and *y necessarily* have exactly the same parts at all times that they exist.

Phil. 4360
Notes #17: Review

<u>Know what these things are:</u>
Infinity:
- ω (omega)
- c (the continuum)

problem of transworld identity
counterparts
essential/accidental properties
rigid/nonrigid designators
bare particulars
bundle theory
Causation:
- regularity theory of
- counter-factual analysis of

Counter-factual conditional
Events:
- Diff. between Davidson & Kim events
- Kim events: Existence & Identity conditions

Fusions
- & doctrine of arbitrary fusions

4-dimensionalism
- & temporal parts

<u>Know what these people thought of these things:</u>
Aristotle & Galileo: infinity, infinite objects
Cantor & modern mathematicians: infinity
Lewis: possible worlds
- transworld 'identity'
- Meaning of:
  - "possible", "necessary", "impossible"
  - "actual"
- what is causation
- analysis of counter-factuals

Kripke: transworld identity, counterparts
- notion of essential properties
- bare particulars, bundle theory

Davidson: individuation of events
Kim: the constituents of events
- the individuation of events

Thomson: temporal parts
- identity conditions of physical objects
- the parthood relation
- the house & the fusion of tinkertoys (are they identical)

4-dimensionalists: ship-of-theseus type problems

<u>Understand these important arguments:</u>
Galileo's argument for: why infinity isn't a number
The Diagonalization Argument
Lewis: why you should believe in possible worlds
Kim: why there are many tables here
- why there are many events

Thomson's tinkertoy paradox (know what the paradox is)
Thomson's main argument against 4-dimensionalism

Phil. 4360
Notes #18: Is Time Unreal?

Basic concepts relating to time:

- *Moments* are positions in time.
- *Events* are the occupants of time ("the contents of a position in time").
- *"Past", "present", "future":* These are (1-place) properties of a moment/event. These are not permanent (an event is future at one time, then present, then past).
- *x is earlier than y:* This is a two-place relation between moments/events. It is "permanent" (if x is ever earlier than y, then it is always so; they can't change their relation).
- *The A-Series:* The series of past, present, and future moments.
- *The B-Series:* The series of moments, as ordered by earlier-later relations.
- *The C-Series:* The series of events, as ordered, but not necessarily in a temporal order.
  - The C-series is not *per se* temporal, because the existence of an 'order' does not necessarily involve *change*, and change is necessary to have time.
  - Supposedly, the C-series plus the A-series yields the B-series. (Once you know the order of events, and you know how far past/future each is, then you can infer the existence of all the earlier/later relations.)
  - Also, the C-series has no preferred order.
  - [Yes, it is really unclear what the C-series is, or what it is needed for. Why isn't the A-series sufficient?]

The Argument against Time:

I. The A-series is essential to time.

A. Time cannot exist without change.

B. Without the A-series, there can be no change.

General idea: because the only thing that ever really changes (if anything does) is that future events become present, and then past. Elaboration:

1) An event cannot change its position in the B-series. (Premise)
2) An event cannot become or cease to be an event. (Premise/From 1?)
3) An event cannot become another event. (Premise)
4) The characteristics of an event never change, other than its position in the A-series. (Premise)
5) Therefore, the only thing about an event that can change is its position in the A-series. (From 1-4)
6) Therefore, without the A-series, events cannot change. (From 5)
7) If the things that are in time do not change, then there is no change. (Premise)
8) Events are the things that are in time. (Premise)
9) Therefore, if events do not change, there is no change. (From 7, 8)
10) Therefore, without the A-series, there is no change. (From 6, 9)

II. The A-series does not exist.

- What are pastness, presentness, & futureness?

A. If they exist, they are properties/relations that apply to events.

B. They are not relations of an event to anything.
   1. They are not relations between events.
   2. They are not relations of an event to a time.
   3. Therefore, they are not relations of an event to anything. (From 1, 2)

C. They are not properties of events.
   1. If they are properties, they are incompatible properties.
   2. If events have these properties, then every event has all three of them. (Because the same event is future, then present, then past.)
   3. Thus, if they are properties of events, then events have incompatible properties (which is absurd).

D. Therefore, they do not exist. (From A, B, C)

III. Therefore, time does not exist.

Conclusion: Even though there is no time (no A-series or B-series), there is probably a C-series.

Objections:

- To I:
  - Fictional times: events in a fictional story form a B-series, but no A-series.
  - *Answer:* Events in the story do not exist, so they are not in time either.
- To II.C:
  - There is no contradiction, because events do not have these incompatible properties *at the same time*. An event is not past, present, and future, at once. Rather

    An event may be present now*, while it* used to be *future and* will be *past.*

  - *Answer:* This creates an infinite regress, because it means:

    "An event is present in the present, future in the past, and past in the future."

    This means that you are introducing a second A-series, to explain away the contradiction in the first A-series.

The first A-series:	*The second A-series:*
• Events: Ra, Gb, etc. (a's being red, b's being green, etc.)	• Events: P(Ra), F(Gb), etc. (Ra's being in the past, Gb's being in the future, etc.)
• Their A-series properties: Ra is past, Gb is present, etc. I.e., we have P(Ra), Pr(Gb), etc.	• The new A-series properties: F(Ra) is past, P(Ra) is future, etc. I.e., we now have P(F(Ra)), F(P(Ra)), etc.

  - The problem reappears, because just as Ra appeared to be past, present, and future, now F(Ra) appears to be past, present, and future. The same thing will happen to F(F(Ra)), and so on.

Phil. 4360
Notes #19: Time Travel

Review from last time: A-series, B-series. Why A-series is required for change. Why A-series is contradictory, the infinite regress argument.

Question: Is time travel metaphysically possible? (Is there a possible world . . . ?)
- Lewis says yes.
- There are at least three problems with time travel (below).

I. The discrepancy between times

• *Problem:* The traveler leaves on his 'journey.' 'An hour later,' he arrives . . . in the past. Thus:
(a) His arrival is after his departure (after getting in the time machine, you wind up in the past).
(b) His arrival is before his departure (because it is *in the past*).
This is a contradiction.

• One answer: 2-dimensional time. His arrival is after his departure in one time dimension, but before in the other time.

• *Lewis says:* The arrival is after the departure in his personal time, not in external time.
- *External time:* Real, ordinary time, as measured by processes outside the traveler's body.
- *Personal time:* Roughly, amount of aging & other internal processes in the traveler.
- Thus, the traveler's arrival in 100 A.D. is before (in external time) his departure in the time machine in 2100 A.D. But he is a bit older at his arrival in 100 A.D. than at his departure in 2100 A.D., has 'memories' of the events of 2100, etc., making the arrival later in his 'personal time.'

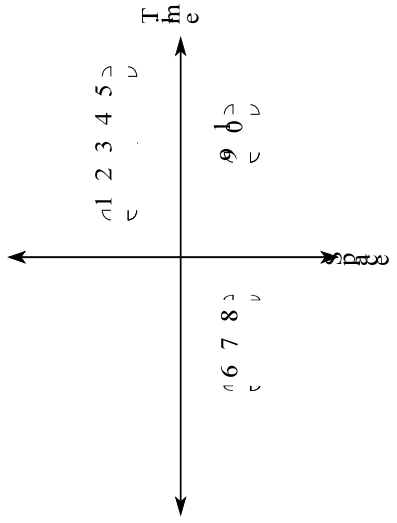


Figure 1. A time traveler. He is a disconnected spacetime worm. The numbers represent stages in his body's development (his physiological age). From his point of view, he started out at 1, then got into a time machine at 5 and went into the past. Then he got into a time machine again at 8 and went into the future, at which point he and his earlier self both existed in different places at the same time. He dies at age 10.

II. Alternate descriptions of what happened

• *Objection:* In figure 1, here is what really happened: There are three people:
- First, a (physiologically) 6-year-old person appeared out of nowhere (call him A). Three years later, he got into a machine labeled "time machine", and the machine destroyed him.
- Some time later, a very similar person was born. Call him B. He lived for 5 years before foolishly getting in another machine labeled "time machine," which destroyed him.
- Two years into B's life, a third person, very similar to the first two but 9 years old, appeared out of nowhere. Call him C. C lived 2 years, and then died.

• Question: Why isn't this a permissible description of the events depicted in figure 1?

• *Lewis says:* What distinguishes the case of time travel from the above description is:
  • A, B, and C are all (temporal parts of) the same person, not three different people.
  • Why? Two criteria of personal identity: (i) Continuity of physical and physiological properties, (ii) C causally depends on B, who causally depends on A.
  • Remember that on Lewis' analysis of causation, backwards causation is possible. Thus, nothing prevents time travel.

