Epistemic Argumentation Framework

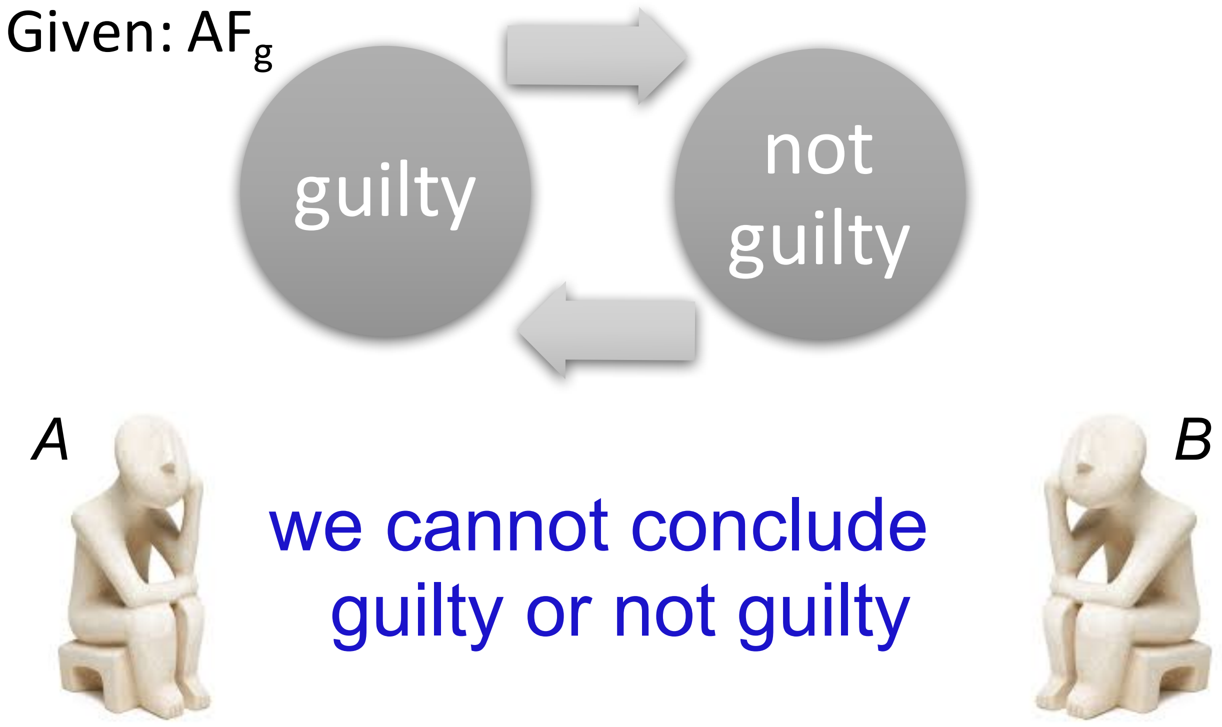
Chiaki Sakama
(Wakayama University, Japan)

Tran Cao Son
(New Mexico State University, USA)

