Problem Set #1: Preliminaries

Philosophy 413: Formal Methods, Fall 2019

Each problem set for this course is divided into two sections—a handful of warm-up problems, followed by ten more challenging problems. Each group must submit *exactly ten solutions* that they wish to be graded for credit. Groups of undergraduates may submit solutions to warm-up problems; groups including Philosophy Ph.D. students may do these problems for fun but will not get credit for them.

1 Warm-up problems

1. Rewrite (a) and (b) using corner-quote notation. Rewrite (c) and (d) using regular quotes and the concatenation sign:[1]

a. $\phi \frown$ '+' $\frown \psi \frown$ '=' $\frown \psi$
b. '∀' $\frown \alpha \frown \psi$
c. For any sentence ϕ, the sentence $\ulcorner \phi \supset \phi \urcorner$ is longer than ϕ.
d. If the sentence ϕ is a logical contradiction, then the sentence $\ulcorner \phi \supset \neg\phi \urcorner$ is true.

2. Add quotes and/or corner quotes to the following sentences in order to transform them into true sentences, or explain why no such alteration is possible. The sentences that you produce should not contain any vacuous quantification, i.e. quantifiers that do not bind any variable in their scope. Give a "best" solution, i.e. one that involves as few additional symbols as possible.

a. This very sentence is made up of three words.
b. It takes longer to read this problem set than it takes to read this problem set.
c. To refer to the president's name, use the president's name.
d. For several definite descriptions D, Winston Churchill said We shall fight on D.
e. For several sentences S, Martin Luther King Jr. said I have a dream that S.

3. Explain why the following are true or false, using everyday English as much as possible.[2]

a. $\{x : x \in B\} = \{B\}$
b. $\{x : x = a\} = \{a\}$
c. $\{x : x \in B\} = B$
d. $\{x : x \text{ is green}\} = \{y : y \text{ is green}\}$

4. For each of the following, say whether the relation is reflexive, whether it is symmetric, and whether it is transitive. Provide a counterexample in order to demonstrate each negative answer.[3]

a. the relation of being identical (with)
b. the relation of being far (from)
c. the relation of being perpendicular (to)
d. the relation of being approximately equal (to)

1. Examples (1a)–(1b) are due to MacFarlane 2016.
2. Examples (3b)–(3d) are due to Heim & Kratzer 1998, p.9–10.
3. Examples (4b)–(4d) are due to McCawley 1993, p.160

2 Additional problems

1. In their legendary "Who's on First?" routine, Abbott and Costello sometimes use the expression 'who' and sometimes mention it. Transcriptions of the routine usually do not follow any careful use-mention conventions. Watch a video recording of the routine,[4] and identify at least one instance where a speaker intends to use 'who', an instance where a speaker intends to mention it, and an occasion where there is confusion about whether the term is being used or mentioned. Defend your answers.

2. Add quotes and/or corner quotes to the following sentence in order to transform it into a true sentence, or explain why no such alteration is possible. The sentence that you produce should not contain any vacuous quantification, i.e. quantifiers that do not bind any variable in their scope. Give a "best" solution, i.e. one that involves as few additional symbols as possible.[5]

> For every sentence ϕ, the last word of the last word of ϕ is polysyllabic is polysyllabic.

3. Follow the same instructions as in (2) for the following sentence:

> The first letter of the Greek alphabet is α is satisfied by an object β only if β is identical with α.

4. Follow the same instructions as in (2) for the following sentence:[6]

> For every sentence ϕ, ϕ implies $\neg\phi$ implies ϕ implies ϕ and $\neg\phi$.

5. Follow the same instructions as in (2) for the following pair of sentences:

a. The last word of #5a is obscene.
b. The last word of #5a is obscene.

6. Say that a sentence is *incorrigible* just in case there is no way of supplying quotes and/or corners such that the result is neither false nor senseless. It appears that:

(1) #5a is incorrigible.
(2) #5b is not incorrigible.
(3) #5a is identical with #5b.

But at least one of (1), (2), and (3) must be false. Which is false, and why?

7. Explain why the following are true or false, using everyday English as much as possible.[7]

a. $\{x : x \in \{y : y \in B\}\} = B$
b. $\{x : \{y : y \text{ likes } x\} = \varnothing\} = \{x : \{x : x \text{ likes } x\} = \varnothing\}$

4. A fair use copy is available here: `https://www.youtube.com/watch?v=kTcRRaXV-fg`.
5. Problems (2)–(6) in §2 of this problem set are due to Cartwright 1987, p.257–8.
6. In the context of this problem, the English term 'implies' is synonymous with the logical term '$\supset$' as it appears in the context of Warm-up Problem 1d.
7. These examples are due to Heim & Kratzer 1998, p.10.

