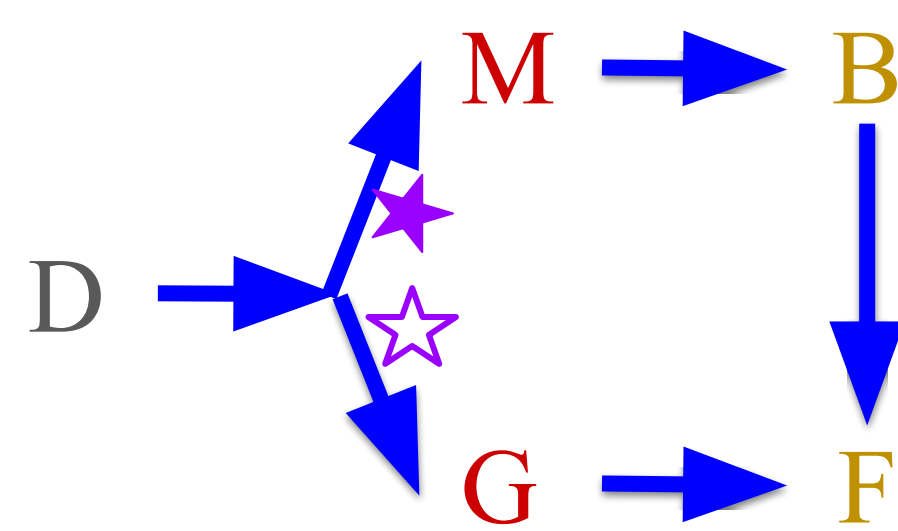


Plans that remain private, even in hindsight

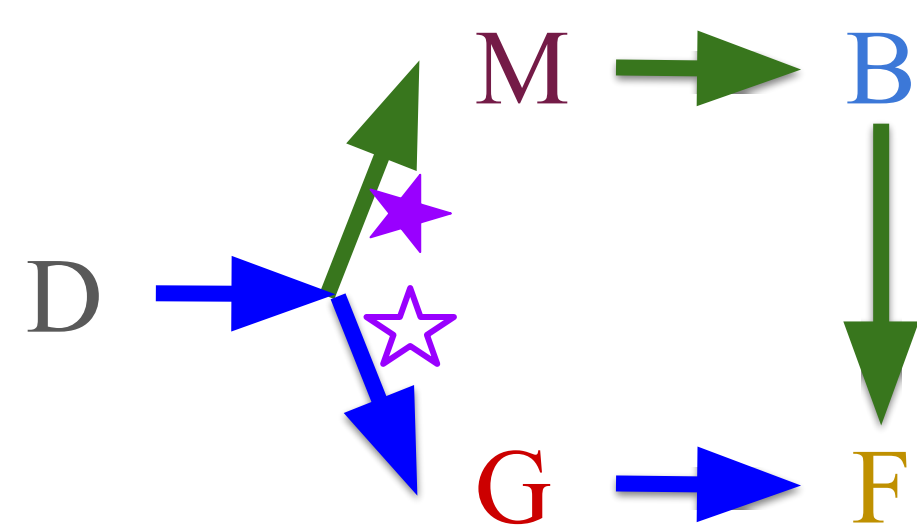
Motivation: privacy-aware scenarios Formalism



Looking forward:
(Filtering)

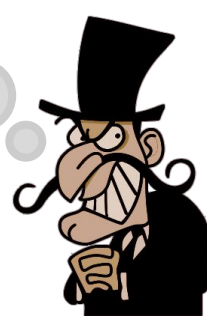


Looking forward and backward:
(Smoothing)