PRICAI 2019@FIJI

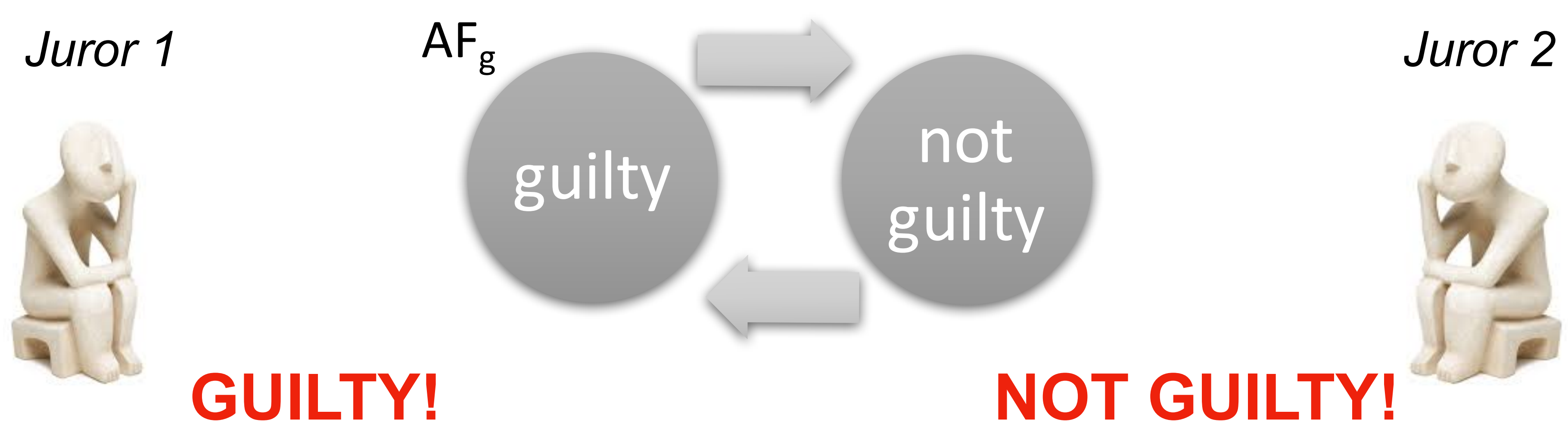
Background and Motivation

- An **abstract argumentation framework (AF)** represents arguments and attacks provided by players in a debate or a dialogue game.
- Arguments and attacks in AF are open and shared, then AF semantics provides a result that is to be agreed by **all** players.



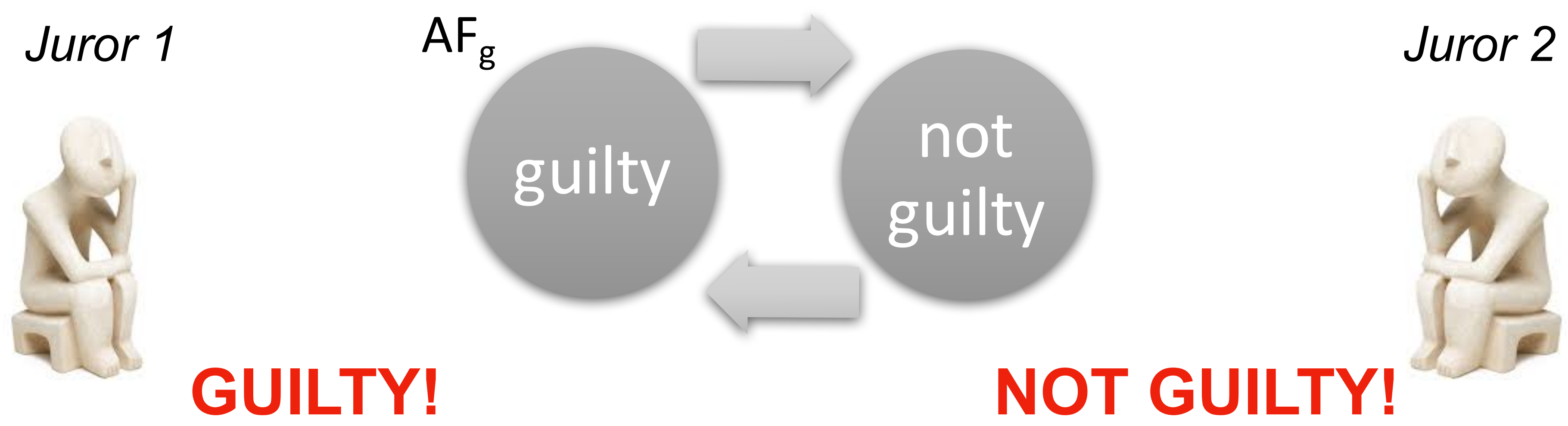
Background and Motivation

- A player may not agree with the result or may have a **bias** towards a particular argument.
- For instance, in a court case jurors share the same open AF while could reach different conclusions based on their biases.



Background and Motivation

- A player may not agree with the result or may have a **bias** towards a particular argument.
- For instance, in a court case jurors share the same open AF while could reach different conclusions based on their biases.



- **Why so?**
- **How do we formulate this phenomenon?**

Contribution

- We introduce an **Epistemic Argumentation Framework** (**EAF**) that incorporates agent's beliefs into an argumentation framework.
- We apply EAFs to representing **preferences** and decision making in **multi-agent** environments.
- We analyze computational complexity of EAFs.

Argumentation Framework

- An **argumentation framework** is a pair $AF=(Ar, att)$ where Ar is a finite set of arguments and $att \subseteq Ar \times Ar$ is attack relations.

- A **labelling** of (Ar, att) is a total function $\mathscr{L} : Ar \rightarrow \{ in, out, und \}$, where *in* (accepted), *out* (rejected), and *und* (undecided).
  It is represented by a set $S(\mathscr{L}) = \{\lambda(x) \mid \mathscr{L}(x) = \lambda \text{ for } x \in Ar\}$.

- We consider the **complete labelling** (*co*), **stable labelling** (*st*), **preferred labelling** (*pr*), and **grounded labelling** (*gr*).
  (Often referred to ω-labellings of *AF* where $\omega \in \{co, st, pr, gr\}$).

Epistemic Formula

- Given *AF*=(*Ar*, *att*), define $\mathscr{A}_{AF}$ = { *in*(a), *out*(a), *und*(a) : $a \in Ar$ }.