8. For each of the following, say whether the relation is reflexive, whether it is symmetric, and whether it is transitive. Provide a counterexample in order to demonstrate each negative answer.[8]

a. the relation of being adjacent (to)
b. the relation of being the brother (of)
b. the relation defined on pairs of integers that holds just in case both integers are even
c. the relation defined on natural numbers of having a common factor greater than 3

9. Consider the following passage from Lewis Carroll's *Through the Looking Glass*:

> "You are sad," the Knight said in an anxious tone: "Let me sing you a song to comfort you."
>
> "Is it very long?" Alice asked, for she had heard a good deal of poetry that day.
>
> "It's long," said the Knight, "but it's very, *very* beautiful. Everybody that hears me sing it—either it brings the *tears* to their eyes, or else—"
>
> "Or else what?" said Alice, for the Knight had made a sudden pause.
>
> "Or else it doesn't, you know. The name of the song is called '*Haddocks' Eyes*.'"
>
> "Oh, that's the name of the song, is it?" Alice said, trying to feel interested.
>
> "No, you don't understand," the Knight said, looking a little vexed. "That's what the name is *called*. The name really *is* '*The Aged Aged Man*.'"
>
> "Then I ought to have said 'That's what the *song* is called'?" Alice corrected herself.
>
> "No, you oughtn't: that's quite another thing! The *song* is called 'Ways And Means': but that's only what it's *called*, you know!"[9]
>
> "Well, what *is* the song, then?" said Alice, who was by this time completely bewildered.
>
> "I was coming to that," the Knight said. "The song really *is* '*A-sitting On A Gate*': and the tune's my own invention."
>
> . . . She stood and listened very attentively, but no tears came into her eyes.[10]

If necessary, add quotation marks to (a) and (b) in order to restore truth:

a. *Haddock's Eyes = The Aged Aged Man*
b. *The Aged Aged Man = Ways and Means*

Also, what is the first word of the song? Explain your answer.[11]

10. In "Slurring Words," Luvell Anderson and Ernie Lepore develop a *silentist* view of slurs, according to which it can be inappropriate to use the quote-name of a slur in order to mention that slur.[12] After reviewing the arguments in §8 of their paper, address the following question: under what circumstances could mentioning a slur share some of the moral features that make it wrong to use that slur? Explain and defend your answer.[13]

---

8. Examples (8a) and (8b) are due to McCawley 1993, p.160
9. See Fara 2011 for a thought-provoking philosophical investigation of appellative-'called' constructions.
10. Carroll 1871, p.205–6
11. This problem is taken from the first homework assignment for the Fall 2013 MIT Proseminar in Philosophy.
12. Anderson & Lepore 2013, p.38
13. A prompt for your discussion: suppose that in certain circumstances, it would be morally inappropriate to share a photograph of a homicide victim. Then it might also be morally inappropriate to share a photograph *of a photograph* of the victim. Under what circumstances might representations of representations share morally relevant features of those representations?

Problem Set #2: Formal Semantics

Philosophy 413: Formal Methods, Fall 2019

1 Warm-up problems

1. Consider the tree structure on p.33 of Heim & Kratzer 1998. As specified in the text, assume that the preposition **to** is semantically vacuous. Identify the semantic type of each of the following:

a. **Ann**
b. **Jacob**
c. **to Jacob**
d. **introduces Maria to Jacob**
e. **introduces Maria**
f. **Maria**
g. **introduces**

2. Consider the following passage from Lewis Carroll's *Through the Looking Glass*:

> 'I see nobody on the road,' said Alice.
> 'I only wish *I* had such eyes,' the King remarked in a fretful tone. 'To be able to see Nobody! And at that distance, too! Why, it's as much as *I* can do to see real people, by this light!'[14]

What mistake is the King making here? Refer to multiple semantic types in explaining your answer.

3. Draw a tree diagram representing the following sentence, record the semantic values of each terminal node of the tree, and use those semantic values to derive the truth conditions for the sentence:

Some philosopher smokes.

4. Follow the same instructions as in (3) for the following sentence:

Clark Kent is Superman.

Feel free to assume any reasonable semantic value for **is**, as long as it yields the correct truth conditions.

2 Additional problems

1. Consider the following structurally ambiguous sentences:

a. **Jones told the person that Smith liked the story.**
b. **Jones listened to the person with the radio.**
c. **Jones put the present in the box under the tree.**
d. **Jones said her mother died yesterday.**

For each sentence, resolve its ambiguity by drawing trees corresponding to each of its readings.