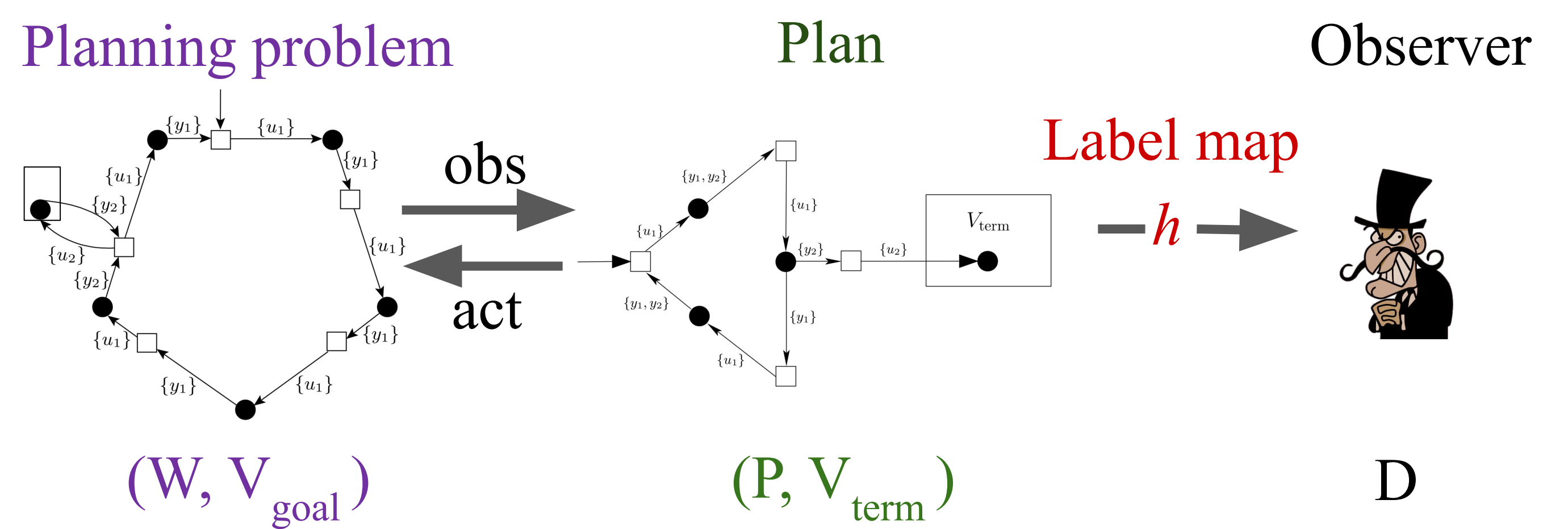
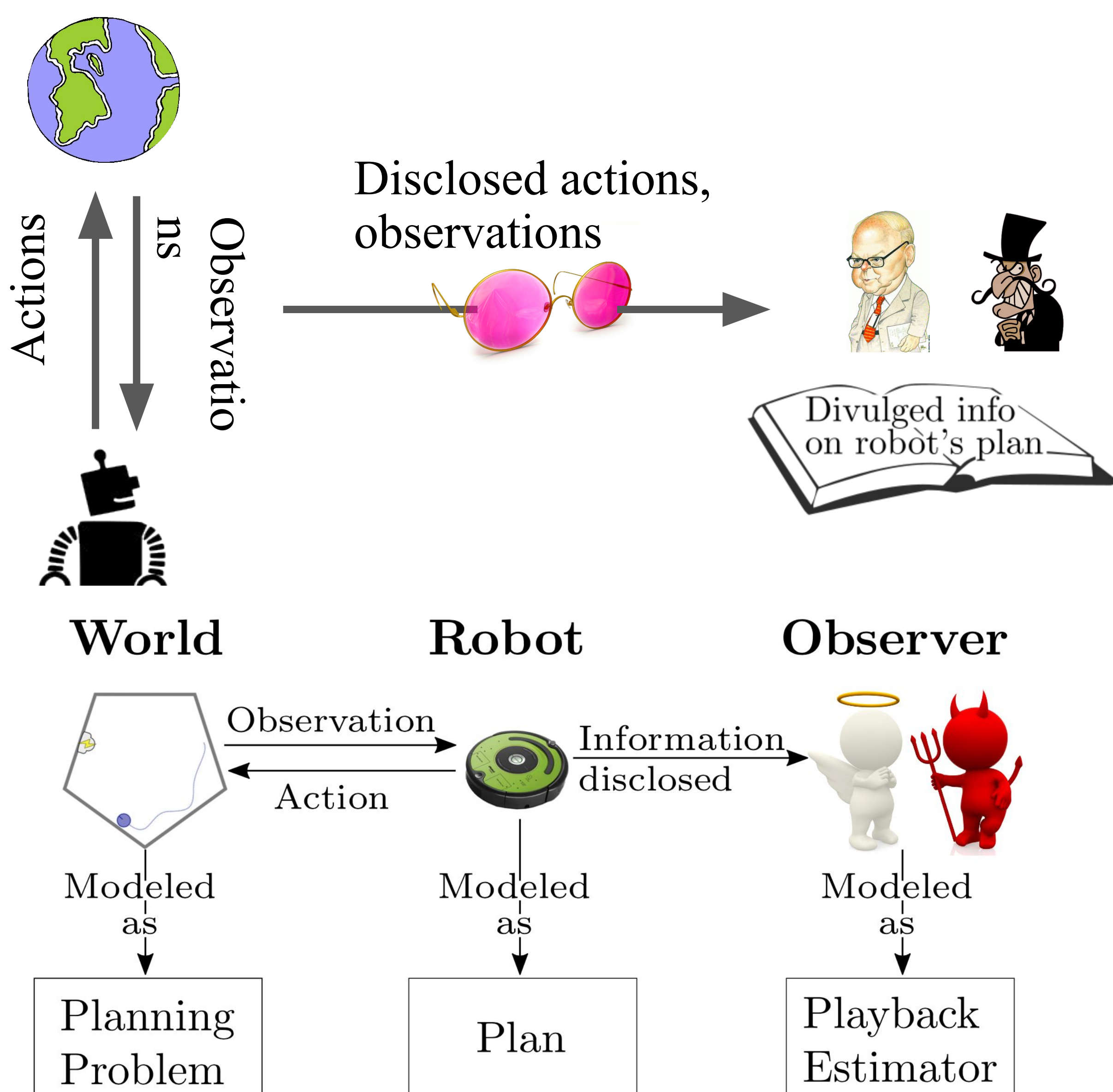
By looking forward, I know that the robot could get through both bedrooms. [1]