III. The grandfather paradox

• Tim is a time traveler, who travels back to 1921 to kill his grandfather (before he, Tim, was born).
• *Problem:*
  1. If Tim can travel back in time, then Tim can kill his grandfather before Tim's father was conceived.
     Why not? Is there some kind of "time police" that are going to stop him from doing it?
  2. Tim cannot kill his grandfather before Tim's father was conceived. For:
     a. This would be inconsistent with Tim's existing in the first place.
     b. Anyway, it is logically impossible to change the past. 1921 only happens once (in 1921). It doesn't happen one way "the first time" and then, later, happen a different way.
  3. Therefore, Tim cannot travel back in time.
• *Lewis says:*
  • The argument is an equivocation.
  • In general: X can do A = X doing A is consistent with certain facts. (Notice how you can get logical possibility, physical possibility, and epistemic possibility.)
  • Two senses of "can":
    • Tim's killing grandfather is consistent with the facts about his current state in 1921, his rifle, his training, etc. (Tim has the right equipment for killing grandfather.)
    • Tim's killing grandfather is *not* consistent with the facts in his, Tim's, personal past.
    • Thus, Tim can kill grandfather, and he can't (in two different senses).

IV. Alternate view of time travel:

• The branching view of time:
  In 1921, the universe splits into two identical universes, but in branch A, Tim shows up with a rifle and kills grandfather. In branch B, grandfather lives until 1957, having begotten father, who begat Tim. Tim gets into a time machine in 1980, disappears from branch B, and appears in branch A in 1921. No contradiction.

Phil. 4360
Notes #20: Absolute Space

Review/continue from last time: 3 objections to time travel; Lewis' responses. The 'branching' view of time.

The traditional dispute: Two conceptions of space, time:

- *The Relational Conception of Space:* All that exists are spatial relations between bodies.
- *The Substantival Conception of Space:* Space exists independently of bodies. Bodies merely 'occupy' space.
- *The Relational Conception of Time:* There are only temporal relations between events.
- *The Substantival Conception of Time:* Time exists independently of events. Events merely 'occupy' time.
- "bodies": Physical objects.
- "substance": Something that 'exists independently'; everything else that exists depends on substances. Also: ultimate subjects of predicates; not predicated of anything. Hence the term "substantival conception of space."
- *Note:* Please do not confuse "the relational theory of space" with the theory of relativity, which will be discussed in a later class.

Related concepts:

- Location:
  - For absolutists: The part of space a body occupies. A statement of location has the form *x*O*y*, where *x* is a body and *y* is a part (a region or point) of space.
  - For relationists: A body's spatial relation to other bodies. A statement of location has the form *x*R*y*, where *x* and *y* are both bodies. R is a spatial relation, e.g. "inside", "next to", etc.
  - Similar points apply to 'location' in time.
- Duration of an event:
  - For absolutists: The measure of the region of time the event occupies.
  - For relationists: A relationship between the beginning of the event (which is itself a small event) and the end of the event. Or, alternately: the number of cycles of a clock that pass during the event (hence, a relationship between the event and some other process).
- Motion:
  - *Absolute motion:* change in the part of space a body occupies, over time.
  - *Relative motion:* change in the spatial relations between bodies.
  - Note: Substantivalists believed in absolute motion; relationists believed in only relative motion, for obvious reasons. Thus, traditional substantivalists were also *absolutists.*
  - Example: Suppose you walk at 5 mph on the deck of a ship. The ship is moving at 20 mph in the ocean. Finally, the earth is moving at 1000 mph *in space*. All these motions are in the same direction. Then your *absolute motion* is 1025 mph (the sum of these motions). Alternately: Suppose your absolute motion is 1025, and the ship's absolute motion is 1020. Then your relative motion, relative to the ship, is 5 mph (the difference in the absolute motions).
  - *Question:* Is absolute motion "motion relative to absolute space"? No; that tries to define absolute motion in terms of relative motion. Relative motion is "difference in absolute

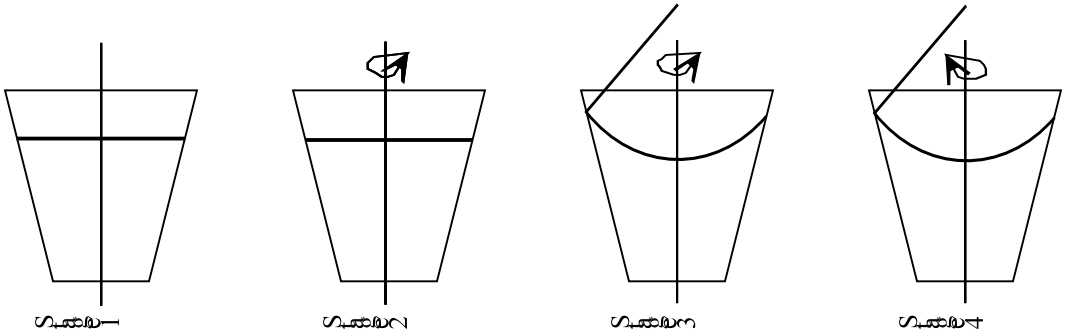
motions."

Measuring:

• We measure the *size* of a body by comparing it with another, *rigid* body (a body that doesn't change its own size).
  • For absolutists: This means it does not change the amount of space it takes up.
  • For relationists: It retains the same "bigger than"/"smaller than" relations to a lot of other bodies.
• Similarly, we measure the *duration* of an event by comparing it with a *uniform* motion/process (a process that continues at a constant speed).
  • For absolutists: It does not change the amount of absolute time that each cycle takes up.
  • For relationists: It keeps a constant relationship (keeps synchronized) with lots of other processes.

Newton's Bucket:

• A bucket of water is suspended from a rope. The bucket is turned around several times, twisting the rope up.



Stage		Relative motion of water w/ respect to bucket	Surface of water
1.	Before releasing bucket.	0	Flat
2.	Bucket is released, starts spinning as rope unwinds.	←	Flat
3.	Water picks up motion of the bucket, starts 'spinning with the bucket.'	0	Concave
4.	Rope has twisted up in the other direction, starts unwinding again. Bucket spins in the opposite direction from stage 2.	→	Concave

• *Newtonian account of the events:*
  • *Newton's First Law:* Bodies at rest tend to remain at rest, and bodies in (absolute) motion tend to remain in (absolute) motion in a straight line, unless compelled to change their state by forces impressed on them.

- ‘Centrifugal forces’ appear for rotating bodies: The ‘force’ pushing towards the outside of the circle is really just the tendency to continue in a straight line.
- Centrifugal force exists in stage 3, 4, because the water is rotating absolutely.

• *Relationist cannot explain this.*

- Newton’s first law is malformed, according to them: there is no such thing as absolute motion in a straight line.
- Suppose you substitute “relative motion”: “Bodies in relative motion tend to remain in relative motion in a straight line unless compelled to change their state by forces impressed on them.” Motion relative to *what*? Everything? This is false.
- The relative motions in stages 1 and 3 are the same (0), but there is centrifugal force in stage 3, not in stage 1. Why?
- The relative motions in 2 and 4 are the same (in opposite directions, but that shouldn’t matter), but there is centrifugal force in 2 but not in 4. Why?

• Thus, the centrifugal force is completely unrelated to the *relative* motions. It can only be explained by the absolute motion of the water.

Phil. 4360
Notes #21: Leibniz vs. Clarke

Leibniz' views:

- Relational theory of space, time.
- Deduces this from the *Principle of Sufficient Reason:* For every fact, there is a reason why it is so, rather than otherwise.
- A vacuum is impossible, because then there would be 'space' existing apart from matter.

Leibniz, Against Absolute Space:

1. If space is absolute, then the following are two different situations:
   a. The world as it actually is.
   b. The world as it is except everything moved over 1 foot to the right. (There are many other variants on this.)
2. There could be no reason for (a) rather than (b).
3. Hence, if space is absolute, then there is a fact for which there is no reason.
4. For every fact, there is a reason. (The PSR.)
5. Hence, space is not absolute.

Clarke:

6. If the relational theory of space & time is true, then
   a. If God moved the universe to the right at 1 million miles per hour, the universe would remain in the same place.
   b. If he suddenly stopped this motion, there would be no shock (no effects at all).
   c. If God had created the universe 1 million years earlier, it would not have been created sooner.
7. (a), (b), and (c) are absurd.
8. The relational theory of space is false.

Leibniz:

Reply to (6):

- It is logically impossible for God to move the universe, or stop it, or to have created it sooner. (See above argument.)
- Further argument:
  9. Such a change could not be observed.
  10. Anything that can't be observed can't exist.
  11. So, such a change is impossible.

Leibniz on Newton's bucket:

- Admits that there is such a thing as "true motion," denies that it should be explained by ref. to absolute space:

  "For when the immediate cause of the change is in the body, that body is truly in motion; and then the situation of other bodies, with respect to it, will be changed consequently, though the cause of that change be not in them." (74)

Is this a good reply?

- What is "the change"? Change in relative motions.
- What causes the water's spinning? (The bucket.)
- L's reply is very obscure. How can he admit the concept of "true motion"??

The modern reply:

- Velocity is relative.
- Acceleration is absolute.
- Rotation is a form of acceleration.
- The water is 'really' rotating in stages 3 and 4.
- *Note*: Is this compatible with the relational theory of space?