14. Carroll 1871, p.186

2. Draw syntactic trees for the following garden-path sentences:

a. **Cotton shirts are made from comes from India.**
b. **Cheese mice cats catch love stinks.**
c. **The horse raced past the barn fell.**

3. Draw a tree diagram representing the following sentence, record the semantic values of each terminal node of the tree, and use those semantic values to derive the truth conditions for the sentence:

Every day is Christmas.

As in Warm-Up Problem 4, you may assume any reasonable semantic value for **is** that yields the correct truth conditions.

4. Follow the same instructions as in (3) for the following sentence:

Lois hopes that Clark Kent is Superman.

5. Develop a theory about the semantic value of **or** as it is used in the sentence **John or Mary smokes**. Then brainstorm an alternative hypothesis about its semantic value. Give at least one argument for the claim that your preferred theory is better than the alternative.

6. In class, we decided that the expression **a** in the sentence **Fido is a dog** is semantically vacuous. The name **Fido** picks out something of type e, while the predicate **dog** picks out something of type et. In other words, the sentence has the logical form represented in a first-order language by an expression such as 'Dog(Fido)'. Hence neither **is** nor **a** contributes anything to the semantic value of the sentence.

Consider the expression **a** in the sentence **A dog barked**. Is this occurrence of **a** also semantically vacuous? If it is not vacuous, what semantic type should it have? Explain and defend your answer.

7. The Predicate Modification rule does not work well for certain combinations of adjectives. For example, use the rule in order to work out the semantic value of the phrase **fake diamond** in the sentence **That is a fake diamond**. Why is the result of your calculation *not* an intuitive semantic value for this phrase?

8. Develop a hypothesis about the semantic type of numerals, such as **three** in the sentence **Three students smoke**. Give at least one argument in defense of your proposal.

9. Compute the truth conditions of **In the world of Sherlock Holmes, Holmes is quick and Watson is slow** using the extensional meaning for **and** introduced on p.5 of von Fintel & Heim 2005, taking **in the world of Sherlock Holmes** as a primitive constituent of the sentence. Then compute the truth conditions using the intensional meaning introduced on p.11. Comment on your results.

10. Develop a hypothesis about the semantic value of **doubts**. Use your hypothesis to derive the correct truth conditions for the following sentence: **Lois doubts that Clark Kent is Superman**. Comment on any property you observe that distinguishes the semantic value of doubts from the semantic values of attitude verbs such as **believes**, **hopes**, **wants**, **desires**, **perceives**, **remembers** and **knows**.

Problem Set #3: Modal Logic

Philosophy 413: Formal Methods, Fall 2019

1 Warm-up problems

1. Evaluate which of the following make sense, and which are mis-uses of modal logic vocabulary:

   a. That formula is true in that model.
   b. This model is a countermodel to the claim that that formula is valid in the class of all models.
   c. The class of all models with transitive accessibility relations is valid.
   d. The accessibility relation of this frame has some unusual properties.
   e. The interpretation function of this frame has some unusual properties.
   f. That formula is valid in the class of all models with reflexive accessibility relations.
   g. This formula is valid in every world of that model.

Explain your answers in detail, including examples of correct uses of vocabulary where relevant.

2. Consider the model consisting of exactly one world which cannot see any other worlds, where the propositional variable P is true at that world. State some claim of the following form: *such-and-such formula is valid in the class of models with such-and-such property*, such that the cheesy model just mentioned serves as a countermodel to your claim.

3. State a claim with the same form such that the cheesy model is *not* a countermodel to your claim.

2 Additional problems

1. In May 2011, security researchers filed an FTC complaint alleging that Dropbox deceived users about the security of its services. *WIRED* reporter Ryan Singel summarized the complaint as follows:

> Dropbox employees can see the contents of a user's storage, and can turn over the nonencrypted files to the government or outside organizations when presented with a subpoena. . .
>
> Up until April 13, the site promised this:
>
> > *Dropbox employees aren't able to access user files, and when troubleshooting an account, they only have access to file metadata (filenames, file sizes, etc. not the file contents).*
>
> Now the site says:
>
> > *Dropbox employees are prohibited from viewing the content of files you store in your Dropbox account, and are only permitted to view file metadata (e.g., file names and locations).*
>
> The complaint alleges that at least two of Dropbox's competitors, SpiderOak and Wuala, make security promises similiar to those of Dropbox, but actually can't get at the data because they don't hold the encryption keys. . . That, according to the complaint, lets Dropbox promise total security without paying the costs, while putting its competitors at a disadvantage.[15]

15. Singel 2011, para. 9–10; my emphasis.

A spokesperson for Dropbox argued that the company never misled clients, saying:

> In our help article we stated 'Dropbox employees aren't able to access user files.' That means that we prevent such access. . . [using] strict policy prohibitions.

