One hundred prisoners and n number of light bulbs

Jasmine O'Callaghan, Joost Franssen, Pedro Rodriguez

June 27, 2022

1 TO DO Simulation:

- Create solver that solves all three protocols
- Gives days how long it takes for each protocol to finish, and if it is possible
- random generator sequence for prisoners (with length n? just a variable)
- do we want three separate solver files and then a master file that runs them? Yeah i guess so
- add names in dictionary
- output order and who is in groups/counter
- maybe be clearer about the steps
- create protocol 4 solver

Introduction

Over the years, logical research on riddles has proven to be a fascinating topic. Theories and concepts can be put to the test by applying them to real-world problems. The riddle of 100 prisoners and one light bulb has been solved numerous times and is well known to the general public. The riddle was even extended with several additional puzzles by [1]. For this research we will modify the 100 prisoners and one light bulb puzzle into a riddle with more light bulbs and see how we can extend the existing protocols and create protocols for new solving strategies.

The riddle

A group of 100 prisoners, all together in the prison dining area, are told that they will be all put in isolation cells and then will be interrogated one by one in a room containing a light with an on/off switch. The prisoners may communicate with one another by toggling the light switch (and that is the only way in which they can communicate). The light is initially switched off. There is no fixed order of interrogation, or interval between interrogations, and the same prisoner may be interrogated again at any stage. When interrogated, a prisoner can either do nothing, toggle the light switch, or announce that all the prisoners have been interrogated. If that announcement is true, the prisoners will (all) be set free, but if it is false, they will be in jail forever. While still in the dining room, before the prisoners go to their isolation cells (forever), can the prisoners agree on a protocol that will set them free? [1].

We will change the number of light bulbs into n amount for our variation. Therefore, there will be n amount of light bulbs available to switch on or off in our interrogation room. We will first explain why a protocol is needed to solve this problem. After this we will explain the steps [1] have taken to solve the problem for one light bulb after which we will increase the explanation for more light bulbs. We will need to build a Kripke model and examine the many states and actions to solve this problem. Another important assumption we adopt from [1] is that this question is not a trick question. Therefore, we only have logical solutions.

Initial riddle analysis

We need a protocol to solve the light bulb problem because the individual prisoners cannot identify when all prisoners have been interrogated. We can explain why we need protocols by modelling the prisoners' knowledge and relations. Such that we can see what every prisoner thinks are the possible states after several days. We use the Kripke semantics. First, we must state the assumptions we are making in the riddle. These are:

- All prisoners wish to escape

- 1 prisoner enters the room each day

- The lightbulb's initial state is either off

- O = light bulb is on

- E = everybody has been interrogated

We consider the Kripke structure $M = \langle S, \pi, R_A, R_B, R_C \rangle$ with $S = \{(\neg O, \neg E), (O, \neg E), (\neg O, E), (O, E)$ and π, R_A, R_B, R_C as indicated in Figure 1 - 1 In this Kripke model we show what states can be possible after the number of days increase. We show in the Kripke model that it is not possible to deduce for one single prisoner to know when everybody has been interrogated. It is common knowledge to the prisoners that one prisoner enters the interrogation room per day, and when in the interrogation room the prisoner can either turn on or off the light bulb. If we take 3 prisoners: Alex (A), Bennie (B) and Charlie (C). Here is the progression of the Kripke model:
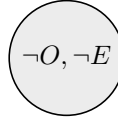

Figure 1: Initial state of the model

The initial state of the Kripke model in Figure 1 just contains the state in which the light bulb is turned off, and nobody considers it possible that everybody has been interrogated. In Figure 1 we can see that from the initial state in one step every agent considers it possible that either we are still in the same state or we have moved to the state in which the light bulb is now turned on. But because the number of days is lower than the amount of prisoners we can not consider it true that everyboyd has been interrogated.
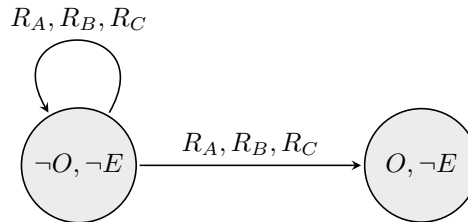

Figure 2: After the first day

In Figure 1 we add two more relations because we can either go back to the state in which the light bulb is off or stay in the state in which the light bulb is on. In Figure 1 we see exactly the same possible states as the day before.

However, at day 4 we can add several new states. Because we up until day 4 we do not need to consider the possibility that all the prisoners have been in the room, since the number of prisoners
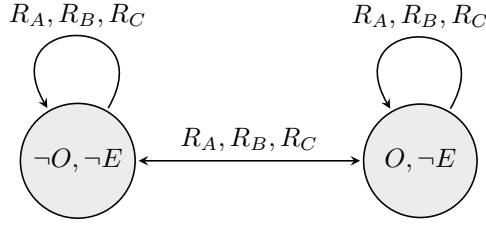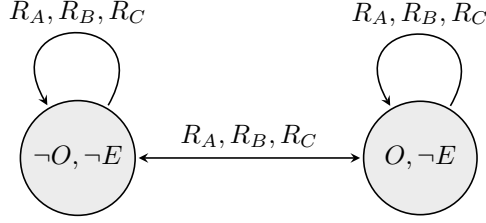
Figure 3: After the second day