Phil. 4360
Notes #22: Spacetime & Relativity

Review: Leibniz' 2 arguments against absolute space/time. The Pr. of Suff. Reason.

Spacetime:

- Two concepts of a space:
  - *Physical space:* the space you are moving around in.
  - *Mathematical space:* a set of things ('points') that have certain mathematical properties. Basically, they have relations to each other that enable them to be arranged along one or more dimensions. *Ex.:* logical space, the color space, the IQ-height space, various spaces in statistics. Physical space is also a mathematical space.
- Another (mathematical) space:
  Spacetime: the 4-dimensional 'space' in which the points are ordered quadruples giving (physical) spatial and temporal coordinates. Spacetime is a mathematical space; it includes physical space.
- Learn to enjoy spacetime diagrams.

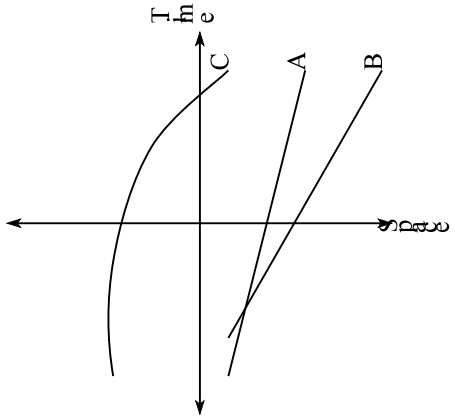


Figure 3. Spacetime. We suppress two spatial dimensions to make it possible to draw on a flat piece of paper. Vertical axis is time; horizontal axis is space. Line A represents an object moving to the right. B is an object moving *faster* to the right. C is an object *accelerating*.

What is the Special Theory of Relativity (STR)?

- It is a theory of the structure of spacetime.

Newtonian spacetime. Various features:

- Shortest path between 2 points is a straight line. Distance between points:
  $D^2 = \triangle x^2 + \triangle y^2 + \triangle z^2 + \triangle t^2$
- Distinguishes:
  Straight lines / curved lines
  Vertical lines / slanted lines
- Spatial & temporal coordinates are separable. Spatial and temporal distances are both objective.
- No speed limit.
- See figure 3.

Minkowski spacetime. Various features:

- Shortest path *not* a straight line (but the path of a light ray). The invariant spacetime interval:
  $I^2 = \triangle x^2 + \triangle y^2 + \triangle z^2 - (c\triangle t)^2$
  Note the minus sign!
- Distinguishes:

Straight lines / curved lines

• Space & time are inseparable. Spacetime intervals are objective, but how they divide into spatial and temporal components is not.
  • There are multiple equally acceptable specifications of the time axis.
  • Hence, no absolute simultaneity.
• Has a 'light cone structure' (any given point has a forward & backward light cone); this is objective. Which s-t points are in the light cone is invariant.
• See figure 4.

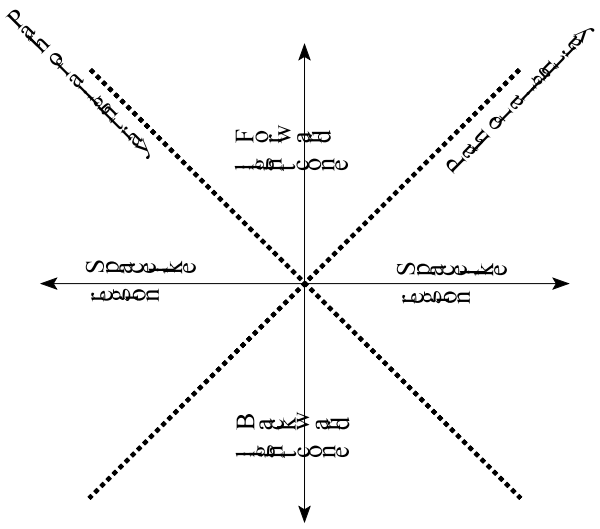


Figure 4. Minkowski spacetime. For any given spacetime point, there is a set of points that would be connected to it by a light pulse sent out in all directions: this set of points is the *forward light cone*. Similarly, there's a backward light cone. Outside the light cones are the points at 'spacelike separation.'

Famous features of STR:

• All inertial reference frames are equally good.
• The speed of light (c) is constant; i.e., every r.f. must agree on whether a thing is traveling at c.
• All of the following are 'relative':
  • Velocity of an object (if below c)
  • Length of an object
  • Time-order of two (spacelike) events
  • Shape of an object
  • Mass of an object
  • Duration of an event
• Nothing can travel faster than c. Why:
  • The putative 'stages' of the spacteime worm of such an object would be *spacelike* related to each other.
  • There is no objective time order to spacelike separated events.
  • Also, it would require infinite energy.

About relative vs. absolute quantities:

• 'Relative' quantities are those which differ between reference frames. They are not in objective reality; they are convention-dependent.
• 'Absolute' quantities are *invariant* ones: i.e., all rf's agree on them. They are in objective reality.

Phil. 4360
Notes #23: More relativity

Review from last time: Invariants in STR: spacetime distance, speed of light, light cones. Acceleration/inertial motion. Relative quantities: velocity, length, shape, mass, time-order, duration. Why: nothing can travel faster than *c*.

About the history of relativity theory:

There are three general arguments for the theory of relativity:

1. The argument from authority:
   a. Physicists have said that STR is true.
   b. Therefore, STR is true.
   • This is the chief argument relied on by most people, including physicists.
2. The philosophical argument:
   • STR rests on the following *philosophical* theses:
      • *The verification criterion of meaning:* A statement is meaningless unless it can be verified (through experience).
      • (Related to the first.) All concepts are dependent on observations; all concepts refer to experiences one would have in certain conditions.
   • These theses amount to "logical positivism," a philosophy popular in the early-mid 20th century.
   • Notice how they appear in Einstein.
   • Notice how they are assumed dogmatically: "I would ask the reader not to proceed farther until he is fully convinced on this point." (Einstein, 22)
3. STR accounts for the null result of the Michelson-Morley experiment.

Why STR is important for metaphysics:

• Very widely accepted.
• Radically revises concepts of space & time. No separation between 'space' and 'time.' No unique time-series. Practically every thought you have about space or time is incoherent, if relativity is true.
• Also radically revises concepts of physical objects. Most of their 'intrinsic' properties are convention- or observer-dependent.
• Has been used to argue for philosophical theses:
   • In epistemology: we should listen to any absurd idea
   • Determinism
   • Against sense data

Lovejoy: Criticism of Einstein

A. Recap of Einstein's argument

Stage 1:

The radically experimental theory of meaning:
(i) A predicate has 'meaning' only if its definition identifies some directly observable event

that would occur under definite conditions, and that would verify the applicability of the term.

(ii) The meaning of a term *is* the occurrence of that event (or rather: to predicate a term is to assert the occurrence of such observable event).

Stage 2:

The definition of "simultaneous." Three kinds of event-pairs:

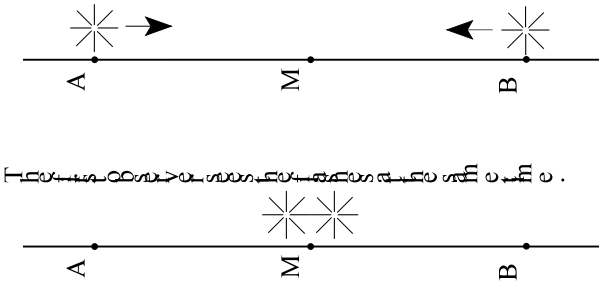
(i) Sense-data or mental events. These can be directly observed to be simultaneous or not.

(ii) Physical events occurring in the vicinity of the observer. These can also be observed to be (approximately) simultaneous.

(iii) Distant physical events. These cannot be directly observed to be simultaneous. Hence, a definition of simultaneity for them is needed:

E_A is simultaneous with E_B iff: If an observer were put at the midpoint of AB, and a light signal were sent from A at the time of E_A and another signal from B at the time of E_B, then the observer would see the light signals at the same time.

(where A is the location of event E_A and B is the location of event E_B.)



Stage 3:

The relativity of the time-order of events:

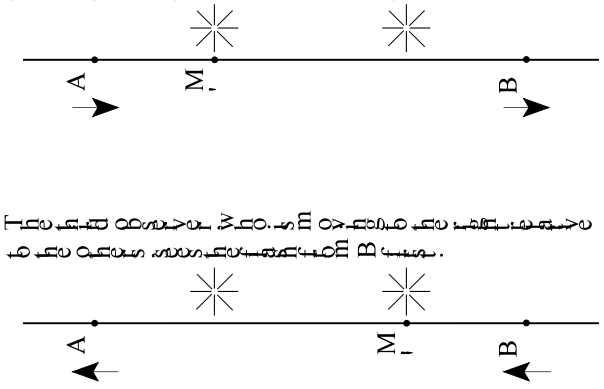
• The above def. of simultaneity leads to the following apparent contradiction (see Fig. 5):

  • The light flashes at A and B are simultaneous.

  • The one at A happened first.

  • And the one at B happened first.

• Solution: time order is 'relative' to a frame of reference.



