Truthful Elicitation of Imprecise Forecasts

Anurag Singh[1], Siu Lun Chau[2] and Krikamol Muandet[1]

[1]Rational Intelligence Lab, CISPA Helmholtz Center for Information Security, Saarbrucken Germany

[2]Nanyang Technological University, Singapore

CISPA HELMHOLTZ CENTER FOR INFORMATION SECURITY

RATIONAL INTELLIGENCE

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Rational Intelligence Lab in CISPA

Krikamol Muandet
(PI)

Siu Lun Chau
Moved to NTU and recruiting!

Come visit us in Saarbrücken!

Rational Intelligence Lab

CISPA Helhmholtz Center for Information Security, Saarbrucken

We currently have an open position for a postdoctoral researcher (our call)

What is an imprecise forecaster?

50% 50%

$p \in \Delta(\mathcal{O})$

Precise Belief

Precise Forecaster

$p = 0.5$

$20 - 80\%$ $80 - 20\%$

$\mathcal{P} \subseteq \Delta(\mathcal{O})$

Imprecise Belief

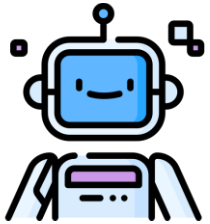
Imprecise Forecaster

$\mathcal{P} = [0.2, 0.8]$

What is an imprecise forecaster?

$p \in \Delta(\mathcal{Y})$

Precise Belief

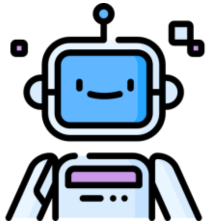


Precise Forecaster

$p(Y) = 0.5$

$\mathcal{P} \subseteq \Delta(\mathcal{Y})$

Imprecise Belief

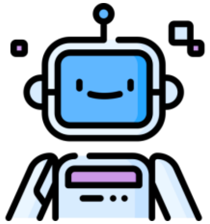


Imprecise Forecaster

$\mathcal{P}(Y) = [0.2, 0.8]$

What is an imprecise forecaster?

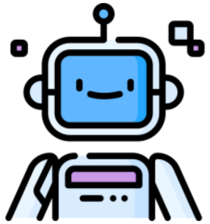
$p \in \Delta(\mathcal{Y})$
Precise
Belief



Precise
Forecaster

$p(Y) = 0.5$

$\mathcal{P} \subseteq \Delta(\mathcal{Y})$
Imprecise
Belief

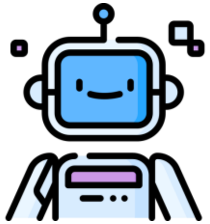


Imprecise
Forecaster

$\mathcal{P} = \{p_1, p_2, p_3\}$

What is an imprecise forecaster?

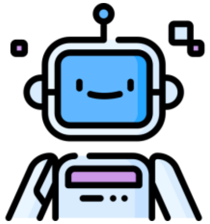
$p \in \Delta(\mathcal{Y})$
Precise Belief



Precise Forecaster

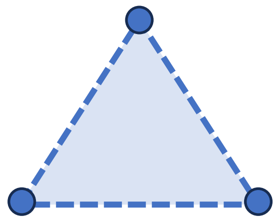
$p(Y) = 0.5$

$\mathcal{P} \subseteq \Delta(\mathcal{Y})$
Imprecise Belief



Imprecise Forecaster

convex hull

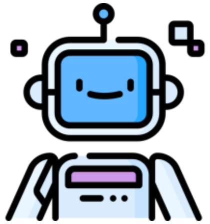


$\mathcal{P} = \{p_1, p_2, p_3\}$

Credal set of an imprecise forecast

$p \in \Delta(\mathcal{Y})$

Precise Belief

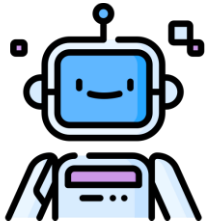


Precise Forecaster

$p(Y) = 0.5$

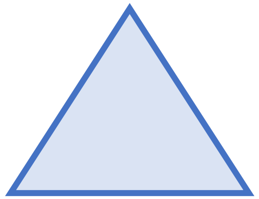
$\mathcal{P} \subseteq \Delta(\mathcal{Y})$

Imprecise Belief



Imprecise Forecaster

closed convex hull

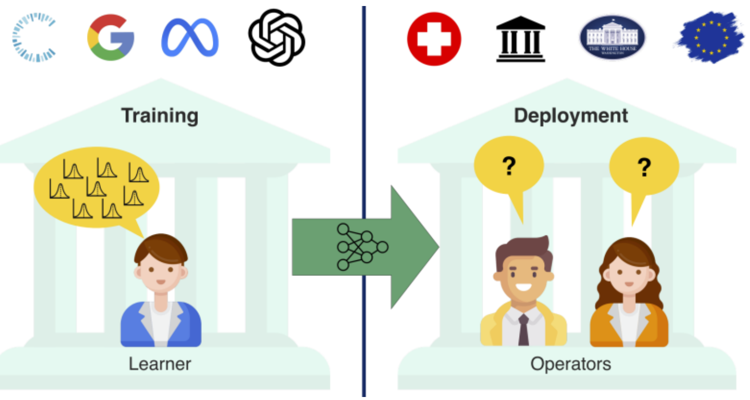


credal set $\mathcal{P}$

Why do we care about credal sets?

Can we train predictors that generalise for all downstream decision makers ?

Yes, with a credal set of models.



Institutional Separation

Domain Generalisation via Imprecise Learning

Anurag Singh [1] **Siu Lun Chau** [1] **Shahine Bouabid** [2] **Krikamol Muandet** [1]

Abstract

Out-of-distribution (OOD) generalisation is challenging because it involves not only learning from empirical data, but also deciding among various notions of generalisation, e.g., optimising the

(LLM) that surpass human-level generalisation capabilities in specific domains.

Despite notable achievements, these systems may catastrophically fail when operated on out-of-domain (OOD) data because theoretical guarantees for their generalisation

2024

ICML International Conference On Machine Learning

Appeared in ICML 2024 as spotlight

Abstracting the same question

What loss to use in ERM to obtain the credal set of models.

Wait....

Was this not impossible?

The impossibility is only for the real-valued scoring rules! Ours is *random*

Truthful Elicitation of Imprecise Forecasts

Anurag Singh[1] **Siu Lun Chau**[2] **Krikamol Muandet**[1]

[1]Rational Intelligence Lab, CISPA Helmholtz Center for Information Security, Saarbrücken, Germany
[2]College of Computing & Data Science, Nanyang Technological University, Singapore

Abstract

ing and Raftery, 2007]. They also serve as a backbone for

Abstracting the same question

What loss to use in ERM to obtain the credal set of models.

Wait....

Was this not impossible?