Let us grant that 'aren't able to' and 'are prohibited from' should be formalized using the same modal logic vocabulary, namely '$\neg\Diamond$'. Are these expressions synonymous? Why or why not? Suppose that you are writing a brief to guide the FTC decision about whether Dropbox misled clients. What are some relevant facts about modal language that you would mention, and why are they relevant?

2. For each of the following formulas, if the formula is S-valid for the given formal system S, then prove its S-validity. Otherwise, construct a countermodel to demonstrate its S-invalidity.

a. in the formal system **K**: $\Box p \supset \Box\Box p$
b. in the formal system **T**: $\Box p \supset p$
c. in the formal system **S5**: $\Box p \supset p$

3. For each of the following formulas, if the formula is D-valid, then prove its D-validity. Otherwise, construct a countermodel to demonstrate its D-invalidity.

a. $\Box p \supset \Box\Diamond p$
b. $(\Box p \wedge \Box(\neg p \vee q)) \supset \Diamond q$

4. For each of the following formulas, if the formula is S4-valid, then prove its S4-validity. Otherwise, construct a countermodel to demonstrate its S4-invalidity.

a. $(\Box p \wedge \Box(\neg p \vee q)) \supset \Diamond q$
b. $\Diamond\Diamond(p \wedge q) \supset \Diamond q$

5. Prove that the following variants of the S4 axiom are logically equivalent:

a. $\Box P \supset \Box\Box P$
b. $\Diamond\Diamond Q \supset \Diamond Q$

6. Prove that the formula $\ulcorner\Box\Diamond\alpha\urcorner$ is not a theorem of **K**, where α is any formula of propositional modal logic. You may assume that **K** is sound with respect to the class of all models, which entails that if the negation of a formula is true in some world in some model, then that formula is not a theorem of **K**.

7. Show that the rule 'if $\vdash \Box\alpha \supset \Box\beta$, then $\vdash \alpha \supset \beta$' preserves validity in **K**.

8. Show that the rule 'if $\vdash \Box\alpha \supset \Box\beta$, then $\vdash \alpha \supset \beta$' does not preserve validity in **T**.

9. Which of the formal systems we have studied—**K**, **D**, **T**, **B**, **S4**, or **S5**—seems best suited to model deontic necessity, i.e. moral obligation? Which seems best suited to model epistemic necessity? In both cases, justify your answers. If none of the systems we have studied seems adequate, explain what axioms are missing.

10. Consider the following *Idle Argument*, as found in Origen, *Against Celsus* (*Cels* II 20):

> If it is fated that you will recover from this illness, then, regardless of whether you consult a doctor or you do not consult [a doctor] you will recover. But also: if it is fated that you won't recover from this illness, then, regardless of whether you consult a doctor or you do not consult [a doctor] you won't recover. But either it is fated that you will recover from this illness or it is fated that you won't recover. Therefore it is futile to consult a doctor.[16]

Suggest a way of formalizing this argument using our formal modal logic vocabulary, and assess whether your argument is valid.

16. Translation from Bobzien 1998, p.182.

Problem Set #4: Conditionals

Philosophy 413: Formal Methods, Fall 2019

1 Warm-up problems

1. Describe a realistic instance of the or-to-if inference, and explain why our intuitive judgment about the inference poses a problem for Stalnaker 1968.

2. State two distinct challenges for the material conditional analysis of the indicative conditional, and give an example of an indicative conditional that simultaneously illustrates both of them.

3. Give an example that intuitively seems to demonstrate that the following rule of inference is invalid:

If A, would B.
If B, would C.
Therefore: if A, would C.

How would advocates of the dynamic strict conditional analysis explain the apparent invalidity of your example inference?

4. Lewis 1973 defends the following *Duality Thesis* relating might- and would-counterfactuals:

$$p \mathbin{\Box\!\!\rightarrow} q \equiv \neg(p \mathbin{\Diamond\!\!\rightarrow} \neg q)$$

Given this thesis and the principle of Conditional Excluded Middle for 'would' counterfactuals, prove that '$p \mathbin{\Diamond\!\!\rightarrow} q$' entails '$p \mathbin{\Box\!\!\rightarrow} q$'.

5. Is the equivalence that you derived in (4) intuitively acceptable? If it isn't, then which premise should we reject? Give arguments in support of your answer(s).

2 Additional problems

1. Construct a model that demonstrates that Stalnaker's semantics for the subjunctive conditional does not validate Antecedent Strengthening.