B. Problems:

1. We already have a concept of simultaneity in Stage 2, (i) and (ii). Why do we need *another* definition for distant events?
   a. The def. in (iii) contains the expression "at the same time," presupposing that we know what this means. This appears to be circular.
   b. E. must think that there are two *different* relations called 'simultaneity', for nearby vs.

distant events. Distant events cannot be 'simultaneous' in the same sense as nearby events.

c. What is the argument for this? The Radically Experimental Theory of Meaning → In circumstances in which the way of verifying that "F" applies to an object are different, the meaning of "F" is different.
   - This implies that, e.g., "rain" has a different meaning when someone says "it will rain tomorrow" based on a weather forecast, and when someone says "it is raining" based on seeing the rain falling.

d. Thus, there is no reason for introducing Einstein's definition, or for thinking that distant events can't be 'simultaneous' in the same sense as nearby events.
   - It may be that the means of *knowing* whether distant events are simultaneous is different.
   - It might also be that we lack such a means of knowing. Both of these would be irrelevant.

   > "The experimental theory of meaning is, in fact, radically opposed to the spirit of scientific empiricism, in so far as it declares that a quality or relation which, in certain instances, *is actually* exemplified in experience *can not* exist in instances beyond the reach of experience." (629, emphasis Lovejoy's)

e. Other consequences of radical experimentalism:
   - Concept of light traveling is meaningless. There are only 'illuminated bodies'.
   - Notion of stuff happening on Arcturus 'now' is meaningless.
   - 'Distance' has a different meaning when applied to stars versus objects on the earth.
   - Statements about the past are either meaningless or just statements about the present 'traces' of past events. [Dummett actually adopted the latter view, decades after this article.]

2. Is Einstein's def. an *arbitrary* verbal definition, or is it supposed to correspond to the pre-existing meaning of the term?
   - E. accepts only (at most) three constraints on the def.: (a) that it should define a relation between events, (b) that it should not be a spatial relation, (c) that it should supply a verification criterion.
   - There are many other definitions satisfying these conditions. E.g., substitute "sound" for "light" in the definition. Or, substitute "donkey" for "light signal."
   - E's def. is an arbitrary verbal definition.
   - There is thus no profound significance to his discovery about "simultaneity."

3. Another possible way of defining 'simultaneous':
   - E's def. requires only that the observer *have been* at the midpoint of AB.
   - Why not modify the def. to require that the observer *be* at the midpoint of AB *when the light signals arrive*?
   - [Problem: Suppose the two events occur in space, so there is no rigid body between them. What are points A and B? Points in absolute space? Does the first def. mean that the observer should have been halfway between A and B *at the time* E_A and E_B happened (but this presupposes the concept of simultaneity)?]

4. The second def. in (3) is required for the two observers to be talking about the same thing. Otherwise they are not performing *the same experiment* to test the simultaneity of the events.

5. Einstein overgeneralizes. His argument (even if we ignore the preceding objections) would only show that simultaneity depends on one's state of motion parallel to the line between the two events. It doesn't vary with just *any* motion (e.g., consider motion perpendicular to that line).

6. E's definition presupposes the notion of a "time of transit" (or duration). But this presupposes a concept of simultaneity b/c it implies temporal relations between stages in a body's motion.

7. E isn't consistent, for he sometimes talks about something happening on the train "while" something else is happening on the embankment. This implies that events in the two reference frames occupy the same time-series. (Maybe the point here is that E. fails to specify in which reference frame these events are simultaneous.)

Phil. 4360
Notes #24: Non-Euclidean Geometry

I. About Pure Geometry

- Three kinds of geometry:
  ("Parallel" lines: Lines that do not cross.)
  a) *Elliptical geometry* : through a given point outside a given line, there are no parallel lines.
  b) *Hyperbolic geometry* : through a given point outside a given line, there are infinitely many parallel lines.
  c) *Euclidean geometry* : through a given point outside a given line, there is exactly one parallel line. (This is the "axiom of parallels.")
- Two kinds of geometry:
  Pure geometry is a word game with made up, stipulative definitions and rules. No connection to reality needed.
  Applied geometry is the application of a geometrical system to some thing in the world.

a) A model of elliptical geometry: The "plane" is the surface of a sphere. The "straight lines" are great circles.

- Features of this geometry:
  1) No parallel lines.
  2) The interior angles of a triangle will be more than 180°.
  3) C/d of a circle $< \pi$.
- Note: Again, "line", "triangle", etc. are *not* used in the ordinary English sense of the words.
- This surface has "positive curvature".
- This proves: Elliptical geometry is consistent.
- Why is it called "elliptical": it can be modeled on the surface of an "ellipsoid".

b) A model of hyperbolic geometry: The "plane" is a saddle surface (surface of a hyperboloid). "Straight lines" are geodesics.

- Features of this geometry:
  1) Many parallel lines.
  2) The interior angles of a triangle will be less than 180°.
  3) C/d of a circle $> \pi$.
- This surface has "negative curvature".
- This proves: Hyperbolic geometry is consistent.

- *Note* : Two senses of "curvature": a) Physical curvature
  b) Mathematical "curvature"

II. About general relativity:

- Spacetime as a mathematical space. (Note: Lots of things are "spaces" in the mathematical sense, although they have nothing special to do with physical space.)
- Points as ordered quadruples. ("event" locations)
- Newton's spacetime: Euclidean.
- Einstein's theory:

1. Concentrations of mass/energy alter the geometry of spacetime, "curving" spacetime. Note:
   a) *Spacetime*, not *space*
   b) It is not *physically* curved. It is curved in the mathematical sense.
   • In this geometry,
   a) Points are ordered quadruples again
   b) Straight lines are paths traced by light rays
2. Objects travel straight lines through spacetime when not acted on by forces. Gravitational 'force' is replaced by spacetime curvature.
3. Light always traces straight lines through spacetime.

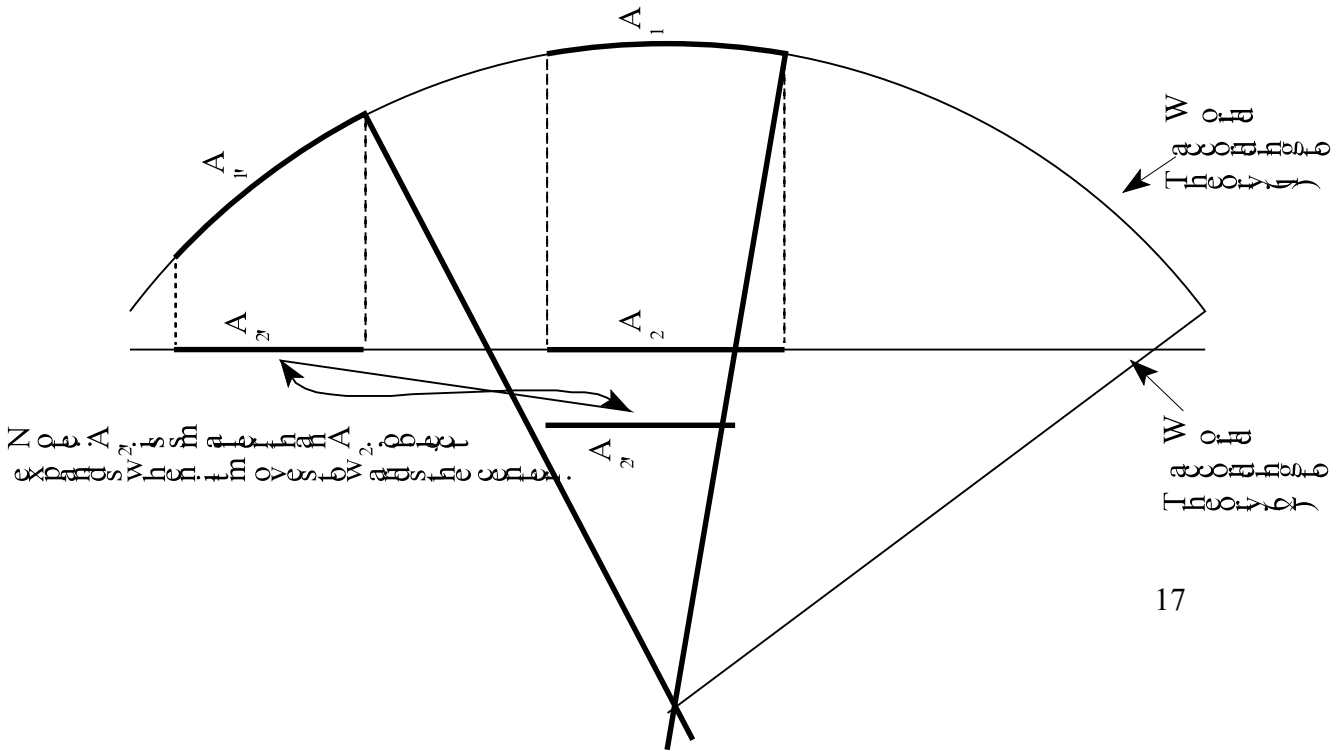
• Empirical evidence:
a) the bending of light around the sun
b) gravitational red shift
c) advance of the perihelion of Mercury

III. Two alternative interpretations (Carnap's example):

• You have two people occupying a 2-dimensional world (see diagram below).
Theory 1 : You have rigid (fixed size & shape) rods moving on a curved surface.
Theory 2 : Rods affected by universal forces, on a flat surface.
• 'Universal forces': Forces that distort everything in the same way and cannot be shielded against.
• A heuristic for seeing the relation between the theories: Imagine the theory-1-world above the theory-2-world, and a light shining directly down from above. In (2), objects expand or contract to be the size of the 'shadow' of the objects in theory (1).