The impossibility is only for the real-valued scoring rules! Ours is *random*

Truthful Elicitation of Imprecise Forecasts

Anurag Singh[1] **Siu Lun Chau**[2] **Krikamol Muandet**[1]

[1]Rational Intelligence Lab, CISPA Helmholtz Center for Information Security, Saarbrücken, Germany
[2]College of Computing & Data Science, Nanyang Technological University, Singapore

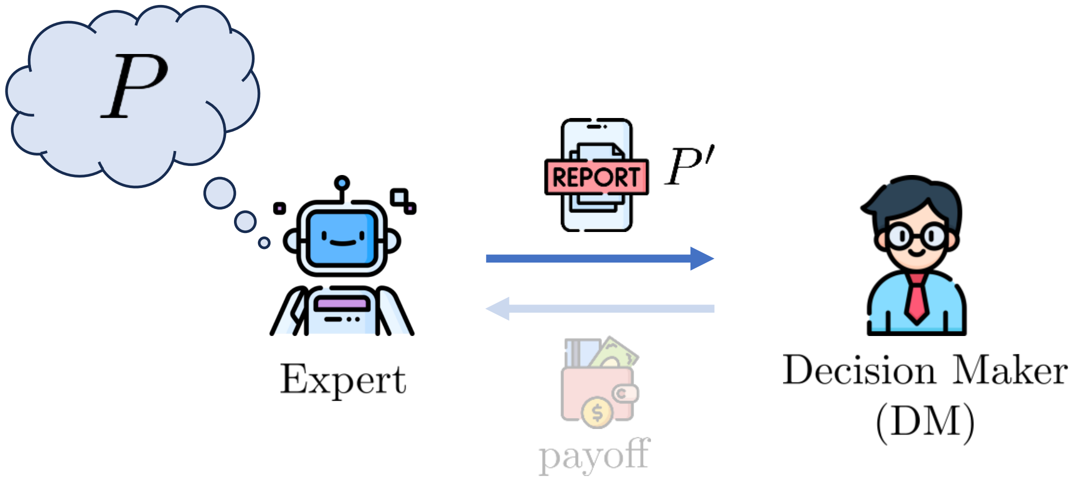
Abstract

ing and Raftery, 2007]. They also serve as a backbone for

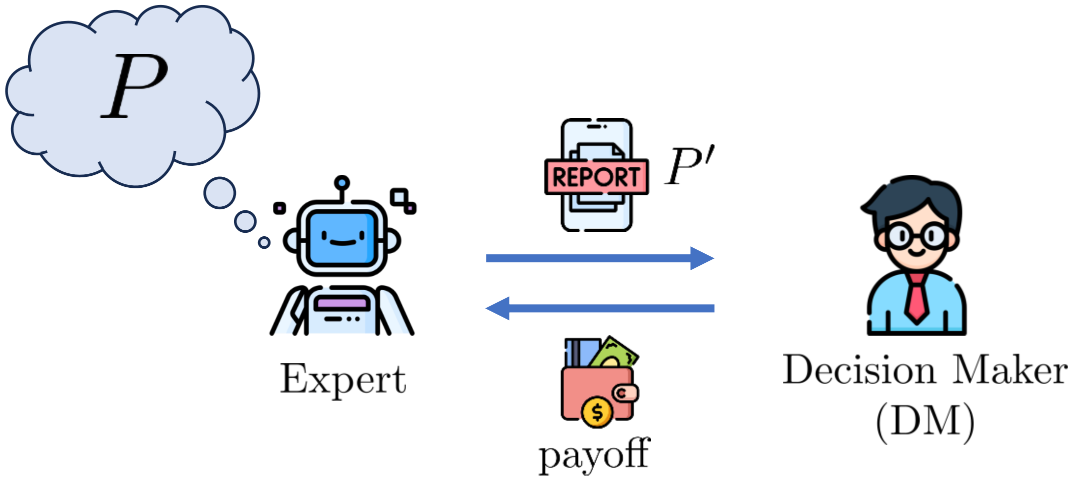
Our Contributions:

- Formalise the role of decision maker in imprecise forecast elicitation.
- Circumvent prior impossibility results to propose a **strictly proper** randomised scoring rule!
- Bonus: Connection to social choice theory