Figure 4: After the third day

was greater than the number of days. However, from day 4 on wards it is possible that everybody has been interrogated. Therefore, we add two more states as all of the prisoners may have entered the room, and at this point both the light bulb can be on or off. In Figure 1 we can see the two states added.
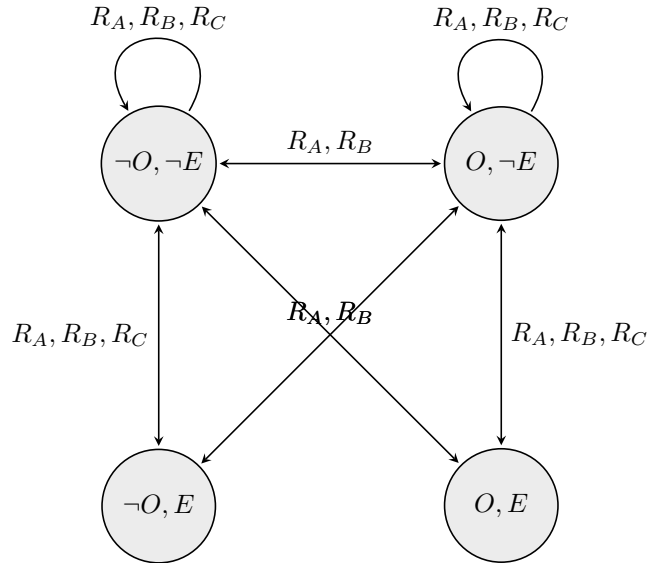

Figure 5: After the fourth day

The last Kripke model is shown in Figure 1 where the relation between the two previously added states. As well as, the reflexive relations are added. Because, we can now either be in those states or can stay in the same state.

The Kripke model does not change anymore from this point on wards. The knowledge an agent has depends on at which state he enters the room either with the light bulb on or off. However,
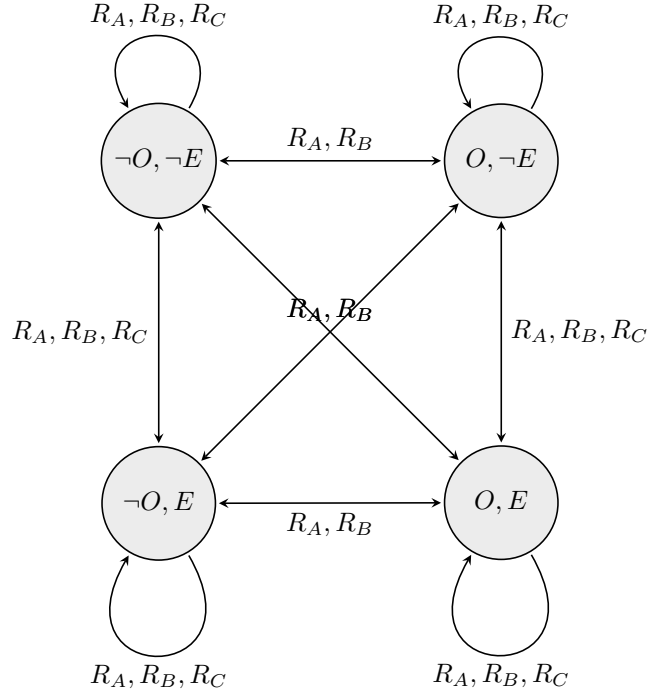
Figure 6: After the fifth day

from that state every other possible state in the Kripke model is reachable. Therefore, the agents never know when everybody has been interrogated, the agents can only guess who has been in the room, and know that if they do not have been in the room for a day someone else has been in the room. But they do not know for certain that the guards have not just been letting two prisoners into the interrogation room or that all prisoners have been in the interrogation room. Therefore, the prisoners need a protocol such that the prisoners know for certain that everybody has been interrogated.

Protocols

In [1] they provide several protocols, and solutions to the one hundred prisoners and one light bulb riddle. During the book they use natural deduction to break down the problem, and first start with a prove for a small number of prisoners before adding more and more prisoners. Therefore, if the problem can be solved for n people, it can also be solved for $n + 1$. In this review of the previous riddle, and the newly created riddle we will also use this strategy of natural deduction. During this section we will use the [1] to give the explanation for the riddle with one light bulb, on which we will describe how the problem works with more light bulbs, and how this changes the protocols that come with it. Furthermore, we will investigate newly created protocols to try to solve the riddle.

One light bulb

First, we will explain the riddle that only contains one light bulb, then increase the number of light bulbs to two or more light bulbs if possible, and then see how the problem works if we increase the number of light bulbs higher than the number of prisoners.

We can see in [1] that the 100 prisoners and one light bulb riddle is solved by choosing one person named the counter, after which every person who enters the interrogation room can turn on the light bulb just once. If the light is already switched on, they do nothing. Only when the light is switched

off, and the prisoner did not yet turn the light on, turn it on. If the only person who turns off the light is the counter, he can add one to the count every time he turns it off. Furthermore, [1] shows that the solution for the problem from two prisoners on wards stays the same. Therefore, we will use an arbitrary number of prisoners in our explanation/examples to solve the riddle that is lower than 100 but higher than two, so that explainability is guaranteed. The protocol for this solution is given in protocol 1.

