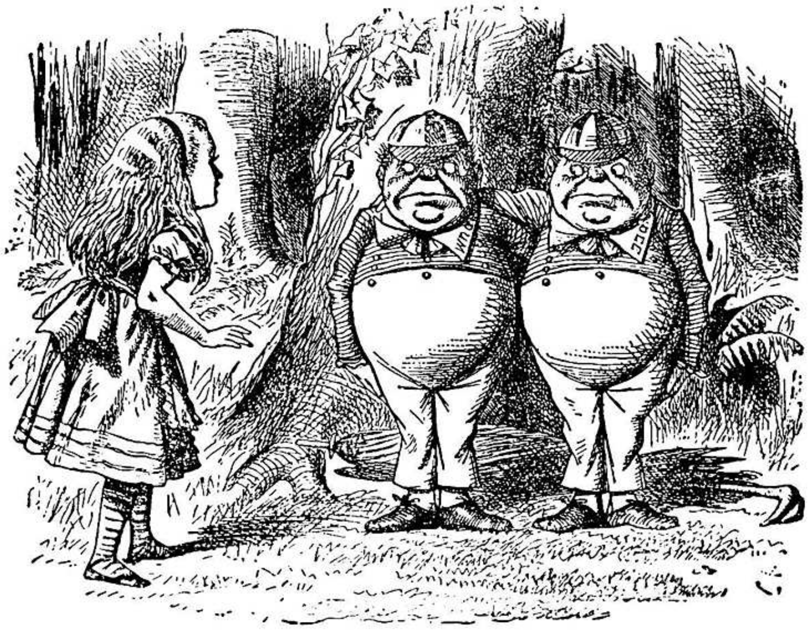
What Kinds of Liars and Truth-tellers Exist, and How Can They Be Identified Through Interaction?



University of Auckland

Philosophy 315

Emma Perry, Forrest Axford, Jason Cairns

October 22, 2018

Contents

Abstract

Truth-telling and lying come in many different forms - we intend to develop a taxonomy through the exploration of some intuitive and counter-intuitive types, and determine how an agent's category can be identified through interaction.

A Note on Belief

Its worth noting that the way in which one determines belief can have drastic impacts on logical systems of liars and truth tellers which consider belief.[1] If a Lockean conception of belief is used, τ being set to anything less than 0.5 can result in the speaker being both a liar and truth-teller simultaneously if using a system in which liars are those who say something they believe not to be true and truth-tellers are those that say something they believe to be true.[2]

1. Krister Segerberg, John-Jules Meyer, and Marcus Kracht, "The Logic of Action," in *The Stanford Encyclopedia of Philosophy*, Winter 2016, ed. Edward N. Zalta (Metaphysics Research Lab, Stanford University, 2016).

2. Vincent Hendricks and John Symons, "Epistemic Logic," in *The Stanford Encyclopedia of Philosophy*, Fall 2015, ed. Edward N. Zalta (Metaphysics Research Lab, Stanford University, 2015).

Chapter 1

Basic Types

	Description	Logical formulation
1	Truth-telling as saying something which is true in reality	$(T\text{♪}\psi \rightarrow \psi)$
2	Lying as saying something which is not the case in reality	$(F\text{♪}\psi \rightarrow \neg\psi)$
3	Truth-telling as saying something sincerely	$(T\text{♪}\psi \rightarrow B_T\psi)$
4	Lying as saying something insincerely	$(F\text{♪}\psi \rightarrow \neg B_L\psi)$
5	Lying as saying something believed to be untrue	$(F\text{♪}\psi \rightarrow B_L\neg\psi)$
6	Truth-telling as saying something that is both true in reality and sincere	$(T\text{♪}\psi \rightarrow (\psi \wedge B_T\psi))$
7	Lying as saying something that is both untrue in reality and insincere	$(F\text{♪}\psi \rightarrow (\psi \wedge \neg B_L\psi))$
8	Lying as saying something that is both untrue in reality and believed to be untrue	$(F\text{♪}\psi \rightarrow (\psi \vee B_L\neg\psi))$
9	Truth-telling as imparting knowledge	$(T\text{♪}\psi \rightarrow K_T\psi)$
10	Lying as imparting what one knows to be false	$(F\text{♪}\psi \rightarrow K_L\neg\psi)$