- A propositional formula φ over $\mathscr{A}_{AF}$ is **true** in a labelling $\mathscr{L}$ (written $\mathscr{L} \models \varphi$) if φ is interpreted to be true under $S(\mathscr{L})$.

- An **epistemic atom** over *AF* is of the form **K**φ or **M**φ where φ is a propositional formula over $\mathscr{A}_{AF}$. **K**φ means an agent believes that φ is true, and **M**φ means an agent believes that φ is possibly true.

- An **epistemic formula** is a propositional formula constructed over epistemic atoms together with $\top$ and $\bot$.

Example

- Consider the AF:

  Guilty $\longleftrightarrow$ Innocent

  where $\mathscr{A}_{AF} =$ { *in*(*G*), *out*(*G*), *und*(*G*), *in*(*I*), *out*(*I*), *und*(*I*) }.

- **K** (*in*(*G*) $\vee$ *out*(*G*)) means an agent believes that *in*(*G*) $\vee$ *out*(*G*) is true.
  (''The accused is either guilty or not guilty'')

- **M (**$\neg$*in*(*G*)) $\rightarrow$ **K**(*in*(*I*)) means if an agent believes that $\neg$*in*(*G*) is possibly true then he/she believes that *in*(*I*) is true.
  (''The accused is innocent unless proven guilty'')

Satisfaction

A set **SL** of labellings **satisfies** an epistemic formula φ (written **SL** $\models \varphi$) if one of the following conditions hold:

- $\varphi = \top$
- $\varphi = \mathbf{K}\psi$ and $\mathscr{L} \models \psi$ for every $\mathscr{L} \in$ **SL**
- $\varphi = \mathbf{M}\psi$ and $\mathscr{L} \models \psi$ for some $\mathscr{L} \in$ **SL**
- $\varphi = \neg\psi$ and **SL** $\not\models \psi$
- $\varphi = \varphi_1 \vee \varphi_2$ and (**SL** $\models \varphi_1$ or **SL** $\models \varphi_2$)
- $\varphi = \varphi_1 \wedge \varphi_2$ and (**SL** $\models \varphi_1$ and **SL** $\models \varphi_2$)

Epistemic Argumentation Framework

- An **epistemic argumentation framework (EAF)** is a triple (Ar, att, φ) where $AF=(Ar, att)$ is an argumentation framework and φ is an epistemic formula (called an **epistemic constraint**).
  An EAF is also written as (AF, φ).

- A set **SL** of labelings is an **ω-epistemic labelling set** of (AF, φ) if
  (i) each $\mathscr{L} \in$ **SL** is an ω-labelling of *AF*, and
  (ii) **SL** is a $\subseteq$-maximal set of ω-labellings of *AF* that satisfy φ,
  where $\omega \in \{co, st, pr, gr\}$.

Example

- A person plans to travel to Fiji or Macau. He/she does not travel to Macau if Hong Kong Airport is closed.

- The situation is represented by the *AF*:

  Fiji $\longleftrightarrow$ Macau $\longleftarrow$ Close $\longleftrightarrow$ Open

- The above *AF* has 3 stable labellings:

  { *in*(*F*), *out*(*M*), *in*(*C*), *out*(*O*) },
  { *in*(*F*), *out*(*M*), *out*(*C*), *in*(*O*) },
  { *out*(*F*), *in*(*M*), *out*(*C*), *in*(*O*) }.

Example (cont.)

- He/she prefers traveling to Macau unless the Airport is closed.
- The belief is encoded by the epistemic constraint:

$$\varphi_1 = \mathbf{M}\, in(O) \rightarrow \mathbf{K}\, in(M)$$

("if *Open* is possibly accepted, then *Macau* should be accepted")

- $EAF_1 = (AF, \varphi_1)$ has the unique stable epistemic labelling set:

{ ~~{ *in(F), out(M), in(C), out(O)* },~~
~~{ *in(F), out(M), out(C), in(O)* },~~
{ *out(F), in(M), out(C), in(O)* } }.

Example (cont.)

- It turns that the Airport is closed. The situation is represented by $EAF_2 = (AF, \varphi_2)$ where

$$\varphi_2 = \varphi_1 \wedge \mathbf{K}\, in(C).$$

- EAF_2 has the unique stable epistemic labelling set:

{ { *in*(*F*), *out*(*M*), *in(C)*, *out(O)* },
~~{ *in(F)*, *out(M)*, *out(C)*, *in(O)* }~~,
~~{ *out(F)*, *in(M)*, *out(C)*, *in(O)* }~~ }.

- As such, an EAF can represent belief change of an agent by revising an epistemic constraint without modifying AF.