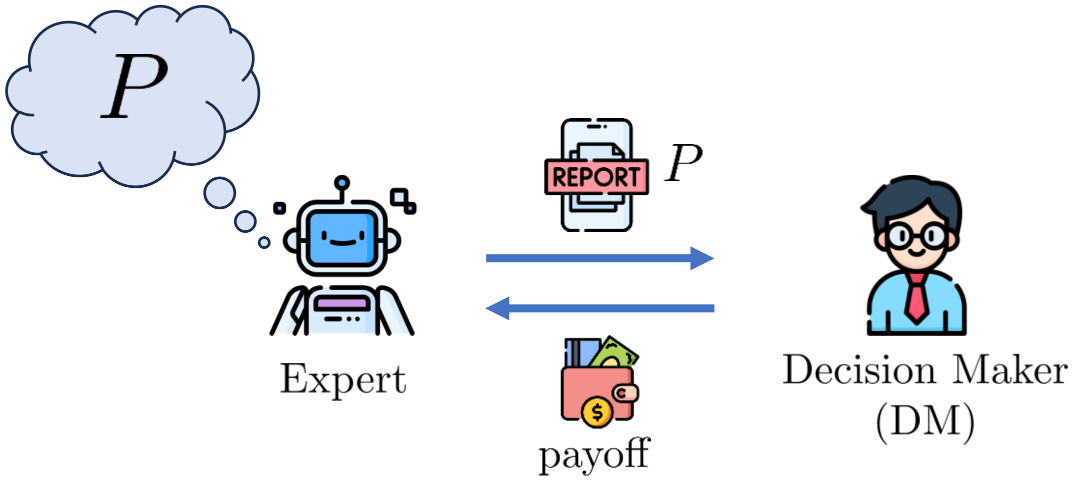
What is Elicitation



What is Elicitation



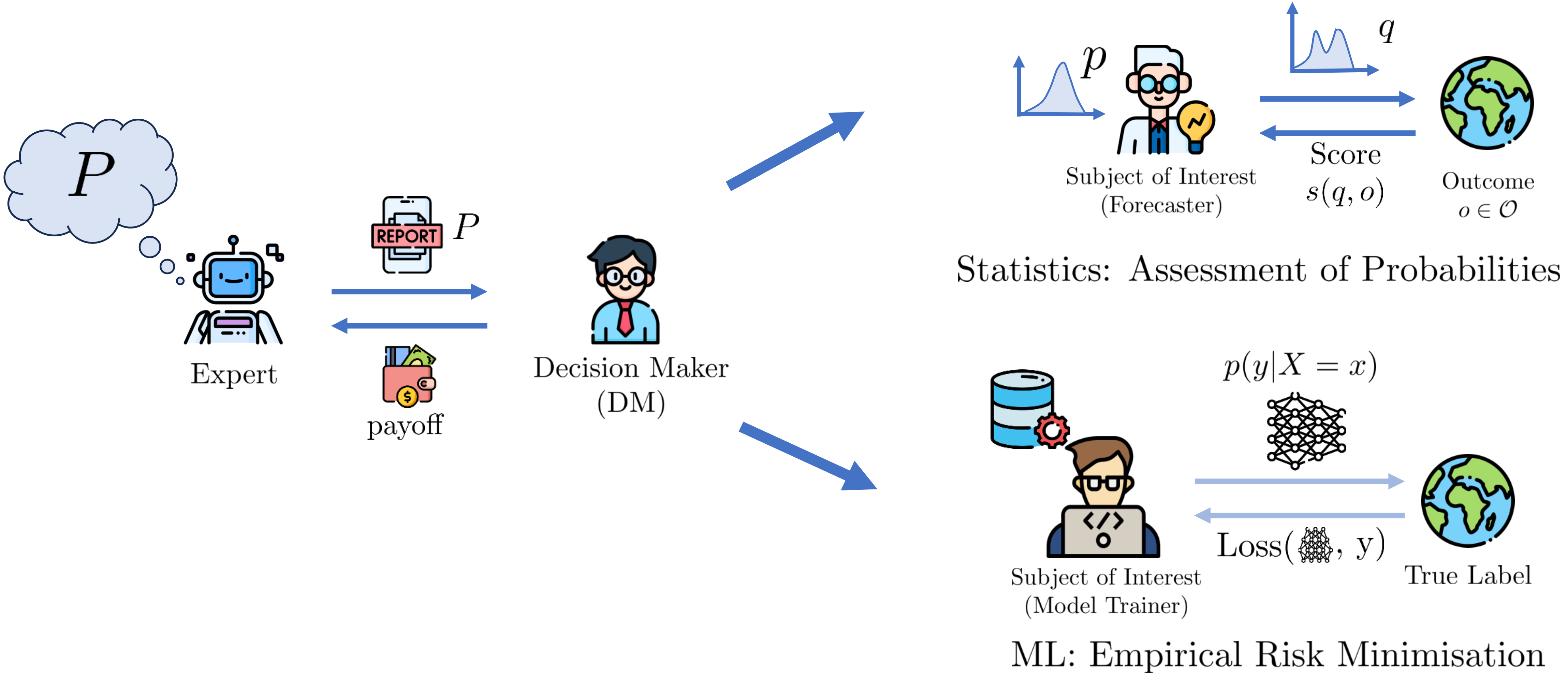
What is Elicitation



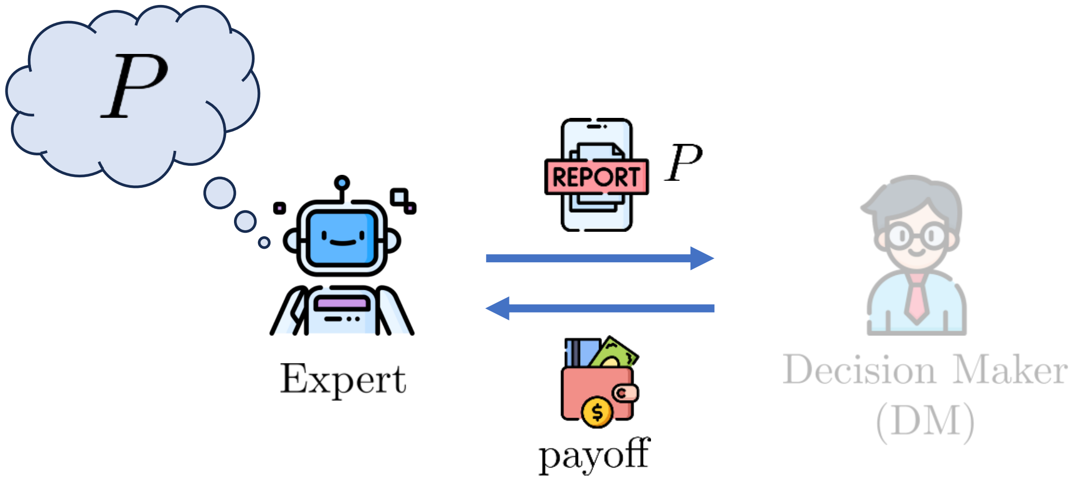
A payoff mechanism performs **truthful** elicitation if it can incentivise the expert to report their true belief.

… In other words, speaking truth is dominant strategy for expert.

Applications of truthful elicitation



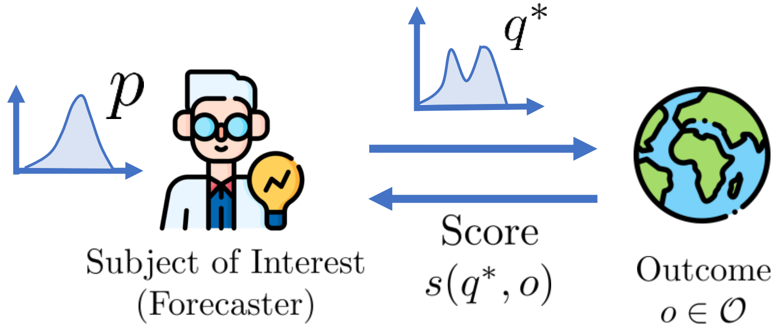
Truthful elicitation of precise forecasts



$$\mathbb{E}_{o\sim p}[s(q^*, o)]$$

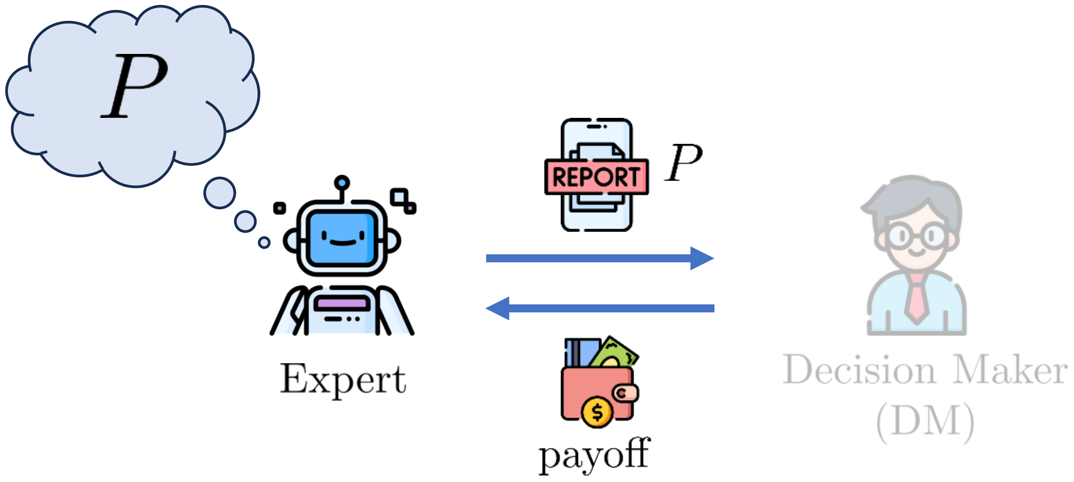
$$q^* \succeq_p q_2 \succeq_p \cdots$$

$\succeq_p$ is complete



Elicitation of precise forecasts ignores the DM!

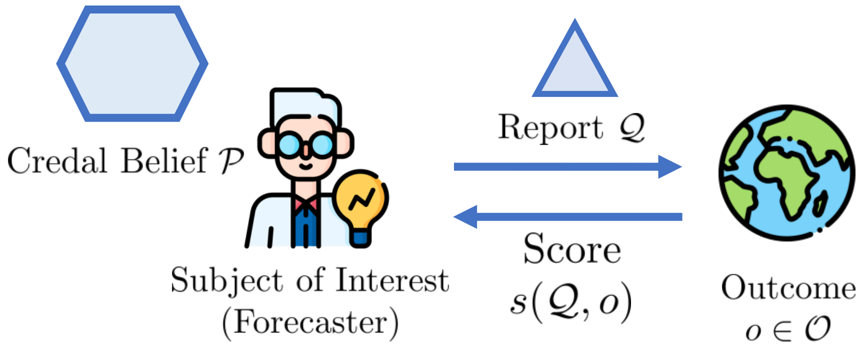
Truthful elicitation of imprecise forecasts