1.1 Identifying Liars and Truth-tellers through an Update Game

A simple method of discernment between liars and truth-tellers is through the scope of an update game. For each of these games the agents must adhere to four key conditions:

- An agent must make a statement to begin the game

- Statements must be made from one agent to at least one other
- The second agent x is taken as a reliable source by the first agent a: $(x♪_a\psi \to (\psi \wedge K_a\psi))$
- An agent must make a statement if possible whenever asked (statements can change, but means of truth-telling cannot change)

1.1.1 Interaction with Alan Truering and Ada Love-Lies

Let an omniscient truth-teller be an agent named Alan Truering, and an omniscient liar be named Ada Love-Lies
If you know an agent is Ada Love-Lies or Alan Truering, what do you require to know the true state of ψ? In this game all you require is the update from the agent, and knowledge of their type, to know the true state of ψ
If Alan Truering states ψ, then you know ψ, due to:

$$(T_a♪\psi \to K_a\psi)$$
$$(K_xK_a\psi \to \psi)$$

The proof of this is as follows:

1.	$(KK_a\psi \to \psi)$	Hypothesis
2.	$(K_a\psi \to \psi)$	Facticity
3.	$(K(K_a\psi \to \psi))$	Necessitation
4.	$(KK_a\psi \to K_a\psi)$	Translates across implication
5.	$(K\psi \to \psi)$	Facticity
6.	$(((p \to q) \wedge (q \to r)) \to (p \to r))$	Tautology
7.	$((KK_a\psi \to K\psi) \wedge (K\psi \to \psi)) \to (KK_a\psi \to \psi))$	Substitution
8.	$((KK_a\psi \to K\psi) \wedge (K\psi \to \psi)$	4,5, Conjunction
9.	$(KK_a\psi \to \psi)$	7,8, Modus Ponens

The same is true for Ada Love-Lies by simply negating her statement.

1.1.2 Identification of Alan Truering and Ada Love-Lies

What if you are presented with an agent a, who is either Alan Truering or Ada Love-Lies, but you do not know which?
This question is answered by the simple game that presupposes your knowledge of the true state of ψ during the agent's announcement.
If ψ is true and the agent announced ψ then you know the agent is Alan Truering.

If $\neg\psi$ is true and the agent announces ψ then you know you are dealing with Ada Love-Lies. This entails the following model, with x as "you":

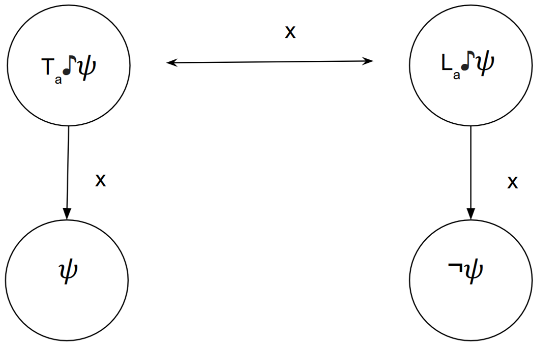


Figure 1.1: Initial space

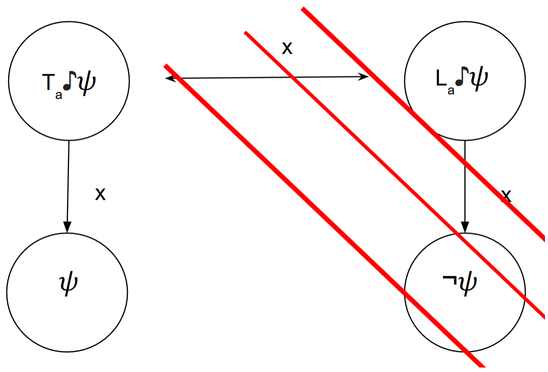


Figure 1.2: After updating with ψ

1.2 Sincere and Insincere Men

So far this game as appeared straightforward, as in each instance there only exists one possible relationship between the agent's statement and the true state of ψ in the world.

Now instead of assuming that you must know ψ in order to be a truth teller, consider instead if all that is required to be a truth teller is to believe ψ, as per types 3,4, and 5, with liars being defined equivalently.