2. As noted in Warm-Up Problem 3, the following rule of inference is intuitively invalid:

If A, would B.
If B, would C.
Therefore: if A, would C.

Consider the following related rule of inference:

If A, would B.
If A and B, would C.
Therefore: if A, would C.

Can one simply extend intuitive counterexamples to transitivity to demonstrate that this rule of inference is also intuitively invalid? Explain your answer.

3. Do variably strict conditional analyses validate the rule of inference introduced in (2)? If not, construct a model that demonstrates that the semantics in Stalnaker 1968 or the semantics in Lewis 1973 does not validate this inference rule. If so, give a proof that one of these analyses does validate the rule.

4. Is the following rule of inference intuitively valid?

If A or B, would C.
Therefore: if A, would C.

If the rule seems valid, give an intuitive explanation for why it is valid. If the rule seems invalid, give an example that illustrates its invalidity.

5. Comment on the importance of your answer to (4) for the debate between strict conditional, variably strict conditional, and dynamic strict conditional analyses of the subjunctive conditional. As part of your answer, you should determine whether the inference is semantically valid according to each of these three analyses, and also whether there are any pragmatic features relevant to its intuitive (in)validity.

6. On the variably strict semantics for the subjunctive conditional, what is the relationship between the subjunctive conditional '$p \mathrel{\Box\!\!\rightarrow} q$' and the necessitated material conditional '$\Box(p \supset q)$'? Are they equivalent, independent, or does one strictly entail the other? Give examples and/or arguments to support your answer.

7. Williams 2008 observes that the indicative analog of a Sobel sequence is felicitous:

> If Sophie went to the parade, she saw Pedro.
> But if Sophie went to the parade and got stuck behind a tall person, she did not see Pedro.

Meanwhile, the indicative analog of a reverse Sobel sequence is infelicitous:

> If Sophie went to the parade and got stuck behind a tall person, she did not see Pedro.
> #But if Sophie went to the parade, she saw Pedro.

Here is a brief description of his response to this observation from Moss 2012:

> Williams accounts for these data by adopting a variant of the strict conditional analysis for indicative conditionals, according to which the domain of the necessity modal is the *context set*: the set of worlds compatible with what is treated as true for purposes of conversation. He says that 'if p, q' is true just in case all p worlds in the context set are q worlds. Like von Fintel and Gillies, Williams then adds another component to the meaning of a conditional: Williams says that 'if p, q' presupposes that the context set contains some p worlds. (565)

Fill in the details of this argument. Why do the two assumptions introduced by Williams suffice to explain the judgments cited above?[17]

8. Consider the following subjunctive conditional embedded under an epistemic probability operator:

> It is .5 likely that if the coin had been flipped, it would have landed heads.

What is your intuitive judgment about this conditional—is it true? False? Acceptable? Unacceptable? How does your intuitive judgment bear on the debate between the variably strict conditional analyses defended by Stalnaker 1968 and Lewis 1973? Brainstorm what Stalnaker and Lewis might say about your judgment, and determine whether they are equally well-positioned to predict or accommodate it.

9. Follow the same instructions as in (8) for sentence (b) of the following monologue:

> a. If the test was really easy and almost all the students passed it, then Jones might have failed, but Smith and Brown couldn't have failed.
> b. But if exactly three students passed the test, then every student might have failed.

Assume that it is clear from context that the possibility modals in (a) and (b) are epistemic. Again, what is your intuitive judgment about the sentence, and how does this bear on the debate between different versions of the variably strict analysis of subjunctives?

A hint to get you started with this question: Could the epistemic possibility modal 'might' in (b) take scope over the subjunctive conditional? Why or why not?[18]

10. In a widely cited manuscript, Hájek 2007 argues that most counterfactuals are false:

> In an indeterministic world such as ours appears to be, lotteries—in a broad sense—abound. . . Two billiard balls colliding may approximate a deterministic system, but even they are not immune from quantum mechanical indeterminism. One ball might spontaneously tunnel through the other, or to China, or to the North Star—incredibly unlikely, to be sure, but possible. Thus, I cannot truly say 'if the cue ball were to hit the 8 ball, the 8 ball would begin rolling'. Or again, whenever I jump in the air, there is a minute chance that I will not come down—I might vaporize instead, for the chance of that happening is non-zero. Thus, I cannot truly say 'if I were to jump, I would come down'. (7)

Assume that Hájek is right that our world is indeterministic, and moreover that there is some positive objective chance that a jumping person will vaporize rather than come down. Do you agree with the conclusion of the above passage? What is the difference, if any, between the various positive-chance outcomes of a lottery and the various positive-chance outcomes of an event such as jumping?[19]

---

17. You may consult Williams 2008 for further background and explanation, but your answer must be in your own words.
18. For further background, see the discussion of quantifiers and epistemic modals in §3 of Swanson 2011.
19. You may consult Hájek's manuscript and even cite it in your answer to this question, but your arguments should be the result of your own reflection on his argument.