$$\{\mathbb{E}_{o\sim p}[s(\mathcal{Q},o)]\}_{p\in\mathcal{P}}$$

$$\overset{???}{Q \succeq_{\mathcal{P}} Q'}$$

$\succeq_{\mathcal{P}}$ is incomplete



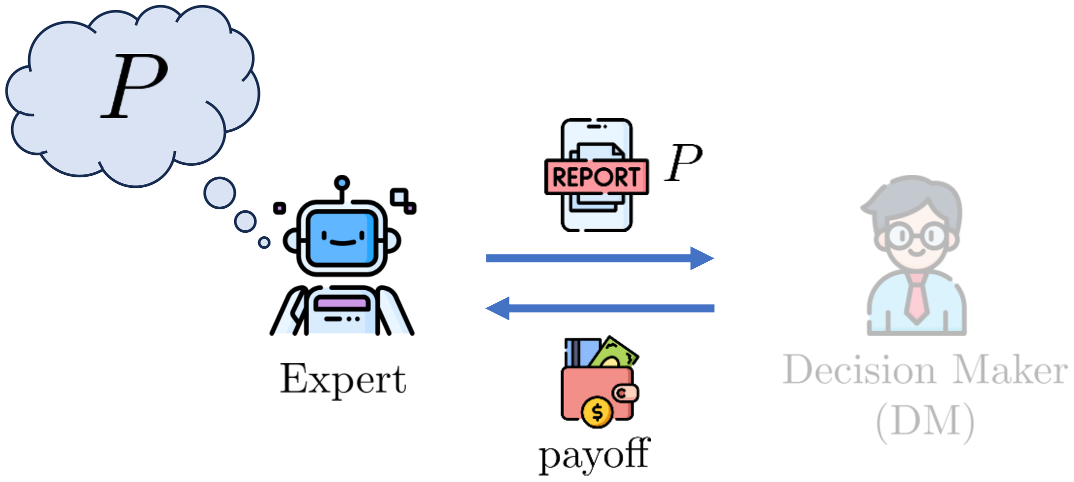
Truthful elicitation of imprecise forecasts

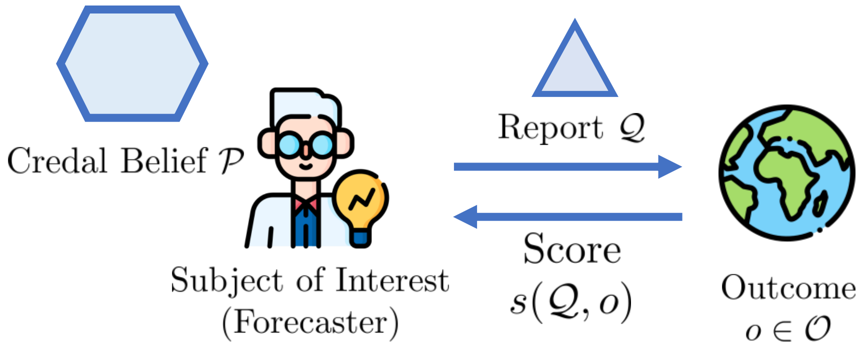
$$\{\mathbb{E}_{o\sim p}[s(\mathcal{Q},o)]\}_{p\in\mathcal{P}}$$

???

$$Q \succeq_{\mathcal{P}} Q'$$

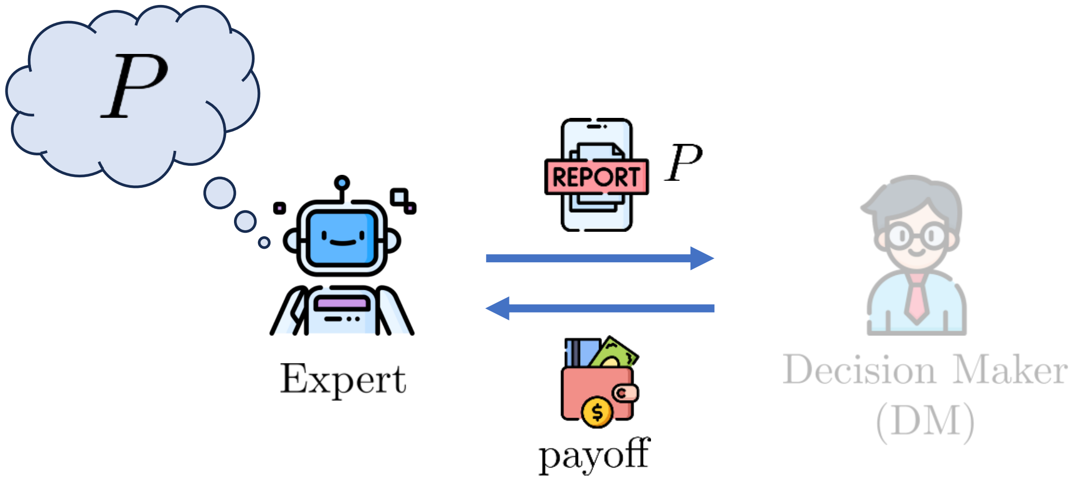
$\succeq_{\mathcal{P}}$ is incomplete





Impossibllity results on IP scoring rules
(Seidenfeld 2012, Mayo-Wilson 2015 and Schoenfield, 2017)

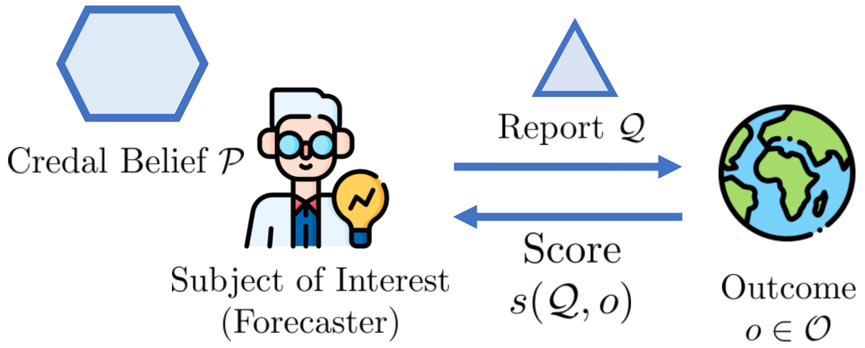
Truthful elicitation of imprecise forecasts



$$\{\mathbb{E}_{o\sim p}[s(\mathcal{Q},o)]\}_{p\in\mathcal{P}}$$

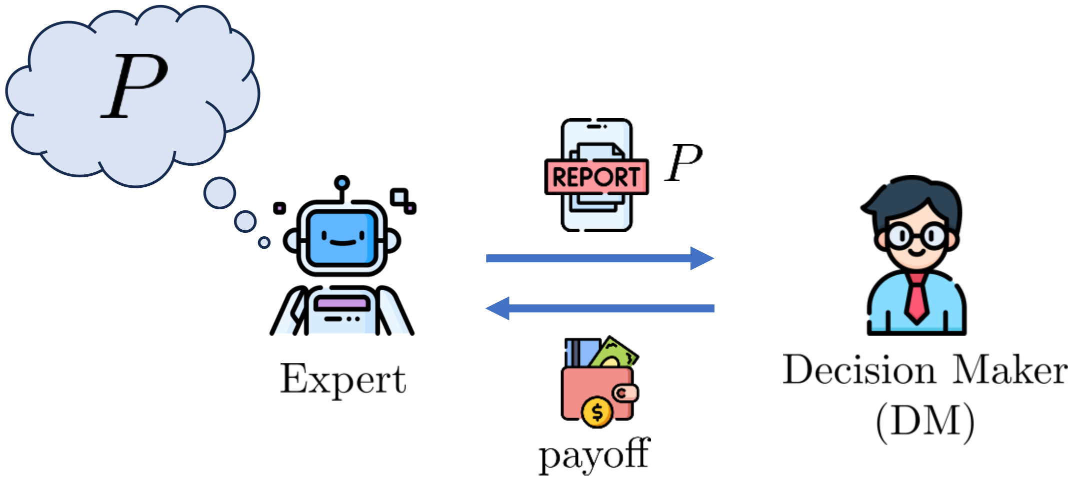
$$\overset{???}{Q \succeq_{\mathcal{P}} Q'}$$