Protocol 1 *The prisoners appoint one amongst them as the counter. All non-counting prisoners follow this protocol: the first time they enter the room when the light is off, they turn it on; on all other occasions, they do nothing. The counter follows a different protocol: if the light is off when he enters the interrogation room, he does nothing; if the light is on when he enters the interrogation room, he turns it off; when he turns off the light for the **(number of prisoners - 1)th** time, he announces that everybody has been interrogated.*

Consider Kripke structure $M = \langle S, \pi, R_A, R_B, R_C \rangle$ with $S = \{(\neg O, \neg E), (O, \neg E), (\neg O, E)$ and π, R_A, R_B, R_C as indicated in Figure 7. The structure contains three agents, namely: Alex (A), Bennie (B) and Charlie (C). There are two propositions used in this model which are $O = light\ bulb$ $is\ on$ and $E = the\ counter\ knows\ everybody\ has\ been\ in\ the\ interrogation\ room$ similar to before. The node containing a purple background is the starting node in which the light is off, and the counter does not know if everybody has been in the room.
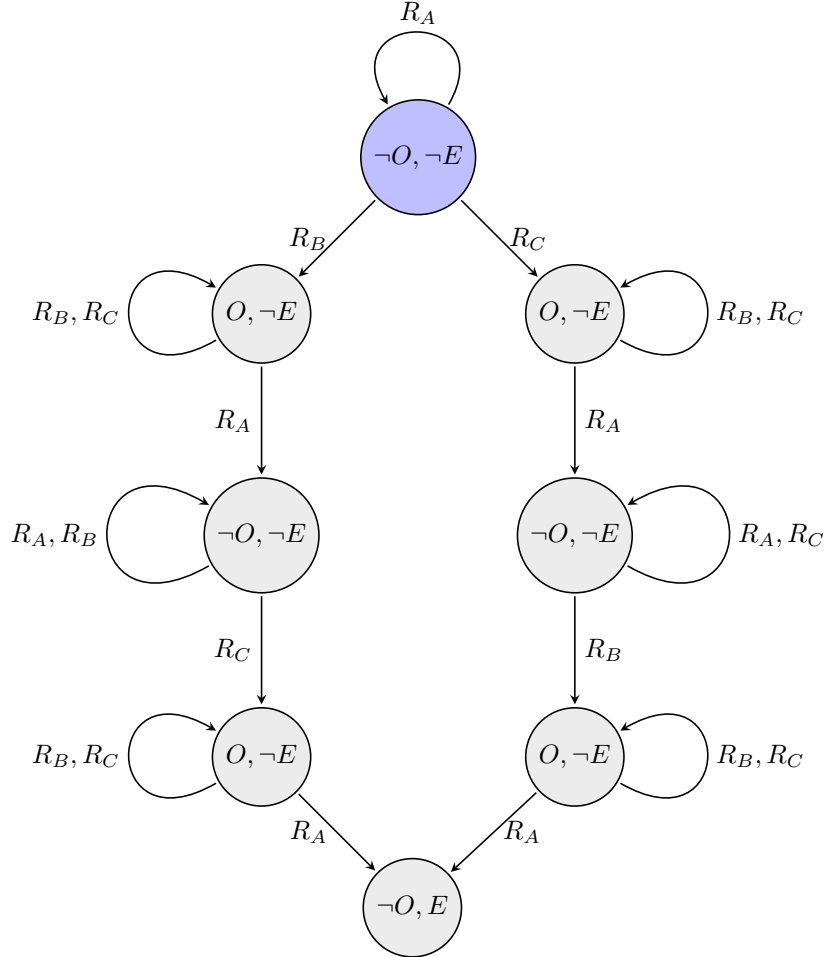

Figure 7: Kripke sequence model, one light bulb using protocol 1.

Step	Name	Action	Current state	Counter
0			$\neg O, \neg E$	0
1	Bennie	Bennie turns on the light bulb	$O, \neg E$	0
2	Alex	Alex turns off the light bulb	$\neg O, \neg E$	1
3	Charlie	Charlie turns on the light bulb	$O, \neg E$	1
4	Alex	Alex turns off light bulb 1, alex announces everybody has been interrogated	$\neg O, \neg E$	3

Table 1: Example sequence for the one light bulb problem using protocol 1. The threshold number for the count is the number of agents minus one. Because then the counter knows everybody has been interrogated. If the count reaches 3 (two times the light bulb is switched off then the counter can announce everybody has been interrogated)

Two or more light bulbs ($n > 1$ and $n \leq$ *number of prisoners*)

In this section, we will explain solving strategies if there are two or more light bulbs, but the number of light bulbs is lower than the number of prisoners. So, strictly $n \leq prisoners$.

Protocol 1

Firstly, we can use the same strategy as displayed in the one light bulb problem, and having one person that is the counter. However, now a person entering the interrogation room can either switch on the first light bulb or can switch on the second light bulb if the first one is already on. If both light bulbs are switched on already and the prisoners did not yet turn it on he has to wait until the lights are turned off by the counter prisoner. For example, if we have a example including three agents Alex, Bennie and Charlie of which Alex is the counter.