• Note that they get the same empirical predictions. (See why.)
• You cannot directly measure distortions made by universal forces.
• They have *effects* indiscernible (by observation) to those of the noneuclidean geometry.

• Which theory is better?
• Einstein: (1) is better.
• Carnap: (1) and (2) are the same theory.
• Huemer: (2) is better.



Phil. 4360
Notes #25: The Dimensions of Space (Swinburne)

Review: Parallel postulate. Elliptical, hyperbolic, euclidean geometry. Postulates of GTR. Universal force theory. Einstein vs. Carnap vs. Huemer.

The dimensionality of a space:

Three definitions:

1. Space is n-dimensional = n is the smallest number such that every point can be specified by n real numbers.
   Problem: Then every space is 1-dimensional. The set of ordered n-tuples of real numbers can be mapped onto the set of (single) real numbers.
2. A recursive definition:
   • A point is 0-dimensional.
   • If every point in a space has arbitrarily small neighborhoods that are bounded by (n-1)-dimensional spaces, then it is n-dimensional.
3. Space is n-dimensional = n is the smallest number such that every point can be specified by n *distance* (or direction) measures.

• Def. 3 is the best.

Earlier philosophers say:

• Everyone up to and including Kant agreed : space is necessarily 3-dimensional. (Incl. Kant, Galileo, Ptolemy)
  • They argue from the fact that *three and only three lines can be mutually perpendicular at a point*.
  • Problem: The necessity of the premise is no more obvious than the necessity of the conclusion.

• Opposite argument:
  1. For any n, there is a formally consistent n-dimensional pure geometry.
  2. Therefore, it is possible that space be n-dimensional.

  • Problem: this is invalid. Must show that any pure geometry could be a physical geometry.

• Another argument:
  1. The 3-dimensionality of space follows from
     a) The principle of total effect: The total change in momentum produced by a gravitational force on a collection of objects completely surrounding the source, is constant (does not vary with distance).
     b) The inverse-square law: Gravitational force between 2 bodies decreases with the square of the distance between them.
  2. Furthermore, (a) combined with a different law, e.g., an inverse-cube law, entails a different dimensionality of space, e.g., 4-dimensional space.
  3. The inverse square law is contingent; a different law is possible.
  4. Hence, the 3-dimensionality of space is contingent; a different dimensionality is possible.

  • Problem: This argument requires assuming that (1a) is necessary.

• Form of the argument:
1. (a) is necessary.
2. (b) is contingent.
3. Hence, (a & ~b) is possible. (From 1, 2)
4. (a & ~b) entails that space is not 3-dimensional.
5. Hence, it is possible that space not be 3-dimensional. (From 3, 4)

• Problem: (2) is false; or, if (2) is true, then (1) is false. E.g.: The principle of total effect is contingent; or, if you think it is necessary, then you should think the inverse-square law is necessary.

Impossibility of a 2-dimensional world:

• Imagine flat 'people' living in a plane. Is this a 2-D world? No, for:
1. It remains *logically possible* for the material objects in the plane to be lifted out perpendicular to the plane.
2. Therefore, the location of those objects also requires specifying their elevation from the plane (even if for all objects, that elevation is '0').
3. Thus, 3 measurements of distance are required to specify the locations of the objects.
4. Thus, the world is 3-dimensional.

• Other comments:

We can only conceive of the alleged 2-D world as a subset of our world (as a plane). Thus, you cannot refer to a 2-D world: "The purported description of the 2-D world fails. The world described ... would in fact be a 3-D world." (150)

Impossibility of a 4-dimensional world:

First argument:
1. If space were 4-dimensional, then it would be impossible for a 3-D space to exist. (By analogy to above argument.)
2. Therefore, if it is possible that space be 4-dimensional, then our world is not 3-D. (From 1)
3. But our world is 3-D.
4. Hence, it is impossible that space be 4-D.

Second argument:
1. If space could be 4-D, then it would be possible to see (from outside) all of the inside of an opaque volume.
2. This is impossible.
3. Hence, space cannot be 4-D.

Third argument:
1. If space could be 4-D, then one could pass through a 3-D object without touching its surface.
2. This is impossible.
3. Hence, space cannot be 4-dimensional.

Phil. 4360
Notes #26: Review of Unit 3

Know what these things are:
Substantival conception of space
Relational view
Newton's bucket & what it's supposed to show
  The Newtonian explanation
  N's 1st Law of Motion
  The modern view of this
Geometry:
  The Axiom of Parallels
  Euclidean, hyperbolic, & elliptical:
    How they differ
    How they can be modeled
  The postulates of General Theory of Relativity
  Difference between 'universal force' theory and GTR
    Einstein's view on them
    Carnap's view
    Huemer's view
Special Relativity:
  The 'invariant interval'
  Constancy of speed of light
  Terms: spacelike, forward light cone, backward light cone
  What is relative & what is invariant in the theory
__Einstein's argument
    Verification criterion of meaning
    E's def. of 'simultaneity'
    How it leads to relativity of simultaneity

Know what these people said about these things:
McTaggart:
  A-series, B-series
  Basic outline of his argument
  Alleged contradiction in concepts of 'pastness', 'presentness', 'futurity'
    The infinite regress
Lewis, time travel:
  Personal vs. external time
  Criterion of personal identity
  The grandfather paradox
  The branching view of time
Leibniz:
  Pr. of Sufficient Reason
  Main arguments against abs. space:
    Arg from Pr of sufficient reason
    Unobservability argument
Lovejoy's view:
  On the meaning of 'simultaneity'.
  His example of the meaning of "rain"--what this shows
  (Recognize other possible examples)
  Why he thinks E's argument is against the spirit of scientific empiricism

Phil. 4360
Notes #27: For Determinism (Blanshard)

I. Definitions

- *Determinism:* Two definitions have been popular:
  1. Every event has a sufficient cause.
     - Events include persistences of states as well as changes.
     - Distinguish 3 kinds of "cause": Necessary cause, sufficient cause, and influence.
  2. If P_0 is a complete and correct description of the state of the universe at time *t*, L is a complete statement of all the laws of nature, and P is a description of any (actual) event occurring after *t*, then (P_0 & L) entails P.
- *The Free Will Thesis:* The view that someone, at least sometimes, has free will.
  - S has free will: Some of S's actions are free.
  - Free action: Two requirements:
    1. Self-control (subject controls own actions).
    2. Alternate possibilities (S could have done otherwise).
  - More about alternate possibilities:
    For some *p*, S has a choice about the fact that *p*.
    S has a choice about the fact that *p*: *p*, but S could have made it the case that ~*p*.
  - FWT does *not* mean:
    S has a choice about everything.
    All of S's actions are free.
    S can do whatever S wants. (These are all absurd.)
- *Compatibilism:* The view that free will is compatible with determinism.
- Three traditional positions:
  1. *Hard Determinism:* Determinism + ~FWT.
  2. *Soft Determinism/compatibilism:* Determinism + FWT.
  3. *Libertarianism:* FWT + ~Determinism.

II. Arguments against Determinism (Blanshard)

A. The stubborn feeling of freedom: No matter how much we learn, we always still feel free in all of our actions.
   - This is just because when you are making a choice, your attention is normally focused on the future consequences of your action, not on the present causes of your choice.
   - But why does this feeling not go away when you learn about determinism?
     • When you are actually making a choice, you are never attending to the causes of your choice. [?]

B. The argument from modern physics: Quantum mechanics has rejected determinism.
   • About the Heisenberg Uncertainty Principle: Four interpretations:
   1. It is a mere limitation on the accuracy of our measuring techniques.
      • No relevance for determinism.
   2. The elementary particles are not the *sort* of thing to have positions/momenta.
      • This also has no relevance for determinism.
   3. They have positions/momenta, but these are indeterminate, i.e., violations of the law of

excluded middle.
• This is nonsense.

4. They have positions/momenta at any given time, but these are causally unrelated to anything that went before.
   • But how could one ever confirm this?

• The practice of science presupposes determinism. When we do not know the causes of things, we have always assumed there were hidden causes.
• Human beings are large enough physical objects to be for all practical purposes deterministic systems (governed by classical physics).

C. The moral argument:
   • The following things presuppose ~determinism: praise/blame, punishment, remorse, duty.
   • Determinism means that people are just machines.