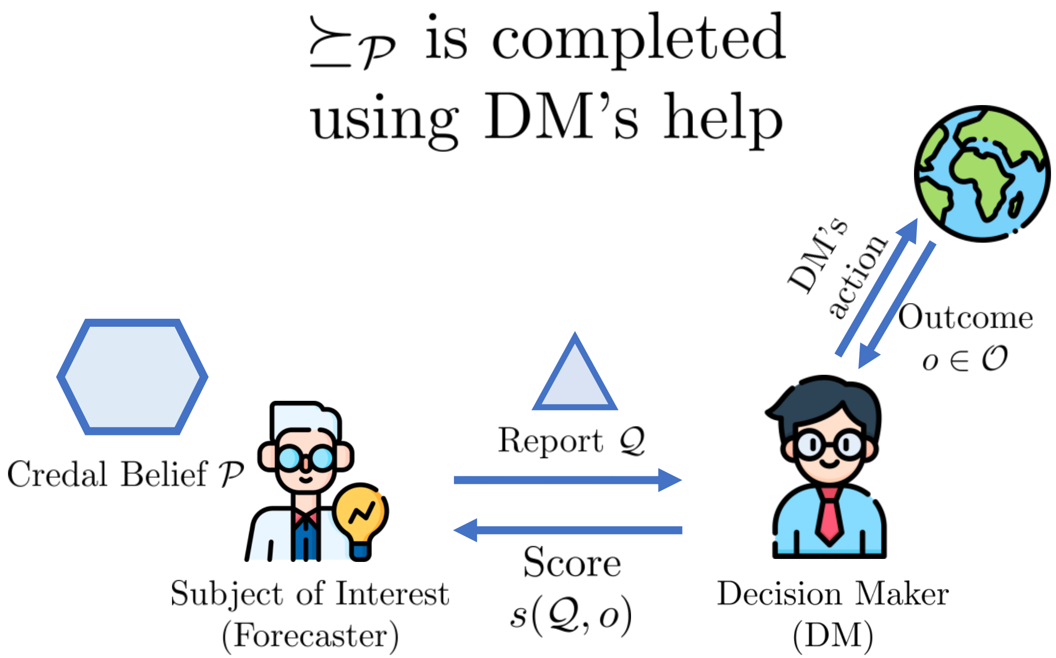
$\succeq_{\mathcal{P}}$ is incomplete



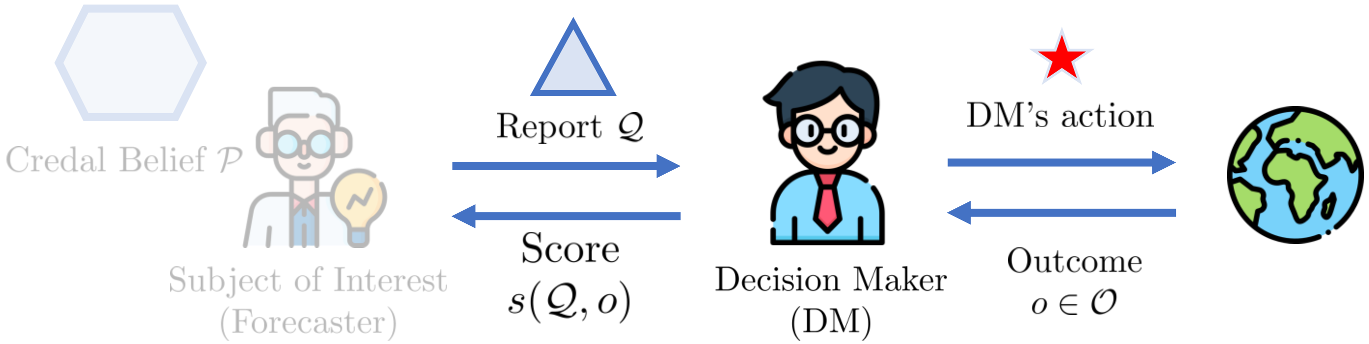
For any belief $\mathcal{P} \in 2^{\Delta(\mathcal{O})}$ the only proper scoring rule $s : 2^{\mathcal{O}} \times \mathcal{O} \to \mathbb{R}$ is for a $k \in \mathbb{R}$, $s(\mathcal{Q},o) = k$ for all $Q \in 2^{\Delta(\mathcal{O})}$

Truthful elicitation of imprecise forecasts

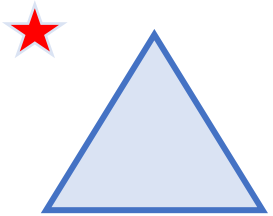




What is the exact form of this help?



What is the exact form of this help?



$\mathcal{Q} = \{q_1, q_2, q_3\}$

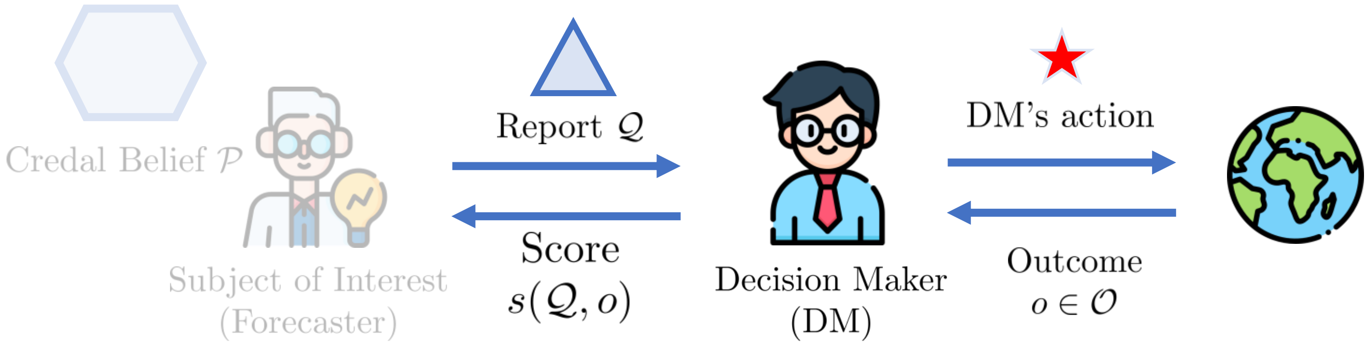
comparing actions

a and a'

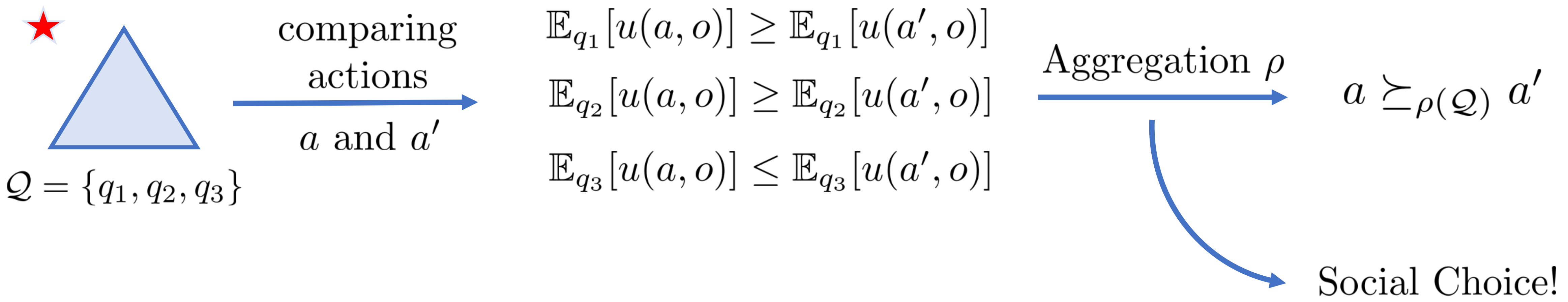
$$\mathbb{E}_{q_1}[u(a,o)] \geq \mathbb{E}_{q_1}[u(a',o)]$$
$$\mathbb{E}_{q_2}[u(a,o)] \geq \mathbb{E}_{q_2}[u(a',o)]$$
$$\mathbb{E}_{q_3}[u(a,o)] \leq \mathbb{E}_{q_3}[u(a',o)]$$