Unlike omniscient liars and truth-tellers, you cannot identify if someone is sincere or insincere based purely off whether their statement is true in reality.

Omniscient Truth-teller:

$$((T_a \text{♪} \psi \rightarrow K_a\psi) \rightarrow \psi)$$

Sincere Men:

$$((T_a \text{♪} \psi \rightarrow B_a\psi) \nrightarrow \psi)$$

Simply put, it is possible for people to have incorrect beliefs.

1.2.1 Identification of Sincerity

If we now use the same framework for our game as we have had previously then we are able to identify if an agent is sincere or insincere;
Once the agent a makes their claim

$$a \text{♪} \psi$$

Then if you know ψ, all you need to do is update a to ψ.
If a knows that all updates are true updates $K_a <!\psi>$

$$K_a\psi$$
$$K_a\psi \rightarrow B_a\psi$$
$$B_a\psi$$

Knowing they now have a belief in ψ, if you ask them to repeat their statement they will have to respond in specific ways.
A sincere man will always state $a \text{♪} \psi$ as they now believe ψ
An insincere man will always state $a \text{♪} \neg\psi$ as they now believe ψ with $\tau = \infty$ and as such unless they have a threshold of 0, they will $\neg B_a \neg\psi$

1.3 A 4-Player Game

Finally, we look at a game in which the agent can exist in four possible states. The agent can be either:

- Alan Truering (omniscient truth-teller)
- Ada Love-Lies (omniscient liar)
- Sincere man
- Insincere man

If the agent announces ψ you are unable initially to distinguish which state they are in, as per the following model: However, if you come to know ψ to be true then you can eliminate Ada Love-Lies from the possible states of a due to her defining proposition:

$$(K_L\text{♪}\psi \to \neg\psi)$$

Based on the following derivation:

$$(K_L\text{♪}\psi \wedge \psi)$$
$$(K_L\text{♪}\psi)$$
$$\psi$$
$$(K_L\text{♪}\psi \to \neg\psi)$$
$$\neg\psi$$
$$\times$$

The same format can be used to eliminate Allan Truering if you know $\neg\psi$.
Now with the knowledge of ψ, you must discern between Alan Truering and the sincere and insincere men.
To do this, you must make an announcement. However, unlike the case in which you are restricted to sincere and insincere men you mus not announce ψ - if you do then ψ becomes public knowledge, thereby allowing the sincere man to hold knowledge of ψ, making them indistinguishable from Alan Truering.
Instead, announce a simple conjunction:

$$(\psi \to K\psi)$$

Chapter 2

Gettier Types

	Description	Logic
11	Gettier Truth-tellers announce truth, for the wrong reasons	$(T\text{♪}\psi \rightarrow ((\neg\chi \wedge \psi) \wedge B_\tau(\chi \rightarrow \psi) \wedge B_\tau\chi))$
12	Gettier Liars announce falsity, for the wrong reasons	$(F\text{♪}\neg\psi \rightarrow ((\neg\chi \wedge \psi) \wedge B_\tau(\chi \rightarrow \psi) \wedge B_\tau\chi))$

Based on Van Bentham's notion of a Gettier Truth-teller,[1] consider a Gettier Truth-teller defined as such; Gettier Truth-tellers announce truth, sincerely, but based on false evidence. They announce ψ, which is true in reality, based on some evidence χ, which is false in reality. They make this based on a modus ponens implication of $(\chi \rightarrow \psi)$, the only falsity being their belief in χ, which is false.
Gettier Liars exist in a similar manner, the primary difference being that they seek to deceive, negating the truth, which they succeed in, but based on evidence that is false.