Problem Set #5: Probability Theory

Philosophy 413: Formal Methods, Fall 2019

1 Warm-up problems

1. A family has two children. You randomly run into one of the two, and learn that she is a girl. What is the conditional probability that both children are girls?[20]

2. An urn contains only red and blue marbles. Two marbles are chosen without replacement. The probability of choosing a red marble and then a blue marble is .2. The probability of choosing a blue marble on the first draw is also .2. What is the conditional probability of drawing a blue marble, given that the first marble drawn was red?

3. Derive $f(p \wedge q)$ as a function of $f(p)$ and $f(q)$, where f is a probability function with respect to which p and q are independent propositions.[21]

4. During the Vietnam War, a fighter jet made a non-fatal strafing attack on a US aerial reconnaissance mission at twilight. Both Cambodian and Vietnamese jets operate in the area. You know the following:

   a. The US pilot identified the fighter as Cambodian. The pilot's aircraft recognition capabilities were tested under appropriate visibility and flight conditions. When presented with a sample of fighters, half with Vietnamese markings and half with Cambodian, the pilot made correct identifications 80 percent of the time and erred 20 percent of the time.
   b. 85 percent of the jet fighters in that area are Vietnamese; 15 percent are Cambodian.

What is the probability that the fighter was Cambodian rather than Vietnamese?[22] Explain your answer by identifying and computing relevant conditional probabilities.

A hint for how to double-check your work on this problem: Suppose that 100 planes are flying in the area. On average, how many will be Vietnamese? Of those planes, how many will be correctly identified as Vietnamese, and how many will be misidentified as Cambodian? How many planes are going to be Cambodian? Of those planes, how many will be identified as Vietnamese, and how many as Cambodian?

2 Additional problems

1. Can a probability function f be such that both of the following hold for some propositions P and Q?

   a. $f(P) = f(Q) = .5$
   b. $f(\neg P \vee \neg Q) = .8$

Give an example to demonstrate that this is possible, or give a proof to demonstrate that it is not.

20. Assume that you are equally likely to run into either child, and that which one you run into has nothing to do with gender.
21. Recall that propositions p and q are independent with respect to f just in case $f(p) = f(p|q)$.
22. This problem is taken directly from *Psychology of Intelligence Analysis*, published by the United States Central Intelligence Agency; see Heuer 1999, p.157–8.

2. Can your credence in a proposition that is compatible with your new information decrease when you update by conditionalization? In other words, can we have $f(q|p) < f(q)$ when p and q are not mutually exclusive? If so, give an example. If not, prove that this never happens.

3. Can your credence in a proposition that entails your new information decrease when you update by conditionalization? In other words, can we have $f(q|p) < f(q)$ when q entails p? If so, give an example. If not, prove that this never happens.

4. Prove that conditionalizing on new information preserves ratios of credences between propositions that entail that information. In other words, prove that if q and r each entail p, $f(q|p)/f(r|p) = f(q)/f(r)$.

5. Prove that conditionalizing on some evidence E preserves the conditional probabilities that you assign to propositions conditional on that evidence, i.e. $f(p|E)$.

6. A jailer is in charge of three prisoners, called A, B, and C. They learn that two of the prisoners have been chosen at random to be executed the following morning; the jailer knows which two they are, but will not tell the prisoners. Prisoner A realizes that his chance of survival is 1/3. But he says to the jailer, "Tell me the name of one of the other two who will be executed; this will not give me any information, because I know that at least one of them must die, and I know neither of them." The jailer agrees and tells A that B is going to be executed.[23]

What credence should A assign to the proposition that he is going to be executed *before* the jailer says that B is going to be executed? What credence should he assign to it *after*? Explain your answers.

7. Give an example to demonstrate that if you compute the arithmetic mean of two probability measures and then conditionalize them on a proposition, you do not always get the same result as if you had first conditionalized each measure on that same proposition and then computed their arithmetic mean.[24]

8. Let us say that propositions p and q are *conditionally independent given proposition E relative to probability function f* when p and q are independent relative to the probability function that results from conditionalizing f on E. Can two propositions be independent relative to f without being conditionally independent given E relative to f? Conversely, can propositions be conditionally independent without being independent? Give arguments or examples to support your answers to both questions.