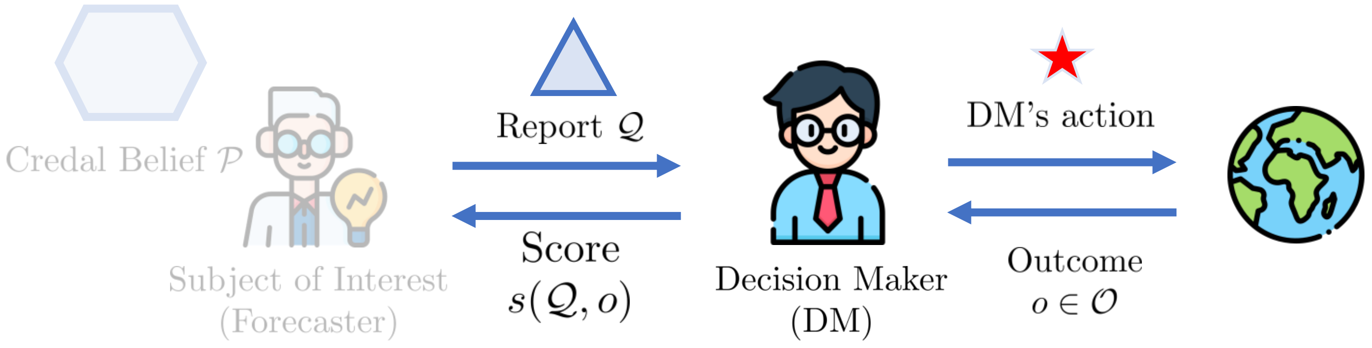
Aggregation ρ

$$a \succeq_{\rho(\mathcal{Q})} a'$$

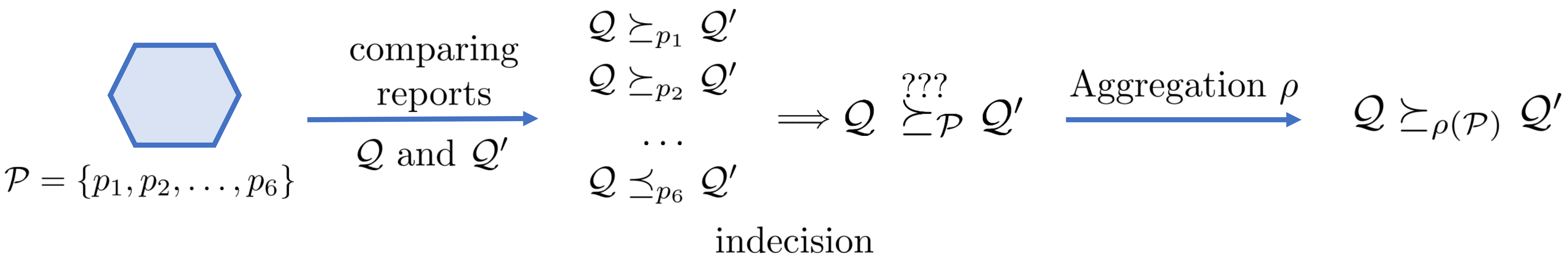


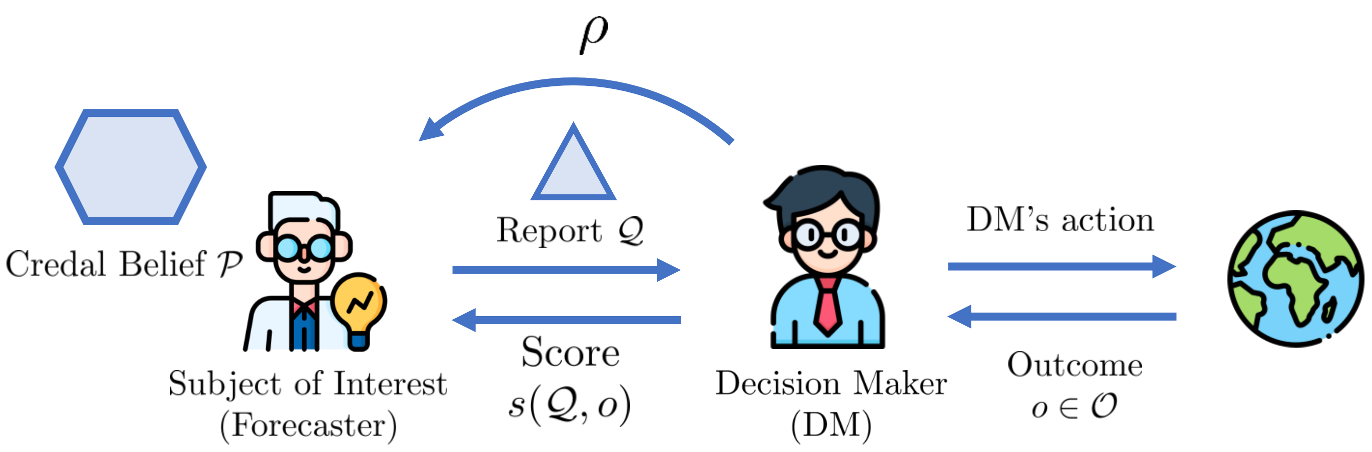
What is the exact form of this help?





What is the exact form of this help?

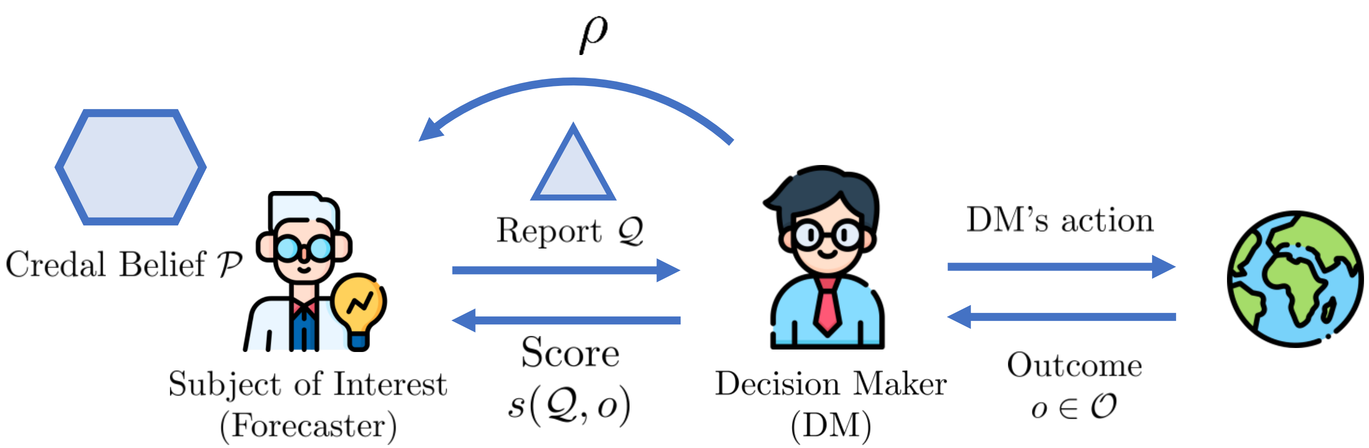




IP Scoring Rules: How to score with a DM

$$s_\rho(\mathcal{Q}, o) = ku(a^*_{\mathcal{Q},\rho}, o) + c \;\text{ where } k, c \in \mathbb{R}_{\geq 0} \text{ and } a^*_{\mathcal{Q},\rho} = arg \max_{a \in \mathcal{A}} \rho(\{\mathbb{E}_q[u(a,o)]\}_{q \in \mathcal{Q}})$$

Proposition: s_ρ is proper for all ρ.