Reply to the latter:

• We need not & should not accept *physicalism*. Mental states are different from physical states.
• Mental states have different kinds of causal relations.
   • We can see why one mental state causes another.
   • Logical relationships between the contents of thoughts are part of how they cause one another.
• To be free is to be moved by an impersonal ideal, viz. the sense of duty. (Compare Kant.) "What we want is control by the objective requirements of the case."

[Comments:
• Are Blanshard's responses on each of these points convincing?
• Did he give any arguments *for* determinism?]

Phil. 4360
Notes #28: Against Free Will (Edwards)

Soft determinists say:

- There is no conflict between 'free action' and determinism.
  - A free action is one which is
    a) Not constrained, compelled, or caused by outside forces,
    b) Not the product of abnormal psychological compulsions, etc., and
    c) The product of the agent's own (conscious) desires.
  - Thus, a free action is caused. It is merely caused in a different way from an unfree action.
- There is no conflict between determinism & moral responsibility.
- Not only can people have FW with respect to their actions; they can also have FW with respect to their characters.

The hard determinist says:

- Two proposed conditions on responsibility:
  a. Lack of constraint, compulsion, abnormal psychology.
  b. Choosing one's own character.
- Softies only recognize (a). Hardies insist on (b) also. In defense of (b):
  1. S is responsible for X ⊃ S is responsible for the cause of X.
  2. S's character is the cause of his actions.
  3. S is responsible for his actions ⊃ S is responsible for S's character. (From 1, 2.)
  4. S is responsible for his character only if S chose his character.
  5. Therefore, S is responsible for his actions, only if S chose his character. (From 3, 4.)
  6. No person chooses his character.
  7. Hence, no person is responsible for his actions. (From 5, 6.)

Possible objections: (not in Edwards)

A. (by Hospers)
- The consequent of (5) is a contradiction. For:
  - To choose something, S must first exist.
  - If S exists, S already has a character; one cannot exist w/o a character.
  - To choose one's character, one must not already have a character.
  - Hence, to choose one's character, one must both exist and not exist.
- But the notion of responsibility is not self-contradictory.
- Hence, the consequent of (5) is not a condition on responsibility.

B. Is premise (1) true?
- Consider alternative premise: S is responsible for X only if S *is* (part of) the cause of X.

C. What does "the cause of" mean? A sufficient cause? Partial cause? Necessary condition?
- Suff. cause → (2) begs the question, may be false.
- Nec. cause → (1) says, "S is responsible for X only if S is responsible for every necessary condition on X" → Obviously false.

D. Is (6) true?

III. Better Arguments Against Free Will (not in Edwards)

A. The argument from physics:
   1. All human actions are constituted by the motions of particles.
   2. The motions of particles are determined, or at least not controlled by FW.
   3. Therefore, all human actions are determined, or at least not subject to FW.

B. Determinism by desires
   1. All human actions are caused by beliefs and desires.
   2. If A causes B, then S has a choice about B only if S has a choice about A.
      • S could avoid B only if S could avoid A.
   3. We do not have a choice about our beliefs and desires.
      • Note: Infinite regress threatens.
   4. Therefore, we do not have a choice about any of our actions.

Phil. 4360
Notes #29: For Compatibilism (Stace)

W. T. Stace: preliminary remarks

- Free will is important to morality. "S should do A" entails "S has free will."
- Responsibility & punishment require free will.
- All deniers of free will really believe in it in real life.
- The problem is merely verbal. It arises from a confusion about the meaning of "free".

The meaning of 'free will':

- It has commonly been assumed that 'free will' implies indeterminism.
- Meanings to be determined by common usage. Observe common usage:

J: I once went without food for a week.
S: Did you do that of your own free will?
J: No. I did it because I was lost in a desert and could find no food.

G: I once fasted for a week.
S: Did you do that of your own free will?
G: Yes. I did it because I wanted to compel the British Government to give India its independence.

J: Did you sign this confession of your own free will?
P: No. I signed it because the police beat me up.

- What is the distinguishing characteristic of free actions?
  - Not the absence of causes. For (a) the free actions all have causes. (b) Even if indeterminism is true, there is no reason to think the unfree actions were any more determined than the free actions.
  - *"Acts freely done are those whose immediate causes are psychological states in the agent. Acts not freely done are those whose immediate causes are states of affairs external to the agent."* (286-7)
    - Objection: What if a thug points a gun at you and demands your money? You hand it over. The cause of your action is the fear of death, which is an internal psychological state. Is your action free?
- What does this mean: "S could have done otherwise"?
  S could have done otherwise if he wanted to ≈ S *would* have done otherwise if he wanted to.
- Understand why all of this means that FW is compatible with determinism.

Punishment:

- Why do we punish people?
  To supply causes to modify their behavior.
- This is the same as the reason why we give fertilizer to a plant. [Is this true?]
  "The only difference is that different kinds of things require different kinds of causes to make them do what they should. Pain may be the appropriate remedy to apply, in certain cases, to human beings, and oil to the machine. It is, of course, of no use to inject motor oil into the boy

or to beat the machine." (290)

- Punishment presupposes determinism--human actions have causes.

Free will requires determinism:

"If there were no determinism of human beings at all, their actions would be completely unpredictable and capricious, and therefore irresponsible." (291) [Is this true?]

1. If human actions lack causes, they are random.
2. If random, they are not free.
3. Therefore, freedom requires determinism.
4. If freedom requires determinism, then it doesn't require *in*determinism.
5. Therefore, freedom doesn't require indeterminism.

Phil. 4360
Notes #30: Against Compatibilism (van Inwagen)

***To prove*:** the existence of free will is incompatible with determinism.

Rough, informal idea:

If determinism is true, then my actions are the (logical) consequences of events in the remote past, together with the laws of nature. But it is not up to me what went on before I was born, and it is not up to me what the laws of nature are either. Therefore, the consequences of these things, including my present actions, are not up to me.

Preliminaries:

N*p* = No one has any choice about the fact that *p*. [Note: "N" is a sentential operator. "N*p*" implies that *p* is true in fact, and no one could have rendered *p* false.]

P_0 = A complete & correct description of the state of the universe at some instant in the remote past.

L = The conjunction into a single proposition of all the laws of nature.

P = An arbitrarily chosen (correct) description of anything happening after that time.

Determinism: The thesis that *at any given time, there is exactly one future course of events that is consistent with the state of the universe at that time and all the laws of nature*. I.e., (P_0 & L) entails P.

Some plausible rules of inference:

Rule α*:* From □*p*, deduce N*p*. ('□' denotes metaphysical/logical necessity.)

Rule β*:* From N*p* and N(*p* ⊃ *q*), deduce N*q*.

Alternate pair of rules (these are equivalent as a pair):

The Conjunction Rule: From N*p* and N*q*, deduce N(*p* & *q*).

The Entailment Rule: From N*p*, deduce N*q*, whenever *p* ⊨ *q*.

Formal Argument:

First Version (van Inwagen):

1.	□[(P_0 & L) ⊃ P]	Assumption, def. of determinism
2.	□[P_0 ⊃ (L ⊃ P)]	1; exportation
3.	N[P_0 ⊃ (L ⊃ P)]	2; rule α
4.	NP_0	Premise
5.	N(L ⊃ P)	3, 4; rule β
6.	NL	Premise
7.	NP	5, 6; rule β
8.	Determinism ⊃ NP	1-7; conditional proof

Alternate Version:

1.	$(P_0 \& L) \vDash P$	Assumption
2.	NP_0	Premise
3.	NL	Premise
4.	$N(P_0 \& L)$	2,3; conjunction rule
5.	NP	1,4; entailment rule
6.	Determinism $\supset$ NP	1-5; conditional proof

Commentary:

- We assume for conditional proof that determinism holds (1).
- We infer that no one has a choice about an arbitrarily chosen fact.
- Thus, if determinism, then no one has a choice about anything. Thus, soft determinism fails.

Examining the premises and rules

There are 4 premises/rules involved:

a) NP_0.

We have no choice about the past.

b) NL.

We have no choice about the laws of nature. This seems to follow from the meaning of "law of nature."

c) If $\Box p$, then Np.

We have no choice about logically necessary truths.

d) If Np and N($p \supset q$), then Nq.

Suppose one of the following:

1. 'S has a choice about the fact that p' = 'S has access to some possible world in which $\sim p$.'
2. 'S has a choice about the fact that p' = 'S can do some act A such that if S did A, it might not be the case that p.'
   Np = For every act A that S can perform, if S did A, it would definitely still be the case that p.

If either of these is correct, then rule β (and the conjunction rule) come out valid.

- PvI thinks (d) (rule β) is the only one that could reasonably be questioned.

A third, simpler argument:

1. If, in order for me to do A, something would have to have happened in the past that did not in fact happen, then I cannot now do A.
2. If determinism is true, then in order for me to do something different from what I actually do, things would have to have happened in the past that did not in fact happen.
3. Therefore, if determinism is true, then I cannot now do anything different from what I actually do. (From 1, 2.)
4. Free will requires alternate possibilities.
5. So, if determinism is true, then I have no free will. (From 3, 4.)