Consider Kripke structure $M = \langle S, \pi, R_A, R_B, R_C \rangle$ with $S = \{(\neg O_1, \neg O_2, \neg E), (O_1, \neg O_2, \neg E), (O_1, O_2, \neg E), (O_1, O_2, \neg E), (\neg O_1, \neg O_2, E)$ and π, R_A, R_B, R_C as indicated in Figure 8. We can observe that the structure is a slight variation of Figure 7. One addition assumption to protocol 1 has to be made for explainability which is:

- If $\neg O_1, \neg O_2$ (both light bulbs are off when the agents enters the room) the agent turns on light bulb O_1

We can now visualize the sequence Kripke model in Figure 8. In which we can observe that every relation has to be used for the model to get into the last state (goal state E). Therefore, we know that in every way possible before the counter can announce that 'everybody has been in the interrogation room' everybody must have been in the interrogation room. We can see a example solving strategy in Figure 1. The solving strategy now takes as many days as before with one light bulb. However, in the Figure we can see that there is a faster way of solving the problem. Therefore, even though in this example it is not faster on average having two light bulbs will increase the speed of solving.

Binary representation

Secondly, the most simple and probably the fastest way to solve this problem is to use the states of the light bulbs to displays numerical numbers via binary representation. With binary representation is meant that a certain number for instance 5 which is in binary representation 101 needs three light bulbs to be displayed. Therefore, this solving strategy only works if the number of prisoners can be binary represented by the number of light bulbs. So, if we have two light bulbs, for instance, we can only use binary representation as a solving strategy if the number of prisoners is lower than 4. We
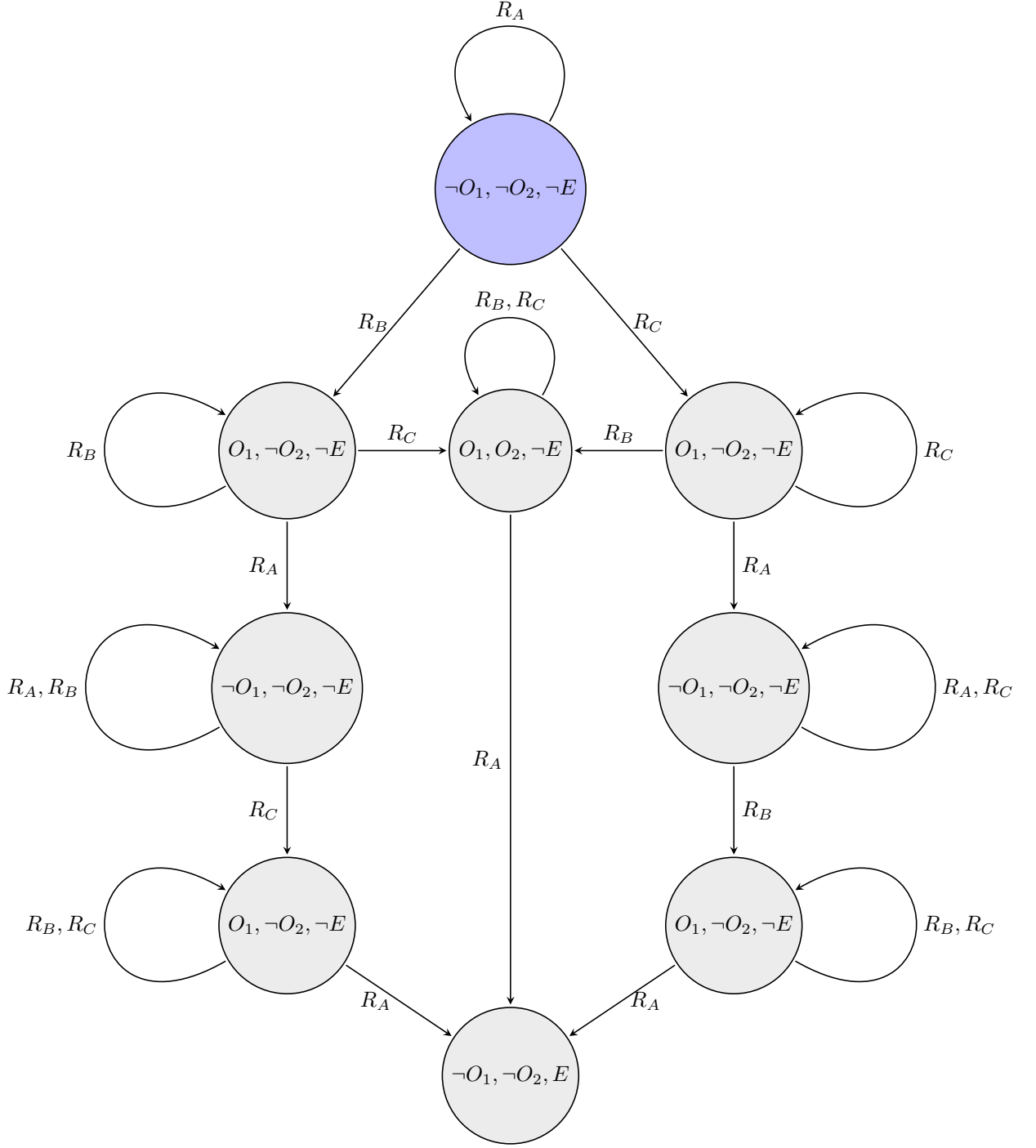
Figure 8: Kripke sequence model two light bulbs protocol 1

can depict the number three using the two light bulbs, but we can suggest that if a person enters the room that has not been interrogated before, and he sees the number three depicted by the two light bulbs he can say with confidence that the other people have been interrogated. Therefore, we can count to four with two light bulbs using the binary representation. In binary representation every light bulb that is added to the number multiplies the number that can be numerical represented by