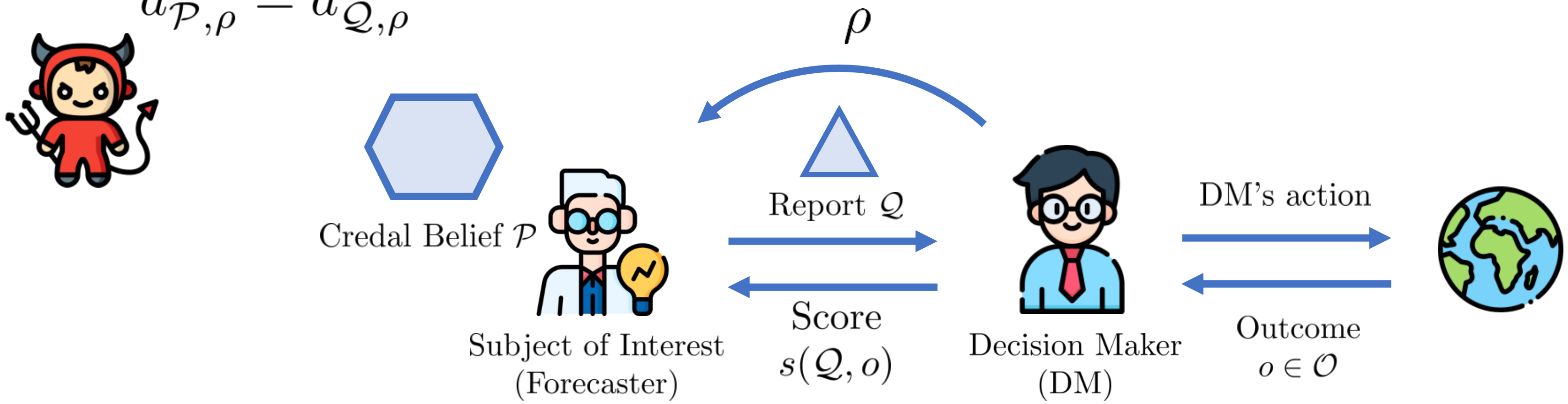


IP Scoring Rules: How to score with a DM

$$s_\rho(\mathcal{Q}, o) = k u(a^*_{\mathcal{Q},\rho}, o) + c \;\text{ where } k, c \in \mathbb{R}_{\geq 0} \text{ and } a^*_{\mathcal{Q},\rho} = arg \max_{a \in \mathcal{A}} \rho(\{\mathbb{E}_q[u(a, o)]\}_{q \in \mathcal{Q}})$$

Proposition: s_ρ is proper for all ρ.

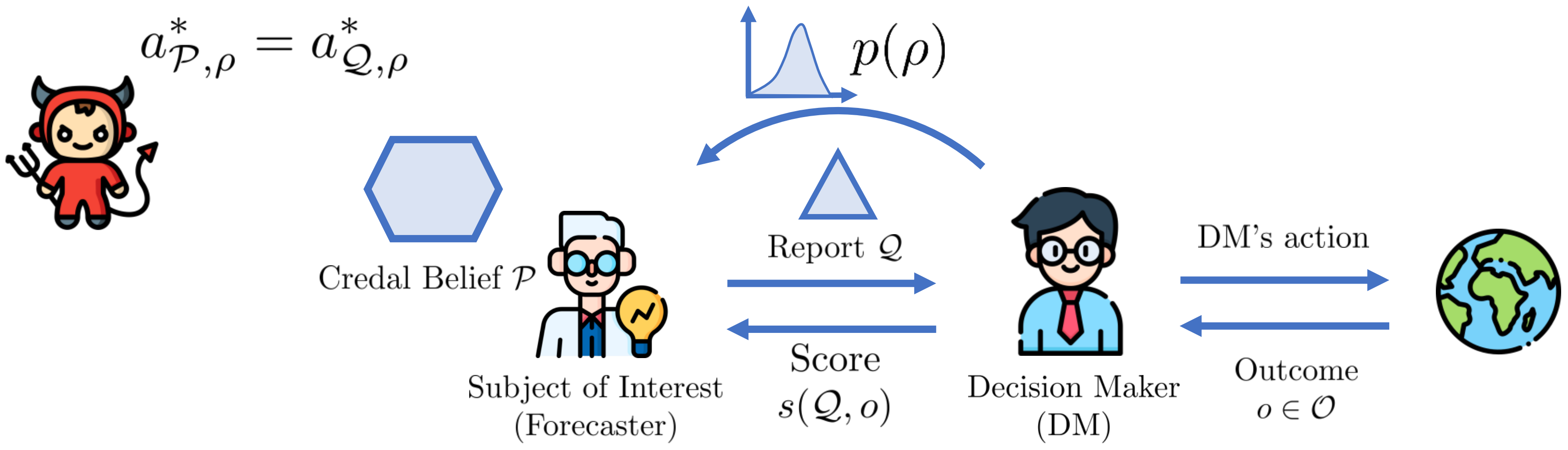
$$a^*_{\mathcal{P},\rho} = a^*_{\mathcal{Q},\rho}$$



IP Scoring Rules: How to score with a DM

$$s_\rho(\mathcal{Q}, o) = k u(a^*_{\mathcal{Q},\rho}, o) + c \;\text{ where } k, c \in \mathbb{R}_{\geq 0} \text{ and } a^*_{\mathcal{Q},\rho} = arg \max_{a \in \mathcal{A}} \rho(\{\mathbb{E}_q[u(a, o)]\}_{q \in \mathcal{Q}})$$

Proposition: s_ρ is proper for all ρ.