By looking backward, I know that there is a guest if it reaches the back yard after coming out from the bedroom.



We want to search for a plan such that its disclosed information to an observer with hindsight always satisfy the information stipulation.

Model: worlds, robots and observers



➤ When observer receives a string x , what does the observer know?

$$S = h^{-1}[x] \cap [\text{Strings in } D] \cap [\text{Strings in } W]$$

➤ Then trace S in W , and get the set of estimated world states W_B

➤ Stipulation represented by the propositional formula:

$$\phi = \text{master bedroom} \Rightarrow \text{guest bedroom}$$

➤ ϕ will be evaluated on every W_B .

Divulged info about the robot's plan

- The observer knows nothing about the robot's executions other than it is operating in the world.
- The observer only knows that the robot is executing a plan that will eventually reaches the goal states in the world within a finite number of time steps.
- The plan to be executed is hidden in a finite number of potential plans.
- The observer knows the exact plan to be executed.

It turns out that we can use a procrustean graph to represent each of these four cases.

Conclusions and future work:

- We can search for a plan that satisfies information stipulation on an observer with hindsight.
- How to search for a plan and information disclosure policy jointly so that information stipulation is satisfied?
- The filtering observer has a finite structure [1]. It seems that the smoothing observer does not. What is the right representation for a smoothing estimator?

[1] Yulin Zhang, Dylan A. Shell, Jason M. O'Kane, *Finding plans subject to stipulations on what information they divulge*, WAFR, 2018