Step	Name	Action	Current state	Counter
0			$\neg O_1, \neg O_2, \neg E$	0
1	Bennie	Bennie turns on light bulb 1	$O_1, \neg O_2, \neg E$	0
2	Alex	Alex turns off light bulb 1	$\neg O_1, \neg O_2, \neg E$	1
3	Charlie	Charlie turns on light bulb 1	$O_1, \neg O_2, \neg E$	1
4	Alex	Alex turns off light bulb 1, alex announces everybody has been interrogated	$\neg O_1, \neg O_2, \neg E$	3

Table 2: Example sequence for the two light bulb problem using protocol 1. The threshold number for the count is the number of agents minus one. Because then the counter knows everybody has been interrogated. If the count reaches 3 (two times the light bulb is switched off then the counter can announce everybody has been interrogated)

two. The number of prisoners counted using this strategy is the binary representation possible with n number of light bulbs plus one. For this protocol to work we need several assumptions about the knowledge the prisoners have:

- It is common knowledge what the order of the light bulbs is (we need to know which light bulb is which)

- It is common knowledge how binary representation works

In Table 3 we can see the answer for the an example sequence. First, Daniel puts on light bulb 2 to form the numerical representation 1. After which Alex turns on light bulb 1 and turns off light bulb 2 to create the numerical representation 2. Charlie then turns on light bulb 2 to create the numerical representation 3. Bennie enters the room for the first time and sees both lights turned on, and knows three people have been interrogated before. Therefore Bennie can announce everybody has been interrogated.

		Light bulbs		
Step	Name	1	2	Numerical representation
0		0	0	0
1	Daniel	0	1	1
2	Alex	1	0	2
3	Charlie	1	1	3
4	Bennie	1	1	3

Table 3: Example sequence for the light bulb problem using protocol 3

Protocol 2 *Prerequisite for this protocol: If the number of prisoners is lower or equal to the number that can be used to depict binary representation. Every prisoner that enters the room adds one to the binary representation using the light bulbs. When a prisoner enters the room that does not have entered the room before, and sees that the binary representation is the number of people minus one he can say that everybody has been interrogated.*

Several counters method

Lastly, we could now try a strategy in which we create groups and appoint one person per group as counter. The strategy is related to the one counter solution. However, now, we have as many counters as that there are light bulbs, s.t. every counter hold the count for one light bulb. We also have to divide the prisoners into several groups depending on the number of light bulbs. The

members of the groups can only turn on the light bulb that is assigned to there group. After, the counter has counted all the people in his group. He waits until he can turn on the light of the other groups. Therefore, if the counter of the other group has counted the number of his group plus the number of other counters, he knows that the other groups have finished counting as well, and he can therefore say that all the prisoners have been interrogated. The complete, and concise version of the protocol can be seen in protocol 3.

Protocol 3 *The prisoners appoint two (or more depending on the number of light bulbs) amongst them as the counter. All non-counting prisoners follow this protocol: the first time they enter the room when the light of their group is off, they turn the light of their group on; on all other occasions, they do nothing. The counter follows a different protocol: if the light of his/her group is off when he enters the interrogation room, he does nothing; if the light of his/her group is on when he enters the interrogation room, he turns it off. If the counter of there group has reached the amount of people is his/her group he then waits until he can turn on the lights of the other groups such that he can signal that his group is done. When one of the counters reaches the amount of his group plus the two other counters, he announces that everybody has been interrogated.*

However, while this protocol sounds solid, and does not deviate much from protocol 1, it is not. There are some cases of sequences in which the protocol does not hold, and a counter example can be found. To show why the model does not hold we created a complete Kripke model that contains all the possibilities for a example with four agents and two light bulbs. We can use Kripke semantics to explain protocol 3. The same assumptions as the Kripke model for protocol 1 hold here, however, two rules are added, and we change the strategy rule to protocol 3. For clarification all the rules are stated again here:

- All prisoners wish to escape

- All prisoners are aware of the strategy set out in protocol 3

- All prisoners are aware of their role (as counter or not)

- All prisoners are aware of their group

- All prisoners are aware which light bulb they can and can not touch

- 1 prisoner enters the room each day

- The light bulb's initial state is off

If we take 4 prisoners: Alex, Bennie, Charlie and Daniel. Alex (A) and Bennie (B) are assigned to group 1, in which Alex is the counter. Charlie (C) and Daniel (D) are assigned to group 2, in which Charlie is the counter. Group 1 has control over light bulb 1, and group 2 has control over light bulb 2. It is common knowledge that everybody knows their groups and the counters of each group. We will use the same propositions as before. However, now we will add numbers to the propositions which represent the groups. For clarification again we show all possible states. There are four propositions used in the model, and four negations of these propositions namely:

- O_1 = light bulb 1 is on

- O_2 = light bulb 2 is on

- E_1 = everybody of group 1 has been in the interrogation room

- E_2 = everybody of group 2 has been in the interrogation room

- $\neg O_1$ = light bulb 1 is off

- $\neg O_1$ = light bulb 1 is off

- $\neg E_1$ = not everybody of group 1 has been in the interrogation room

- $\neg E_2$ = not everybody of group 2 has been in the interrogation room

Consider Kripke structure $M = \langle S, \pi, R_A, R_B, R_C, R_D \rangle$ with $S = \{(\neg O_1, \neg O_2, \neg E_1, \neg E_2),$
$(O_1, \neg O_2, \neg E_1, \neg E_2), (\neg O_1, O_2, \neg E_1, \neg E_2), (O_1, O_2, \neg E_1, \neg E_2), (\neg O_1, O_2, E_1, \neg E_2), (O_1, \neg O_2, E_1, \neg E_2),$
$(O_1, O_2, E_1, \neg E_2), (O_1, \neg O_2, E_1, \neg E_2), (\neg O_1, O_2, \neg E_1, E_2), (O_1, O_2, \neg E_1, E_2), (\neg O_1, O_2, E_1, E_2),$
$(\neg O_1, \neg O_2, E_1, E_2), (O_1, \neg O_2, E_1, E_2)$ and π, R_A, R_B, R_C as indicated in Figure 9. To make the model more readable we divided it into four square boxes in which all states within one square box contain the same values for the two E propositions. The values are displayed either in the top left, top right or bottom left corner. The starting point of the Kripke model is at the top most state $(\neg O_1, \neg O_2)$ in the square $(\neg E_1, \neg E_2)$ which has a purple colored background. Every state in the model has four relations for all four agents that can enter the room at every state. Except the last three states in the lowest square $(\neg E_1, \neg E_2)$ because at this square the counters can announce that everybody has been interrogated. The counter knows that everybody in there group has been interrogated when the light bulb of there group has been switched on once's. Then the counter increases the count to 2 (he knows that he himself also has been in the room). So, the total members of there group which is 2 has been in the room. As the protocol states the counter can now turn on the light bulb of group 2 when it is off. One of the counters can announce that everybody has been interrogated when both propositions E_1, E_2 hold in a state.

In Table 4 we then can see a counter example that is run through the Kripke model and does not satisfy the condition that all agents have been in the interrogation room. Furthermore we can see in the table that Alex announces everybody has been interrogated while Daniel has never been in the interrogation room. Therefore, the protocol does not hold for this sequence, and the prisoners would be doomed to a life in prison if they would have followed this protocol.

| | | | | Count group | |
Step	Name	Action	Current state	A	B
0			$(\neg O_1, \neg O_2, \neg E_1, \neg E_2)$	0	0
1	Bennie	Bennie turns on light bulb 1	$(O_1, \neg O_2, \neg E_1, \neg E_2)$	0	0
2	Alex	Alex turns off light bulb 1, turns on light bulb 2	$(\neg O_1, O_2, E_1, \neg E_2)$	2	0
3	Charlie	Charlie turns off light bulb 2, turns on light bulb 1	$(O_1, \neg O_2, E_1, \neg E_2)$	2	2
4	Alex	Alex turns off light bulb 1. Alex announces everybody has been interrogated.	$(\neg O_1, \neg O_2, E_1, E_2)$	3	2

Table 4: Example sequence for the light bulb problem using protocol 3. The threshold number for the count for each group is 1. Because then the counter knows everybody in there group has been interrogated. Therefore, the count will increase to 2 because the counter also adds himself to the tally. If the count reaches 3 (two times the light bulb is switched off then the counter can announce everybody has been interrogated)

Number of prisoners minus 1 or more light bulbs than prisoners

In this section we will show solving strategies if the number of light bulbs is one lower than the number of prisoners or higher $n \geq number of prisoners - 1$.