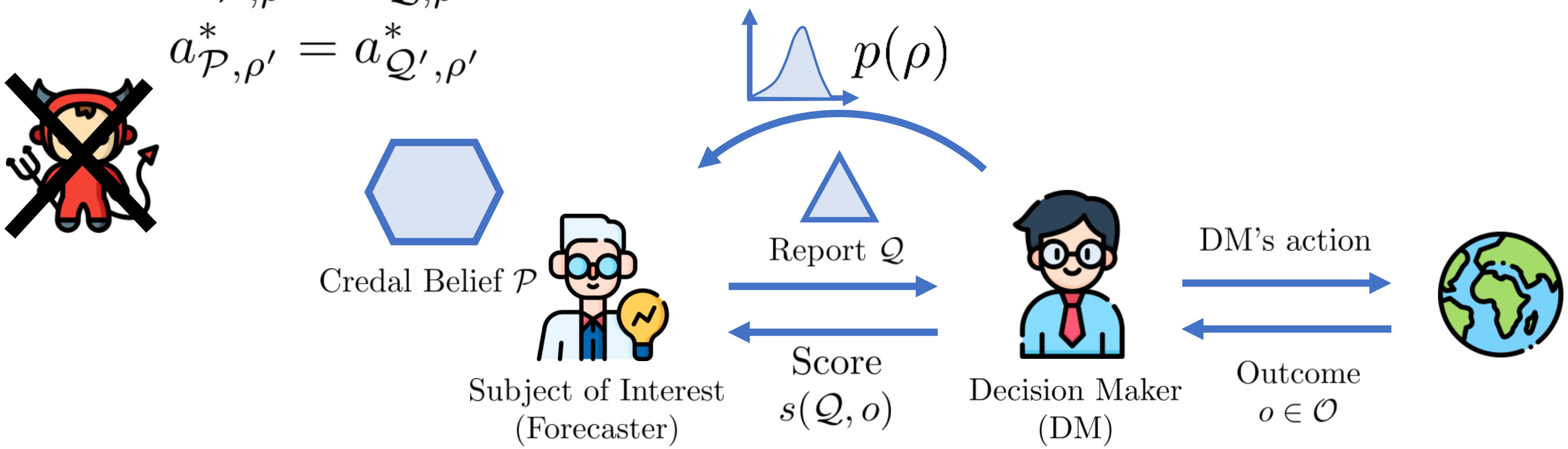


IP Scoring Rules: How to score with a DM

$$s_\rho(\mathcal{Q}, o) = k u(a^*_{\mathcal{Q},\rho}, o) + c \;\text{ where } k, c \in \mathbb{R}_{\geq 0} \text{ and } a^*_{\mathcal{Q},\rho} = arg\max_{a \in \mathcal{A}} \rho(\{\mathbb{E}_q[u(a,o)]\}_{q \in \mathcal{Q}})$$

Proposition: s_ρ is proper for all ρ.

$$a^*_{\mathcal{P},\rho} = a^*_{\mathcal{Q},\rho}$$
$$a^*_{\mathcal{P},\rho'} = a^*_{\mathcal{Q}',\rho'}$$



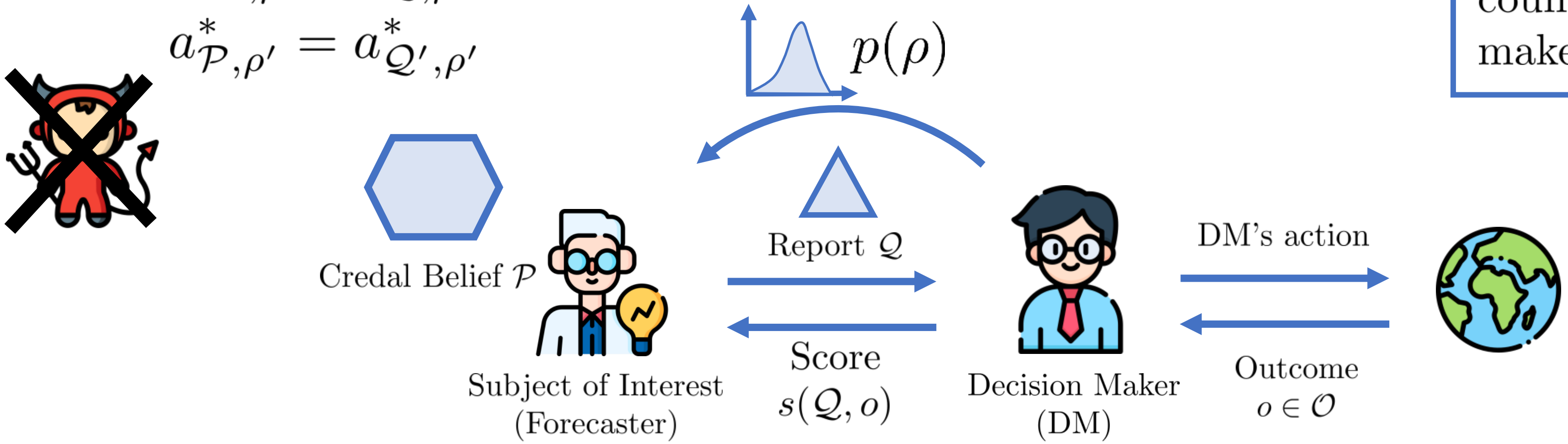
Strictly proper IP scoring rule

Theorem: s_ρ is strictly proper for $p(\rho)$ with full support.

$$s_\rho(\mathcal{Q}, o) = \begin{cases} k_\rho u(a^*_{\rho,\mathcal{Q}}, o) + c_\rho & \text{if } p(\rho) > 0 \\ \Pi_o(\mathcal{Q}) & \text{if } p(\rho) = 0 \end{cases}.$$

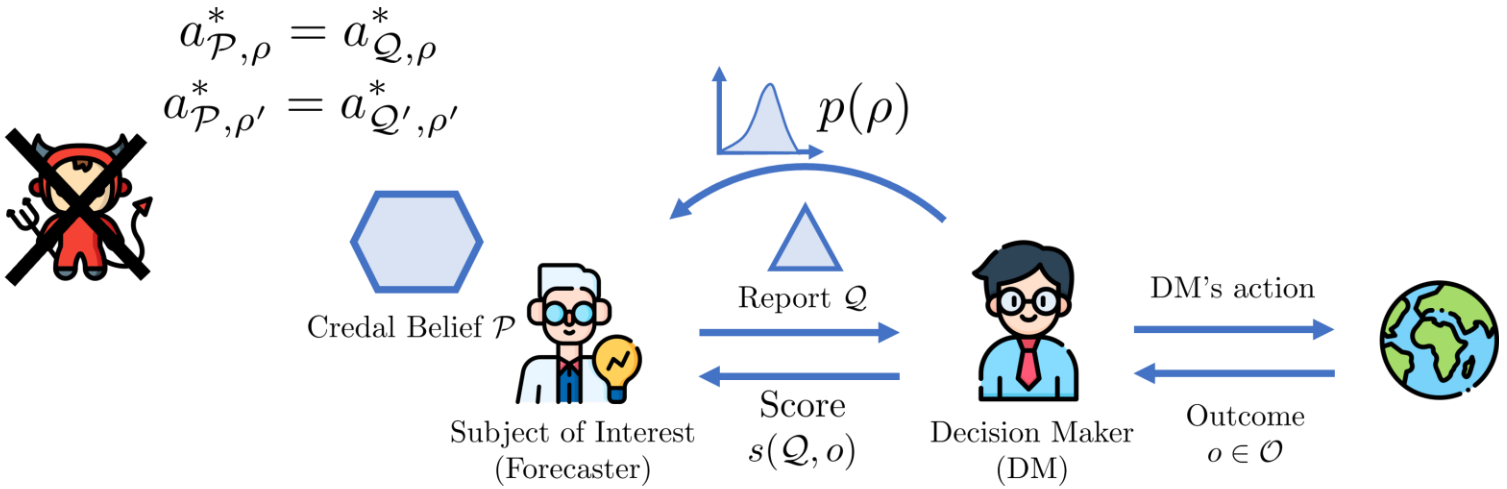
By not telling which assignment counts for the final grade. I can make students do all of them.

$$a^*_{\mathcal{P},\rho} = a^*_{\mathcal{Q},\rho}$$
$$a^*_{\mathcal{P},\rho'} = a^*_{\mathcal{Q}',\rho'}$$



Conclusion

1. Allow experts and algorithms to say "I don't know exactly, but it's between a and b"
2. Design of systems that explicitly embrace—not suppress—epistemic uncertainty.
3. Honest communcation of uncertainty for trustworthy decisions.



Come visit our poster 176!