9. Is the relation of *being independent relative to probability function f* transitive? Defend your answer.

10. A family has two children. Find the probability that both children are girls, given that at least one of the two is a girl who was born in winter. Assume that the four seasons are equally likely and that gender is independent of season, i.e. that knowing the gender gives no information about the probabilities of the seasons, and vice versa.[25]

23. This problem is originally due to Gardner 1959. The above statement of the problem is taken from Kelly 1994.
24. The mean of two functions f and g is defined as the function h such that for every input x, we have $h(x) = .5f(x) + .5g(x)$.
25. This problem is taken from Blitzstein & Hwang 2019, p.32.

Problem Set #6: Decision Theory

Philosophy 413: Formal Methods, Fall 2019

1 Warm-up problems

1. Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?[26]

Say what credences you should have before and after the host opens one of the doors you do not open. You may consult discussions of the Monty Hall problem as you work out your answer, but the solution that you submit should be explained entirely in your own words.

2. Consider the following example described by Resnik 1987:

> Danny, who has been injured by Manny in an automobile accident, has applied to Manny's insurance company for compensation. The company has responded with an offer of $10,000. Danny is considering hiring a lawyer to demand $50,000. If Danny hires a lawyer to demand $50,000, Manny's insurance company will respond by either offering $10,000 again or offering $25,000. If they offer $25,000, Danny plans to take it. If they offer $10,000, Danny will decide whether or not to sue. If he decides not to sue, he will get $10,000. If he decides to sue, he will win or lose. If he wins, he can expect $50,000. If he loses, he will get nothing." (19)

Construct a decision tree representing this situation.

3. Pascal argues in *Pensées*, §233 that it is better to act so as to cultivate the belief in yourself that God exists, since if God exists, you will thereby gain an "eternity of life and happiness" rather than misery, whereas if God does not exist, your decision makes no difference. Construct a decision table that represents this argument, and use it to explain Pascal's reasoning.

4. Recall the game *Chicken*: two players are riding bikes straight at each other. Each player would most prefer to continue riding while the other swerves, would next prefer that both players swerve, would next prefer that he himself "chickens out" while the other player is brave, and would last prefer that neither swerves and both players die. Is this game a variant of *Prisoner's Dilemma*? Explain your answer.

2 Additional problems

1. Create an example similar to the example in Warm-up Problem 1, but which involves *one million* doors instead of three, and set up the example so that the correct answer to your revised example is more intuitive than the correct answer to the original problem.

26. This statement of the Monty Hall problem is taken from vos Savant 1990, para. 1.

2. Reformulate the decision in Warm-up Problem 2 using a decision table, rather than a decision tree.[27]

3. Assess the Pascalian argument described in Warm-up Problem 3. Do you agree with the conclusion? If you disagree, explain what principle supports your own reasoning about the wager. If you agree, respond to the strongest objection that you can imagine someone bringing against his argument.

4. Consider a variant of *Chicken* where the players have the same preferences as the players described in Warm-Up Problem 4, except that each player actually prefers that neither swerves and both die heroically, than that he himself "chickens out" and lives in humiliation. Is the resulting game a variant of *Prisoner's Dilemma?* Explain your answer.

5. In 1998, Sally Clark was tried for murder after two of her sons died shortly after birth. During the trial, an expert witness for the prosecution testified that the probability of a newborn dying of sudden infant death syndrome (SIDS) was 1/8500, so the probability of two deaths due to SIDS in one family was $(1/8500)^2$, or about one in 73 million. Therefore, he continued, the probability of Clark's innocence was one in 73 million.[28]

Was the prosecutor's reasoning sound? Explain why or why not.

6. Consider the following hypothetical from Blitzstein & Hwang 2019:

> A woman has been murdered, and her husband is put on trial for this crime. Evidence comes to light that the defendant had a history of abusing his wife. The defense attorney argues that the evidence of abuse should be excluded on grounds of irrelevance, since only 1 in 10,000 men who abuse their wives subsequently murder them.
>
> Suppose that the defense attorney's 1-in-10,000 figure is correct, and further assume the following facts: 1 in 10 men commit abuse against their wives, 1 in 5 married women who are murdered are murdered by their husbands, and 50% of husbands who murder their wives previously abused them. Also, assume that if the husband of a murdered wife is not guilty of the murder, then the probability that he abused his wife reverts to the unconditional probability of abuse. (75)

Let A be the event that the husband commits abuse against his wife, and let G be the event that the husband is guilty. The defense argues that $P(G|A) = 1/10,000$, so guilt is still extremely unlikely conditional on a previous history of abuse. Is this reasoning sound? Explain why or why not.