Representing Preference

- Given *AF*=(*Ar*, *att*) and a pre-order relation $\sqsupseteq \subseteq \mathscr{A}_{AF} \times \mathscr{A}_{AF}$, define EAF=(*AF*, φ_J) where

$$\varphi_J = \bigwedge_{\lambda(x) \sqsupset \mu(y)} \mathbf{K}\,(\mu(y) \supset \lambda(x))$$

and $\lambda, \mu \in \{in, out, und\}$ and $x, y \in Ar$.

- φ_J states that if the justification state $\lambda(x)$ is preferred to $\mu(y)$, then $\mathscr{L} \vDash \mu(x)$ implies $\mathscr{L} \vDash \lambda(x)$ for any $\mathscr{L} \in$ **SL** where **SL** is any ω-epistemic labelling set of EAF.

Example

- Consider the *AF*:

$$\text{Fiji} \longleftrightarrow \text{Macau} \longleftarrow \text{Close} \longleftrightarrow \text{Open}$$

- Whether Close or Open is undecided, it is specified as

$$\varphi_J = \bigwedge_{x \in \{C,O\}} \mathbf{K}\,(in(x) \supset und(x)) \wedge \mathbf{K}\,(out(x) \supset und(x))$$

- EAF=(*AF*, φ_J) has the unique preferred epistemic labeling set:

{ { *in(F), out(M), und(C), und(O)* } }.

Multiple Agents

- Consider multiple agents who share AF while having different beliefs. The situation is represented by the collection of EAFs: $EAF_i = (AF, \varphi_i)\ (i = 1,...,n)$.

- $EAF_1, \ldots, EAF_n$ **credulously agree on** $\lambda(a)$ for $a \in Ar$ where $\lambda \in \{in, out, und\}$ under ω-epistemic labelling if each EAF_i has an ω-epistemic labelling set $\mathbf{SL}_i$ such that $\mathbf{SL}_i \models \mathbf{M}\,\lambda(a)$.

- $EAF_1, \ldots, EAF_n$ **skeptically agree on** $\lambda(a)$ under ω-epistemic labelling if for any ω-epistemic labelling set $\mathbf{SL}_i$ of EAF_i , $\mathbf{SL}_i \models \mathbf{K}\,\lambda(a)$.

Majority Voting

- Define:

  $M_{\psi}^{\omega} = \{\, i \mid EAF_i$ has an ω-epistemic labelling set $\mathbf{SL}$ s.t. $\mathbf{SL} \models \mathbf{M}\psi \,\}$,
  $N_{\psi}^{\omega} = \{\, i \mid$ for each ω-epistemic labelling set $\mathbf{SL}$ of EAF_i, $\mathbf{SL} \models \mathbf{K}\psi \,\}$.

- $\lambda(a)$ is **credulously (resp. skeptically) adopted by majority voting** under ω-epistemic labelling iff the cardinality of the set $M_{\lambda(a)}^{\omega}$ (resp. $N_{\lambda(a)}^{\omega}$) is greater than the cardinality of $M_{\mu(a)}^{\omega}$ (resp. $N_{\mu(a)}^{\omega}$) where $\lambda, \mu \in \{in, out, und\}$ and $\lambda \neq \mu$.

- When $\mid M_{\lambda(a)}^{\omega} \mid = n$ (resp. $\mid N_{\lambda(a)}^{\omega} \mid = n$), $EAF_1, \ldots, EAF_n$ credulously (resp. skeptically) agree on $\lambda(a)$.

Complexity

- Consider an epistemic formula φ in DNF that has at most *k* disjuncts and each disjunct contains at most *p* conjuncts where *p* and *k* are polynomial in the size of an AF.

- Deciding whether EAF=(*AF*, φ) has a non-empty ω-epistemic labelling set is done in polynomial time for $\omega = gr$ and NP-complete for $\omega \in \{co, st, pr\}$.

Comparisons

- EAF vs. **Probabilistic AF** (PAF):

  - PAF focuses on uncertainty of arguments rather than agent's belief

  - PAF merges objective knowledge and subjective beliefs

  - PAF may produce new extensions

- EAF vs. **AF with Preference** (AFP):

  - AFP specifies preference between arguments or attacks

  - EAF can specify preference over justification states

  - AFP often changes the original argumentation graph

Final Remark

- EAF represents an objective evidence in AF, while encodes subjective belief of individual agents by epistemic constraints.
- Such separation enables agents to produce different conclusions based on their biases toward a common AF.
- EAF is transformed to an **epistemic logic program** and epistemic labelling sets are computed by **answer set solvers**.
- Future study includes extending EAF to reasoning about beliefs of other agents and representing an agent's own belief based on beliefs of other agents.