Phil. 4360
Notes #31: For Free Will: The Self-Refutation Argument

The Presupposition of Thought

• Lucas phrases this as an argument against determinism, but it seems he is really arguing against *physicalism*. Not completely clear what his argument is. Perhaps:

1. If physicalism holds, then all beliefs have only non-rational causes. (Random or mechanistic movements of insentient little particles.)
2. If a belief has only non-rational causes, then
   a. it cannot be true;
   b. it cannot be justified;
   c. we have no adequate reason to think it likely to be true;
   d. the person holding it has no adequate reason to think it likely to be true; or
   e. it cannot be held *because* it is true.
3. Therefore, if determinism is true, then
   a. it is not true;
   b. we are not justified in believing it; or
   c. we don't believe it *because* it is true. (From 1, 2.)
4. If we know that we don't believe *p* because it is true, then we aren't justified in believing *p*/don't know that *p*.
5. Therefore, determinism is
   a. false (from 3a); or
   b. unjustified (from 3b or 3c + 4).

Is Determinism Self-Refuting? Yes.

Preliminaries:

• Presupposition of rational thought: One should believe only what is true (avoid false beliefs).
• 'Ought' implies 'can' principle: 'S should do A' implies 'S can do A.'
• Def. of (hard) determinism:
  No person ever has more than one course of action available. Or:
  (S)(A) (S can do A ⊃ S does A).
• Def. of MFT (the Minimal Free-will Thesis):
  ~Determinism. Or:
  (∃S)(∃A) (S can do A & S does not do A).

Determinism is self-refuting:

1. With respect to the free will issue, we should believe only what is true.
2. If S should do A, then S can do A.
3. I believe MFT.
4. Assume determinism. Then if S can do A, S does A.
5. If S should do A, S does A. (From 2, 4.)
6. With respect to the free will issue, we believe only what is true. (From 1, 5.)
7. MFT is true. (From 3, 6.)

Objections:

A) "Sometimes, we should do the impossible"? Consider some similar situations:
   - x is impossible, & S ought *to attempt* to do x.
   - x is impossible, & x *would be good*.
   - x is impossible, & S is obligated to *do x if he can*.

B) "Instead of (1): we should believe only what is *justified*."

C) "(1) is 'begging the question,' because if determinism is true [given that 2 and 3 are also true], (1) is false." Three conceptions of begging the question:
   1) An argument begs the question iff: If the conclusion is false, then a premise is false.
   2) An argument begs the question iff: The conclusion is, or is contained in, one of the premises.
   3) An argument begs the question iff: The conclusion is used to justify (argue for) a premise.

D) "There are two senses of 'should', in (1) and (2), epistemic and moral."

E) "(1) is false because: people have no control over their beliefs / believing is not an action."

F) "This must be wrong, since it derives a contingent conclusion from necessary premises."

Phil. 4360
Notes #32: For Free-Will: The Lucas-Gödel Argument

I. Lucas' General Position on Free Will

- Lucas is a libertarian.
- Mechanism: The view that the human mind is a "machine" or machine-like. Lucas appears to view this as equivalent to determinism. (p. 113)

  "Our idea of a machine is just this, that its behaviour is completely determined by the way it is made and the incoming 'stimuli': there is no possibility of its acting on its own..."
- Lucas has two arguments against mechanism:
  1) Determinism is self-refuting.
  2) Gödel's Theorem refutes mechanism.

II. About Gödel's Theorem

Definitions:

- *Gödel's Theorem:* Any formal system capable of representing arithmetic on the natural numbers, is either inconsistent or incomplete.
- *Formal systems:* A formal system contains:
  i) A set of symbols;
  ii) Formation Rules: a set of *formal* (syntactic) rules defining how symbols may be combined to form 'sentences';
  iii) Axioms: sentences that the system starts with; and
  iv) Transformation Rules: a set of formally (syntactically) defined rules for when you can derive on sentence from other sentences.
- *Inconsistency:* A formal system is inconsistent iff a formula of the form (P & ~P) can be derived. Also: any proposition whatsoever can be derived.
- *Incompleteness:* A formal system is incomplete iff there is at least one statement which is true in all the intended models, but cannot be derived by the rules of the system.

- G's theorem was initially a proof of the incompleteness of *Principia Mathematica* (PM), the system of Whitehead and Russell.
- *How was Gödel's Theorem proved?* Five major stages:
  1. He showed that every statement of PM could be associated with a unique natural number (its "Gödel number").
  2. He showed that every formal operation on a sentence or set of sentences corresponded to an arithmetical operation on the Gödel number(s) of the sentence(s).
  3. From (2), it follows that there exists a definite arithmetical property that belongs to all and only the Gödel numbers of sentences that can be formally derived in the system.
  4. Finally, he showed that there was a sentence of PM which says, of *its own* Gödel number, that it does not have that property. Call this the Gödel sentence for PM.
  5. Either the Gödel sentence is true, or it is false.
     a. If it is true, then the Gödel sentence is unprovable in PM; hence, there is a true but

unprovable sentence of PM. Hence, PM is incomplete.

b. If it is false, then the Gödel sentence *is* provable in PM. Hence, PM is capable of deriving a false arithmetical statement. Hence, PM is unsound. (In this particular case, its unsoundness would also entail inconsistency.)

- To notice about this:
  - Gödel shows not only that *there is* a true but unprovable statement of PM; he provides *a method for constructing the statement.*
  - The method can be applied to any formal system that can represent arithmetic. Hence, it shows that *any* consistent formal system that can represent arithmetic is incomplete.

III. Lucas' argument from Gödel's Theorem

1. If mechanism is true, then there is a formal system that accurately represents all human thinking.
2. Assume, for reductio, that there is such a system. Call it F.
3. F is not an inconsistent system. (Premise.)
4. F is capable of representing arithmetic. (Premise.)
5. F cannot be used to derive the Gödel sentence for F. (From 3, 4, and Gödel's Theorem.)
6. A human mathematician can derive the Gödel sentence for F. (Premise.)
7. Therefore, F does not correctly represent all human thinking. (From 5, 6.)
8. No formal system correctly represents all human thinking. (From 2-7, RAA.)
9. Mechanism is false. (From 1, 8.)

IV. Objections

(a) Suppose we were to supplement F by adding in its Gödel sentence, as an axiom.
Answer: Then there will be a new Gödel sentence, for the enlarged system.

(b) What if we add an infinite series of Gödel sentences?
Answer: Then there will be a new Gödel sentence, for the enlarged system. (Note that (a) and (b) are objections to Gödel's Theorem itself.)

(c) But computers can do lots of things humans cannot, so computers are better than human minds! Also, no human mind can surpass all machines simultaneously.
Answer: This isn't the issue.

(d) Suppose we added to the system an operation for provisionally adding unproven sentences?
Problem: How to ensure that the system would choose to add the Gödel sentence, and not the negation of the Gödel sentence?

(e) But people are inconsistent. Perhaps human minds are modeled by an *inconsistent* system.
Answer: (i) No, for an inconsistent *formal system* is a system in which *every* statement is derivable. (ii) Human inconsistencies are *mistakes* (like malfunctions), not set policies as in an inconsistent formal system.

(f) Perhaps some future, more sophisticated machines will be developed that are not completely predictable.
Answer: But then these wouldn't be "machines" in the intended sense.

Phil. 4360
Notes #33: The Problem of Personal Identity

Dennett's story:

Yorrick = Dennett's brain.
Hamlet = Dennett's body.

• Where is Dennett?

(1) Dennett is where Hamlet is. (Cf. answer 1 below.)
Problem: Brain transplant cases. In a brain transplant, you want to be the donor, not the recipient. When Dennett gets a new body, he still exists.

(2) Dennett is where Yorrick is. (Cf. answer 2 below.)
Problem: (a) He doesn't seem (to himself) to be in the vat. (b) If Dennett robbed a bank, the state wouldn't lock up his brain and leave his body to roam free.

(3) Dennett's location is determined by his 'point of view.'
Problem:
• Implies that people are infallible about their locations. Can't people get lost? Yes, but even then you still know you're 'here.'
• What about people in the Cinerama? Or people in laboratories operating feedback-controlled mechanical arms to handle dangerous materials?
• This leads to the conclusion that Dennett is an immaterial object.
  • When he loses radio contact between his brain and his body, his location shifts from under Tulsa to Houston.
  • No physical object shifts location.
  • So he must be a nonphysical thing.

(4) Dennett is in a scattered location.

The Problem of Personal Identity

- Q1: What are you?
  Q2: Fill in the blank: "*x* is the same person as *y* iff _____."
- *Not* the question: (a) When are two people identical with each other? (b) When is *x* qualitatively identical with *y*? (c) When are you still you?

Possible answers:

1. (Q1) You are your body. (Q2) *x* has the same body as *y*.
   Problem: Brain transplant cases. (You have your brain transplanted into another body. The recipient body then becomes *your body*.)