The solving strategy is straightforward now. Every prisoner turns on one of the light bulbs the first time he enters the interrogation room. Then if he again enters the interrogation room, he does nothing only if the number of light bulbs turned on is the same as the number of people minus himself he knows that everybody has been interrogated, and therefore announces that everybody has been interrogated.
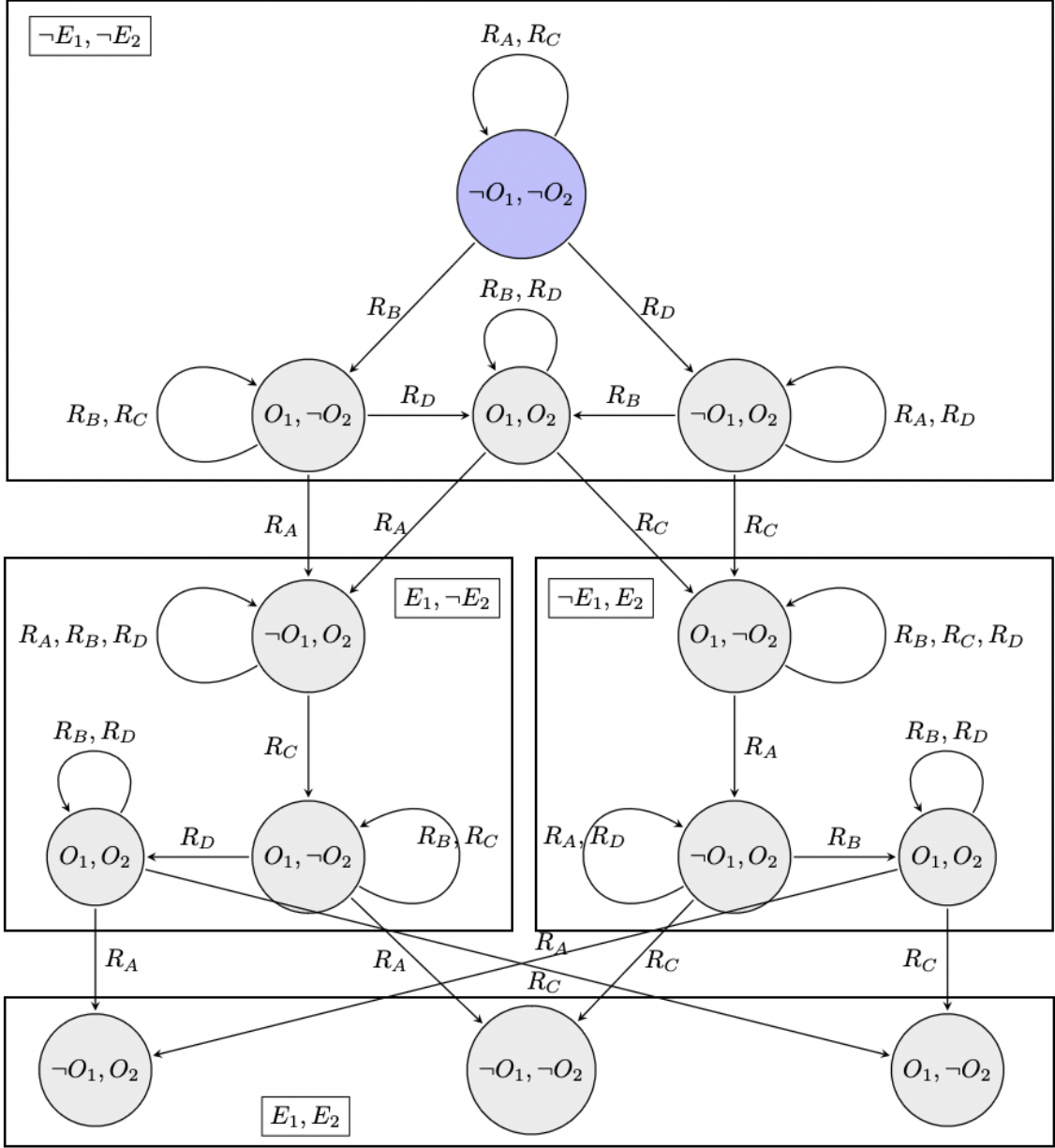
Figure 9: Kripke model of protocol 3. The Kripke model should be read top down. The model contains four squares that represent additional propositions that hold in the states that are in the square, this is done for ease and readability. The first state is the state that contains $(\neg O_1, \neg O_2)$ in the $(\neg E_1, \neg E_2)$ square. From there four different agents can enter the room that can trigger different relations and can reach new other possible states.

Protocol 4 *The first time you enter the interrogation room you switch one light bulb on. When you enter the interrogation room and you already switched on a light bulb you do nothing. If you enter the room and the number of light bulbs that are turned on is the same as the number of prisoner. You can announce that everybody has been interrogated*

We consider the Kripke structure $M = \langle S, \pi, R_A, R_B, R_C \rangle$ with $S = \{(\neg O, \neg E), (O, \neg E), (\neg O, E), (O, E)\}$ and π, R_A, R_B, R_C as indicated in Figure 10. We can use the same example as earlier that contained three prisoners Alex (A), Bennie (B) and Charlie (C) and two light bulbs. However, while those