2.1 Gettier Interactions

From their announcement of ψ as well as your knowledge that they believe χ falsely, then due to the nature of implication, $(\chi \rightarrow \psi)$ holds regardless of the truth value of χ, due to ψ being true. Therefore you can't immediately conclude that they are a Gettier Truth-teller - they may be entirely generic.
In order to ascertain them as Gettier, truthfully announce that their (false, but not necessarily known to you) antecedent χ leads to the negation of a different (true,

1. Johan van Benthem, *Modal logic for open minds* [in eng] (Stanford, Calif.: CSLI Publications, Center for the Study of Language / Information, Leland Stanford Junior University, 2010), ISBN: 9781575865980.

but not known to them as true) consequent ρ, giving the common knowledge $(\chi \rightarrow \neg\rho)$, and they should then be able to make the announcement of their belief in the negation of the (true) consequent ρ, i.e. $\neg\rho$, which you know to be a falsehood. This can be modelled as:

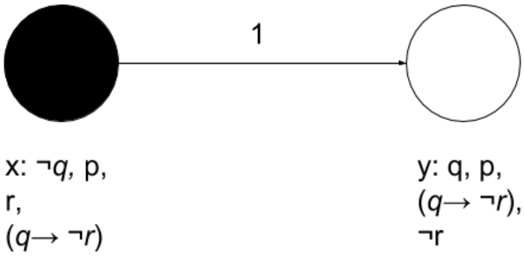


Figure 2.1: Plausibility model

Gettier Liars Interaction is symmetrical to Gettier Truth-tellers - this is not a unique property of Gettier announcers, however it is worth noting that symmetry isn't common to all types, such as 3,4,5

2.1.1 State Plausibility Models

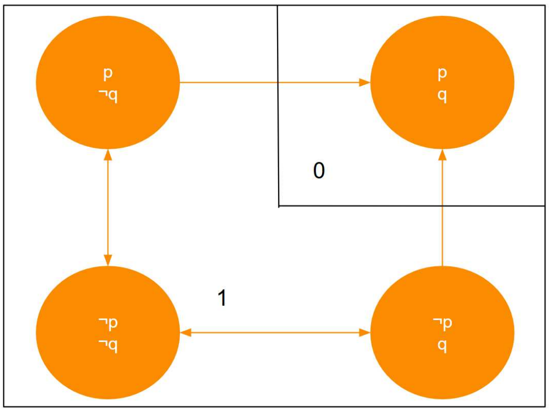


Figure 2.2: Before $(\chi \rightarrow \neg\rho)$ announcement, with the derived Spohn rank of states

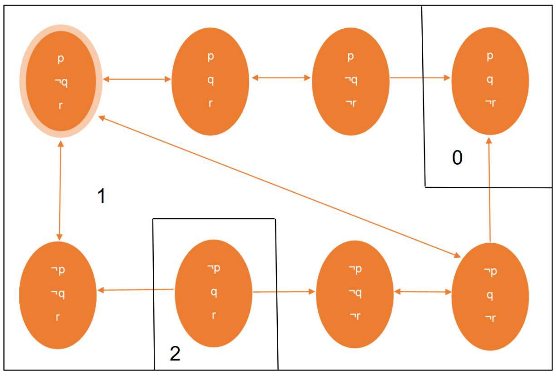


Figure 2.3: After $(\chi \rightarrow \neg\rho)$ announcement, with the derived Spohn rank of states

Chapter 3

Non-Gettier Types

What if you aren't a Gettier Truth-teller? Negating the necessary conditions of Gettier Truth-tellers gives rise to 3 other possible cases when the state of the actual world remains the same. These cases give insight on what other kinds of truth tellers or liars may exist and what might differentiate them.

3.1 The Beliefless Truth-teller

$$((\neg\chi \wedge \psi) \wedge \neg B(\chi \rightarrow \psi) \wedge \neg B\chi)$$

In this case, the speaker has no belief in any of the Gettier conditions. The speaker may very well have other types of belief or knowledge that contribute to the announcement of ψ, but these are further sub-cases to be examined. If there is no other belief or knowledge that pertains to ψ, a speaker satisfying these conditions announces ψ with no rationale or justification. The announcement happens to be true in this world, but that is pure happenstance (when discarding other possible beliefs or knowledge).