7. Churchill and Hepburn are corporate executives at companies that make cigars and cigarettes, respectively. They are arguing about what sort of smoking is worse for your health. Hepburn claims that cigars are more dangerous, citing a recent study that found the overall mortality rate among cigar smokers to be higher than that among cigarette smokers. Meanwhile, Churchill claims that cigarettes are more dangerous. In defense of his claim, he cites a recent study that found that for any given decade, the mortality rate among cigar smokers born in that decade is lower than the mortality rate among cigarette smokers born in that decade.

---

27. This problem is taken from Resnik 1987, p. 19.
28. This description of the Clark case is taken from Blitzstein & Hwang 2019, p.66.

Are Churchill and Hepburn citing inconsistent studies? Who has the more convincing argument?

8. Create a pair of bets that constitute a synchronic Dutch book for an agent who has .3 credence in a proposition and .3 credence in its negation.

9. Suppose I offer you the following game: we toss a fair coin. If it lands heads, I pay you $2. If it lands tails, I toss the coin again. If it lands heads then, I pay you $4. If it lands tails, I toss the coin again. I keep tossing the coin until it lands heads, doubling the amount that I would pay you for a heads result each time. How much should you be willing to pay me to play this game? Explain your answer.

10. Propose a short-answer question that you would like to see on the final exam for this class, and explain what would be required to get full and partial credit for an answer to your question.

References

Anderson, Luvell & Ernie Lepore. 2013. *Noûs*, vol. 47 (1): 25–48.

Blitzstein, Joseph K. & Jessica Hwang. 2019. *Introduction to Probability*. CRC Press, Boca Raton, FL, second edn.

Bobzien, Suzanne. 1998. *Determinism and Freedom in Stoic Philosophy*. Oxford University Press, Oxford.

Carroll, Lewis. 1871. *Through the Looking-Glass*. Bantam Classics, New York. First Printing edition (June 1, 1984).

Cartwright, Richard. 1987. *Philosophical Essays*. MIT Press, Cambridge.

Fara, Delia Graff. 2011. "You Can Call Me 'Stupid',. . . Just Don't Call Me Stupid." *Analysis*, vol. 71 (3): 492–501.

von Fintel, Kai & Irene Heim. 2005. "Intensional Semantics Lecture Notes." Ms., MIT. Available online at `http://semantics-online.org/advsem/IntensionalSemantics.pdf`.

Gardner, Martin. 1959. "Mathematical Games: Problems Involving Questions of Probability and Ambiguity." *Scientific American*, vol. 201 (4): 174–872.

Hájek, Alan. 2007. "Most Counterfactuals are False." Ms., Research School of the Social Sciences, Australian National University.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Blackwell, Malden, MA.

Heuer, Richards J. 1999. *Psychology of Intelligence Analysis*. Central Intelligence Agency Center for the Study of Intelligence, Washington DC.

Kelly, D. G. 1994. *Introduction to Probability*. Macmillan, New York.

Lewis, David K. 1973. *Counterfactuals*. Basil Blackwell Ltd., Malden, MA.

MacFarlane, John. 2016. "Substitutional Quantifiers." Ms., University of California, Berkeley. Available online at `http://docplayer.net/43546470-Substitutional-quantifiers.html`.

McCawley, James D. 1993. *Everything That Linguists Have Always Wanted to Know about Logic*. University of Chicago Press, Chicago.

Moss, Sarah. 2012. "On the Pragmatics of Counterfactuals." *Noûs*, vol. 46 (3): 561–86.

Resnik, Michael. 1987. *Choices: An Introduction to Decision Theory*. University of Minnesota Press, Minneapolis.

vos Savant, Marilyn. 1990. "Ask Marilyn." *Parade Magazine*. September 9, 1990.

Singel, Ryan. 2011. "Dropbox Lied to Users About Data Security, Complaint to FTC Alleges." *WIRED*. May 13, 2011. Retrieved from `https://www.wired.com/2011/05/dropbox-ftc/`.

Stalnaker, Robert. 1968. "A Theory of Conditionals." In *Ifs: Conditionals, Belief, Decision, Chance, and Time*, William L. Harper, Robert Stalnaker & Glenn Pearce, editors, 41–55. D. Reidel Publishing Company, Dordrecht.

Swanson, Eric. 2011. "How Not to Theorize about the Language of Subjective Uncertainty." In *Epistemic Modality*, Andy Egan & Brian Weatherson, editors, 249–69. Oxford University Press, Oxford.

Williams, J. Robert G. 2008. "Conversation and Conditionals." *Philosophical Studies*, vol. 138 (2): 211–23.