2. (Q1) You are your brain. (Q2) *x* has the same brain as *y*.
   Problems: (a) Mind transplant case? (Your brain's information is transferred into another brain.) (b) The 'brain of Theseus' case. (Neurons of your brain are replaced one at a time.)

3. (Q1) You are your mind. (Q2) *x* has the same mind as *y*. *But what is your mind? When does x have the same mind as y?*
   3a. *x* has the same mind as *y* iff *x* and *y* have the same soul (or mind-stuff).

Problems: Is there a soul? Where do they come from? How do they get hooked up to bodies?

3b. *x* has the same mind as *y* iff *x* has (some of?) the same memories as *y*. (**Locke**)
Problem: Circular, because "memory" presupposes personal identity. (**Reid**) See also under (3c).

3c. *x* has the same mind as *y* iff *x*'s quasi-memories are a subset of *y*'s quasi-memories, or vice versa. (modified version of **Locke**)
(*"Quasi-memories"*: mental states that are qualitatively just like memories, but the events they represent need not have actually happened, nor need they have happened to the person who has the quasi-memory.)
Problems: (a) Amnesia case. (b) Normal forgetting. Entails non-transitivity of identity. (**Reid**) (c) Entails that more than one person can be you.

3d. *x* has the same mind as *y* iff *x* and *y* have the same character traits.
Problems: (a) People's character traits can change over time. (b) Entails that more than one person can be you.

3e. *x* has the same mind as *y* iff *x* and *y* have the same beliefs, desires, and character.
Problems: Same as 3d.

3f. Some combination of 3c, 3d, and 3e. For example, *x* has the same mind as *y* iff *x* and *y* have the same character traits, *and* *x*'s quasi-memories are a subset of *y*'s quasi-memories or vice versa.
Problems: Same as 3c *and* 3d.

4. (Q1) You are a mind-body combination. (Q2) *x* has the same body as *y and x* has the same mind as *y*.
Problems: (a) Brain-transplant case. (b) Mind-transplant case? (c) What is required to have the same mind?

5. (Q1) You are a mind-brain combination. (Q2) *x* has the same brain as *y and x* has the same mind as *y*.
Problems: (a) Mind-transplant case? (b) Brain-of-Theseus case? (c) What is required to have the same mind? (See 3a-3f above.)

6. (Q2) The continuity theory:
   6a. *x* is the same person as *y* iff *x* and *y* are connected by a spatiotemporally continuous sequence of person-stages. (You are a 4-D spacetime worm.)
   Problem: (a) Fission case. Entails that more than one person could be you. (b) Impossibility of teleportation? (c) Brain-transplant case (implies that you cease to exist)? (d) Mind-transplant case (implies that you continue to exist in the original body).

   6b. *x* is the same person as *y* iff *x* and *y* are connected by a psychologically continuous sequence of mind-stages.
   Problems: (a) Entails that more than one person could be you. (b) Sleep, unconsciousness.

   6c. *x* is the same person as *y* iff *x* and *y* are connected by a spatiotemporally

and psychologically continuous sequence of person-stages.
Problems: Same as 6a (a, b, c) and 6b (a, b).

7. (Q2) The skeptical theory: there are no facts about personal identity. It's just a matter of convention/a semantic question.
   Problem: General craziness. Implies either (a) that you don't exist (because there's no such thing as personal identity), or (b) that you can make yourself immortal by just adopting an appropriate convention (because personal identity is conventional).

8. (Q2) The closest-continuer theory: Immediately after some change occurs, the entity that is *you* is the entity (if any) that (a) has at least some minimum level of similarity to you, and (b) is the best candidate for being you of all the then-existing entities.
   • The criteria for the best candidate include one or more of the above suggestions (e.g., continuity, having the same character, etc.), possibly a weighted combination.
   Problems: (a) Implies that identity is extrinsic. (b) Implies that identity is not symmetric. More than one thing can be (earlier stages of) you.

A General Problem:

• Most accounts of personal identity seek a *qualitative criterion* of personal identity.
• Any purely qualitative condition (and many not-purely-qualitative conditions) can be satisfied by more than one thing.
• But more than one thing cannot be you.
• Conclusion: There cannot be a purely qualitative criterion of personal identity.
  • Any criterion of personal identity (or criterion of being identical with you) must be such that it is logically impossible for more than one thing to satisfy it with respect to you.

Question:

• Why doesn't this work: Every person is identical with himself and nothing else?
• Because the problem is not to say when two people are identical. The problem is: when are two person-stages stages *of the same person*? (I.e., how must all of a person's stages be related to each other?)

Phil. 4360
Notes #34: Hume's Craziness

Traditional view:

There is a thing called "the self".

- We are directly introspectively aware of it. It is present in all conscious experience.
- It is not a mental state (like a belief, experience, desire, etc.). It is a single thing that has all of a person's mental states.
- Often said to be "simple"; also often said to be immaterial. (See Descartes.)

Hume: The self does not exist. Arguments:

A)

1. All ideas are copies of impressions.
2. There is no impression that the idea of self can be a copy of.
   a. Self is not any one impression; it is supposed to "underlie" all our impressions.
   b. We don't find anything common in all our impressions.
3. Therefore, there is no idea of self.

B)

1. We never observe the self by introspection.
   • We only observe particular perceptions, and the self is not one of these.
2. So the self either does not exist, or is just the collection of perceptions.
   • Therefore: I don't exist when I am asleep.

Conclusion: A person is "nothing but a bundle or collection of different perceptions."

Reid's Response:

> "[I]t is certainly a most amazing discovery, that thought and ideas may be without any thinking being. A discovery big with consequences which cannot easily be traced by those deluded mortals who think and reason in the common track. We were always apt to imagine, that thought supposed a thinker, and love a lover, and treason a traitor: but this, it seems, was all a mistake; and it is found out, that there may be treason without a traitor, and love without a lover [...]: or if, in these cases, ideas are the lover, the sufferer, the traitor, it were to be wished that the author of this discovery had farther condescended to acquaint us, whether ideas can converse together and be under obligations of duty or gratitude to each other [...] It seemed very natural to think that the *Treatise of Human Nature* required an author, and a very ingenious one too; but now we learn, that it is only a set of ideas which came together, and arranged themselves by certain associations and attractions." (35)

1. The following principle is *self-evident*. It neither needs proof nor can be proved.:
   An action cannot exist without a thing that acts. A property or state cannot exist without a thing that has it.

Applications of this principle:

Sensation cannot exist without a mind that senses. Thought cannot exist without a thinker. Further indications of this self-evidence: (a) the ease of making fun of Hume. (b) The fact that even Hume can't keep to his philosophy consistently.

2. What is wrong with Hume's argument:
   - Hume's argument rests on *the theory of ideas:* The view that the (direct) objects of awareness are always "ideas".
   - The history of "ideas" in philosophy:
     - Originally introduced to *explain awareness*, e.g., perception of physical objects.
       - Perception is held to consist of an object's causing an "image" of itself to appear in the subject's mind.
     - Led to rejection of secondary qualities. "Fire is not hot, nor snow cold, nor honey sweet."
     - Then led to rejection of primary qualities. (Berkeley)
     - Finally, led to rejection of the mind itself, leaving nothing but ideas in existence.
   - Theory of ideas should be rejected, for:
     a) No proof for the existence of "ideas" was ever given to begin with.
        • Why can't one just say we are aware of the actual, real objects?
     b) The theory of ideas has lots of absurd consequences.
     c) It undermines the very motivation for introducing "ideas" to begin with.

Phil. 4360
Notes #35: Review of Unit 4

<u>You should know these terms:</u>

- Free will
- Determinism
- Hard determinism
- Soft determinism
- Libertarianism
- Compatibilism/incompatibilism
- Problem of personal identity/Criterion of pers. identity
- Quasi-memory
- The "theory of ideas"

<u>Know what positions these people held (incl. which of the above positions they held):</u>

- Brand Blanshard
  - His explanation for the 'feeling of freedom'
  - The diff. between people & machines, how we are 'free.'
- Paul Edwards
  - Why we're not responsible for our actions/character.
- W. T. Stace
  - Why free-will is important.
  - What 'free' means & his main argument for this.
  - What 'could have' means.
  - Why we should punish people for things.
  - How FW requires determinism.
- Peter van Inwagen
  - Know his basic argument, incl.:
  - The operator 'N'
  - The 2 premises
  - The 2 rules of inference
- Mike Huemer
  - the proof of FW, its 3 main premises.
- J. R. Lucas
  - The presupposition of thought
  - What is Godel's Theorem
  - The Godel sentence, & how it helps refute a mechanistic conception of the mind.
- John Locke
  - on personal identity
- David Hume
  - on persons
- Thomas Reid
  - his criticisms of Locke & Hume

<u>Be able to argue against each of these theories of personal identity:</u>

- the body theory
- the brain theory
- the Cartesian dualist theory
- the memory theory
- the personality/beliefs/desires theory
- the spatiotemporal continuity theory
- the psychological continuity theory
- the conventionalist theory
- the closest-continuer theory
- qualitative criteria in general (the general argument against such)