This case does allow for sub-cases in which the speaker does in fact have a justified true belief in ψ however, as it allows for sub-cases such as:

$$((\neg\chi \wedge \psi) \wedge \neg B(\chi \rightarrow \psi) \wedge B\neg\chi)$$

3.1.1 State Plausibility Models

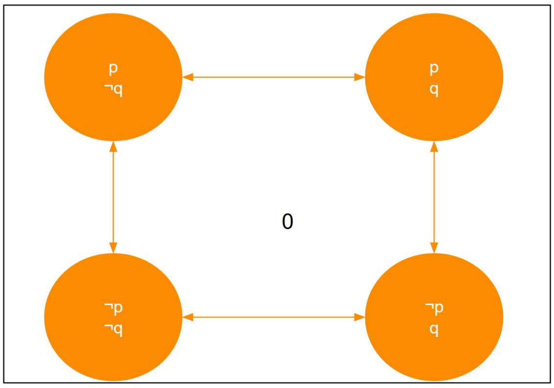


Figure 3.1: Before $(\chi \to \neg\rho)$ announcement, with the derived Spohn rank of states

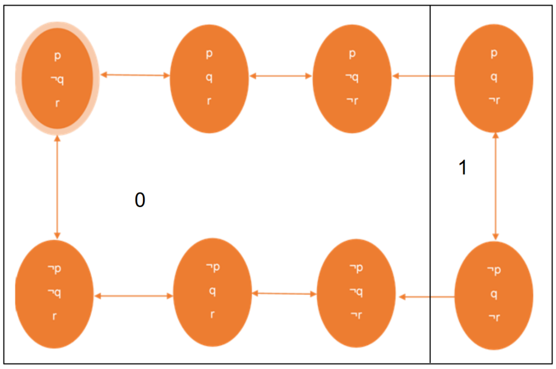


Figure 3.2: After $(\chi \to \neg\rho)$ announcement, with the derived Spohn rank of states

3.2 The Uninformed Accidental Truth-teller

$$((\neg\chi \land \psi) \land \neg B(\chi \to \psi) \land B\chi)$$

In this case, one has an untrue belief about the world ($B\chi$ when $\neg\chi$), but also doesn't have a belief in the implication $(\chi \to \psi)$ that Gettier Truth Tellers have (a belief which holds true in reality in this state of the world). If one announces ψ with these conditions, they are announcing a true thing but have very little rationale in doing so without considering additional beliefs. They are indifferent on ψ if no other beliefs are held, and therefore announcing ψ seems to be announcing something insincerely. One possible sub-case:

$$\begin{aligned}
&((\neg\chi \land \psi) \land \neg B(\chi \to \psi) \land B\chi \land B\neg(\chi \to \psi))\\
\leftrightarrow&((\neg\chi \land \psi) \land \neg B(\chi \to \psi) \land B\chi \land B(\chi \land \neg\psi))\\
\leftrightarrow&((\neg\chi \land \psi) \land \neg B(\chi \to \psi) \land B\chi \land B\neg\psi)
\end{aligned}$$

This sub-case is one in which the speaker announcing ψ announces something true about the world, but does so by saying something she believes to be untrue - implying an intent to deceive and a type of lie, but announcing a true statement due to her incorrect beliefs.

3.2.1 State Plausibility Models

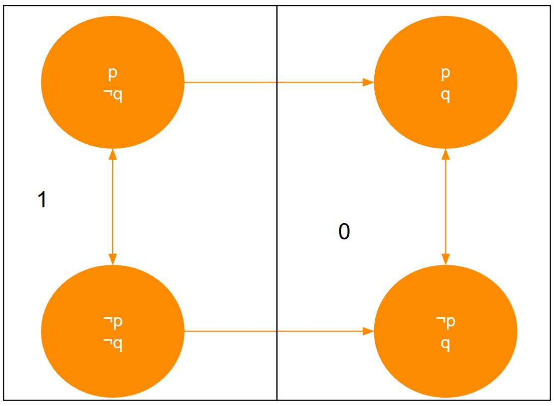


Figure 3.3: Before $(\chi \to \neg\rho)$ announcement, with the derived Spohn rank of states

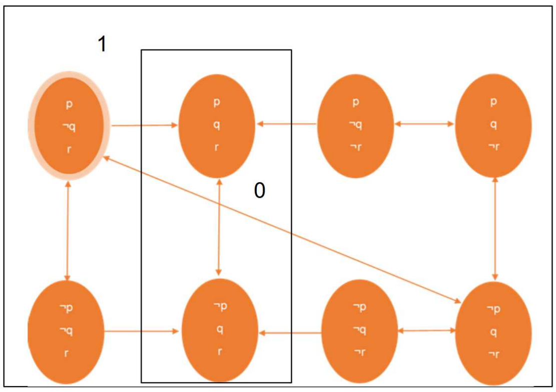


Figure 3.4: After $(\chi \to \neg\rho)$ announcement, with the derived Spohn rank of states