protocols could be scalable to more prisoners with the same amount of light bulbs, this protocol can not be scaled. Therefore, before this protocol can be used the strict restriction of the number of prisoners minus - 1 should be equal or lower than the number of light bulbs must be met before protocol 4 can hold. It is still common knowledge to the prisoners that one prisoner enters the interrogation room per day, and when in the interrogation room the prisoner can either turn on or off the light bulb. Furthermore, still all the light bulbs are switched off as the initial state. Also, we assume that the prisoner that first enters the room turns on light bulb 1 after which prisoner two turns on light bulb 2 etc.
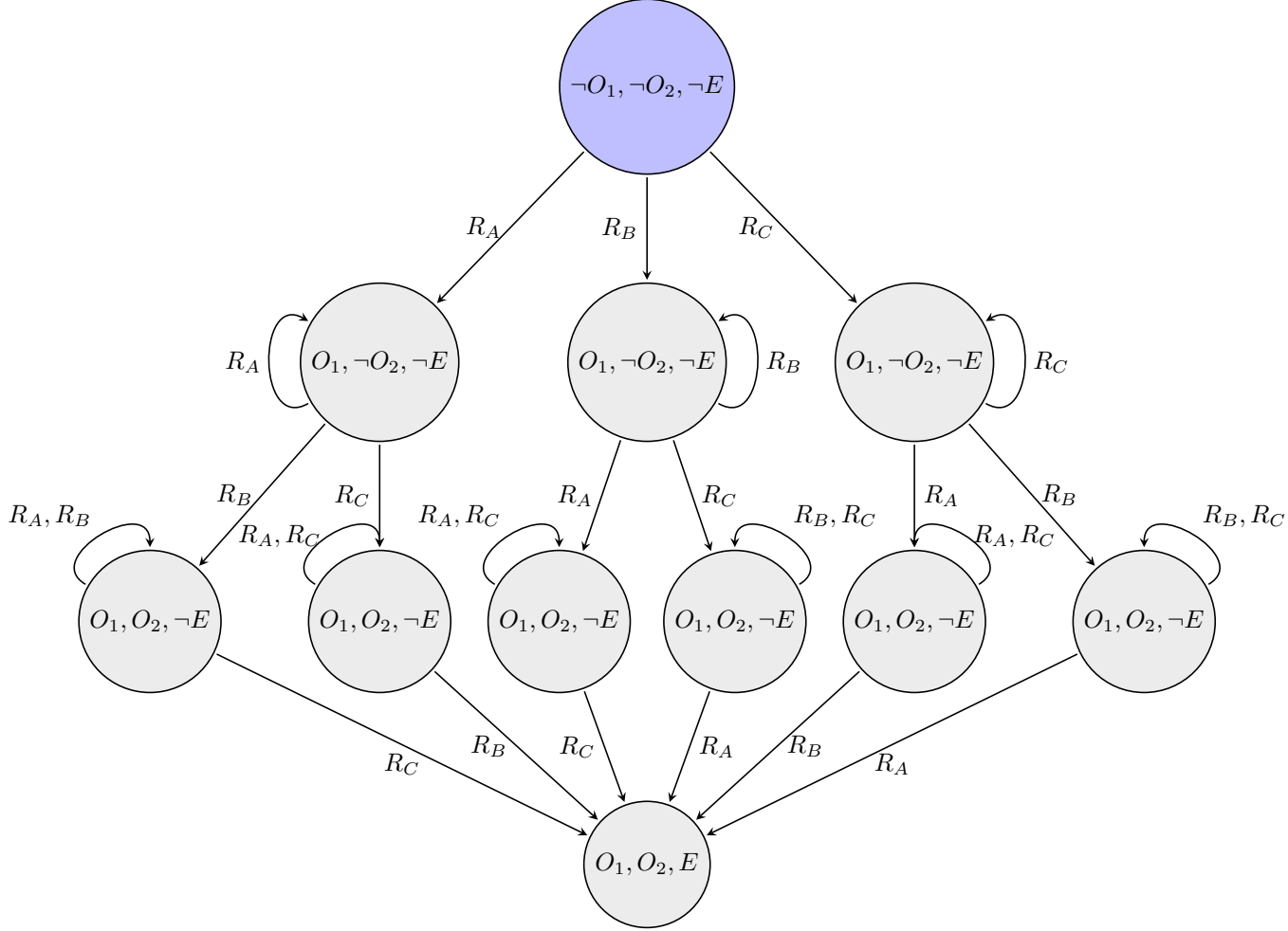

Figure 10: Kripke model for protocol 4

Conclusion

To conclude, we extended the 100 prisoners and 1 light bulb riddle into a version that included more than one light bulb, and showed several old and new solving strategies. We found that the old strategies that already worked on the one light bulb riddle could also be extended to the riddle that included more light bulbs. Furthermore, we made three new solving techniques which all have pros and cons. The binary representation and when the number of light bulbs is one less than the number or prisoners or more are good strategies given that all the assumptions both need do hold. In which the binary representation is a more effective strategy as you need fewer light bulbs to signal

Step	Name	Action	Current state
0			$\neg O_1, \neg O_2, \neg E$
1	Bennie	Bennie turns on light bulb 1	$O_1, \neg O_2, \neg E$
2	Alex	Alex turns on light bulb 2	$O_1, O_2, \neg E$
3	Charlie	announces everybody has been interrogated	O_1, O_2, E

Table 5: Example sequence for the two light bulb problem using protocol 4. If two people have entered the room before the third person, and they both turned on a light bulb the third person can announce everybody has been interrogated

the same thing as in protocol 4. The new solving strategy that included creating several groups that have several counters looked like a simple, and easy extension of the one counter method. However, while it looks feasible on paper the protocol does not hold in certain sequences of prisoners going into the interrogation room. Therefore, this is not a protocol you should use to solve this riddle.

Simulation

For the simulation of the riddle we want to create a interactive dashboard in which many setting can be chosen, and changed. The settings we want to create are: number of prisoners, number of light bulbs, solving protocol, time and sequence. In which the number of prisoners, light bulbs and solving protocol are obvious settings the sequence and time need some extra explanation. We want to let the user type in there own sequence or a random sequence which is generated by a sequence generator. After which the solver can check if the sequence is solvable for the chosen other settings. Then if the sequence is not solvable we also want to create a setting that the sequence can be added with the names needed to solve the sequence. We also want to create a setting for the sequence that the shortest answer is generated. The time setting can be adjusted in which the time is the time between interrogations. Then if the sequence is solved or not the time the prisoners are locked up is also displayed.

References

[1] Hans van Ditmarsch and Barteld Pieter Kooi. *One hundred prisoners and a light bulb*. Springer, 2015.