3.3 The Accurate Believer

$$(\neg\chi \wedge \psi) \wedge B(\chi \rightarrow \psi) \wedge \neg B\chi)$$

In this case, announcing ψ is again saying something true, but based on very little evidence without considering other beliefs the speaker may hold. They dont believe χ, but they do believe $(\chi \rightarrow \psi)$, and announcing ψ may be with the intent to deceive if they don't believe ψ. This case is one in which the speaker has the most accurate set of beliefs. It is the only case in which the speaker both believes the true statement $(\chi \rightarrow \psi)$ and does not falsely believe χ. Interestingly, this case is the only one in which Spohn Rank 0 contains only 5 cases rather than 6. This case allows for sub-cases in which the speaker has a justified true belief of ψ or a justified false belief in $\neg\psi$.

$$(\neg\chi \wedge \psi) \wedge B(\chi \rightarrow \psi) \wedge \neg B\chi \wedge B(\psi \rightarrow \chi))$$

In this case, one cannot be sincere in announcing ψ as it would then result in $B\chi$ but it is given that $\neg B\chi$

3.3.1 State Plausibility Models

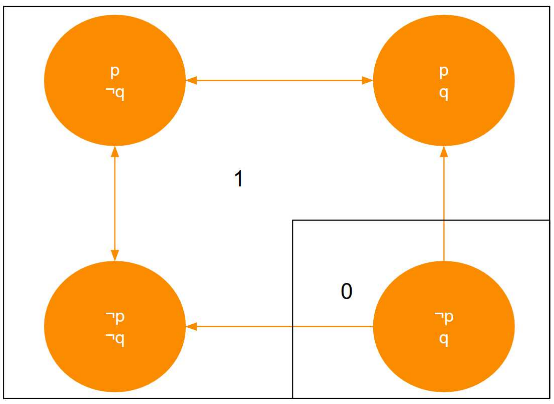


Figure 3.5: Before $(\chi \rightarrow \neg\rho)$ announcement, with the derived Spohn rank of states

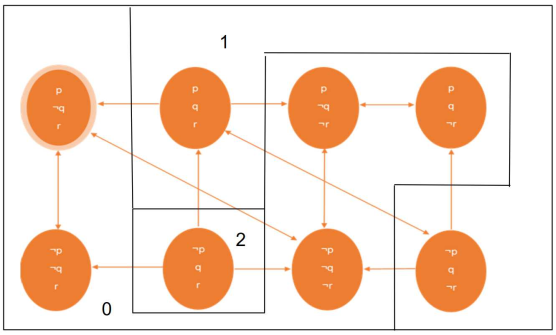


Figure 3.6: After $(\chi \to \neg\rho)$ announcement, with the derived Spohn rank of states

Chapter 4

Further Considerations

Meta:

- Is there an exhaustive list of liars and truth-tellers?
- Is it even possible to arrive at one at all?

Basic Types:

- Interactions between omniscient and insincere men
- How different thresholds affect what can be claimed by sincere and insincere men ($\tau \leq \frac{1}{2}$ and $\tau > \frac{1}{2}$)

Building on Gettier and non-Gettier Types:

- What sub-cases arise from the non-Gettier cases?
- What facets of interaction with these different sub-case agents will differ?
- Is this non-Gettier approach a useful approach or is too much redundancy embedded in it?

Bibliography

Benthem, Johan van. *Modal logic for open minds* [in eng]. Stanford, Calif.: CSLI Publications, Center for the Study of Language / Information, Leland Stanford Junior University, 2010. ISBN: 9781575865980.

Hendricks, Vincent, and John Symons. "Epistemic Logic." In *The Stanford Encyclopedia of Philosophy*, Fall 2015, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University, 2015.

Segerberg, Krister, John-Jules Meyer, and Marcus Kracht. "The Logic of Action." In *The Stanford Encyclopedia of Philosophy*, Winter 2016, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University, 2016.