# Argument and Belief

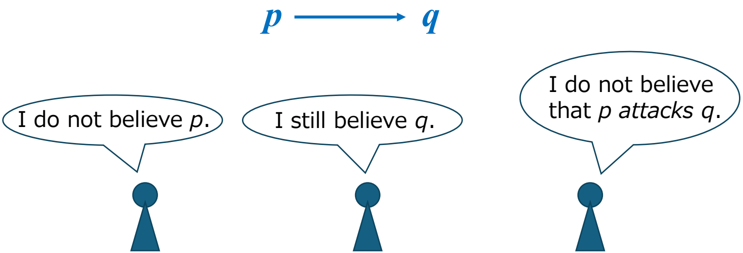
**Chiaki Sakama**

Wakayama University, Japan

COMMA 2024@Hagen

# Background & Motivation

- Given $AF = (\{p, q\}, \{(p, q)\})$, argumentation semantics normally concludes that $p$ is accepted and $q$ is rejected.
- To reject $p$, on the other hand, a counter-argument attacking $p$ is to be introduced.
- A player participating in an argumentation or a person in the audience of a public debate would have opinions s.t. "I do not believe $p$", "I still believe $q$", or "I do not believe that $p$ attacks $q$" without any concrete grounds.



# Contributions

- We introduce the framework of **AF with beliefs (AFB)** to represent interaction between **arguments** and **beliefs**.
- In AFB an agent's beliefs are added to the argumentation graph and **interact** with arguments.
- We introduce **axioms** for interlinking arguments and beliefs, and compute **belief extensions** that represent (dis)believed arguments as well as accepted arguments.
- We apply the framework to modelling the **audience** of argumentation, **dialogue** between two agents, and **inner conflict** of an agent.

# Representing belief in AF

- If an agent $a$ believes an argument $p$ (resp. an attack $p \to q$) to be true, it is represented as $B_a p$ (resp. $B_a(p \to q)$).
- When the agent's identification is unimportant, $a$ is omitted and it is simply written as $Bp$ or $B(p \to q)$.
- An agent's disbelieving $p$ (resp. $p \to q$) is represented by $\neg Bp$ (resp. $\neg B(p \to q)$).

Technically, we handle $p \to q$, $p \leftrightarrow q$, $(\neg)Bp$, $(\neg)B(p \to q)$ or $(\neg)B(p \leftrightarrow q)$ as an atom, so $B$ is not an operator in modal epistemic logic. In this setting, the "atom" $\neg\neg Bp$ is identified with $Bp$.

# AF with belief

Given an argumentation framework $AF = (A, R)$, the set $\mathcal{B}_{AF}$ of belief atoms over $AF$ is defined as
$\mathcal{B}_{AF} = \{ Bp, \neg Bp \mid p \in A \} \cup \{ B(p \to q), \neg B(p \to q) \mid (p, q) \in R \}$.

## AF with belief

Given $AF = (A, R)$, AF with belief (or AFB) is defined as a triple $\Gamma = (A, R, S)$ where $S \subseteq \mathcal{B}_{AF}$. $\Gamma$ is often written as $(AF, S)$.

## attacks over beliefs

Given $AF = (A, R)$, define
$R_B = R \cup \{ (\neg Bp, p), (\neg Bp, Bp), (Bp, \neg Bp) \mid p \in A \}$.

# Attack axiom

**attack axiom**

Let $p$ and $q$ be arguments. Then

$$(\mathbf{AT}) \qquad Bp \wedge B(p \rightarrow q) \supset \neg Bq$$

is called the attack axiom.

(**AT**) is rewritten as

$$Bq \wedge B(p \rightarrow q) \supset \neg Bp \quad \text{or} \quad Bp \wedge Bq \supset \neg B(p \rightarrow q).$$

# Closure

## $cl_{AT}(S)$

Given $S \subseteq \mathcal{B}_{AF}$, define $cl_{AT}(S) \subseteq \mathcal{B}_{AF}$ as the smallest set of belief atoms satisfying the following conditions:

1. $S \subseteq cl_{AT}(S)$.
2. If $Bp \in cl_{AT}(S)$ and $B(p \rightarrow q) \in cl_{AT}(S)$, then $\neg Bq \in cl_{AT}(S)$.
3. If $Bq \in cl_{AT}(S)$ and $B(p \rightarrow q) \in cl_{AT}(S)$, then $\neg Bp \in cl_{AT}(S)$.
4. If $Bp \in cl_{AT}(S)$ and $Bq \in cl_{AT}(S)$, then $\neg B(p \rightarrow q) \in cl_{AT}(S)$.

$cl_{AT}(S)$ is consistent if it does not contain $\{Bp, \neg Bp \mid p \in \mathcal{A}\}$ nor $\{B(p \rightarrow q), \neg B(p \rightarrow q) \mid p, q \in \mathcal{A}\}$ as a subset.

# Belief extension

Let $\sigma$ be an argumentation semantics.

**belief extension**

Given an AFB $\Gamma = (A, R, S)$, a set $E$ is a $\sigma$ belief extension of $\Gamma$ if $E$ is a $\sigma$ extension of $AF = (X, Y)$ with

$X = A \cup cl_{AT}(S)_A$,

$Y = ((X \times X) \cap R_B) \setminus \{(p \rightarrow q) \mid \neg B(p \rightarrow q) \in cl_{AT}(S)_R\}$,

where

$cl_{AT}(S)_A = cl_{AT}(S) \cap \{ Bp, \neg Bp \mid p \in A \}$,

$cl_{AT}(S)_R = cl_{AT}(S) \cap \{ B(p \rightarrow q), \neg B(p \rightarrow q) \mid (p \rightarrow q) \in R \}$, and

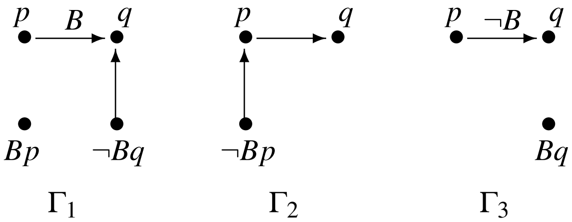
$R_B = R \cup \{ (\neg Bp, p), (\neg Bp, Bp), (Bp, \neg Bp) \mid p \in A \}$.

# Example (1)

Suppose an agent in the audience of a public debate.

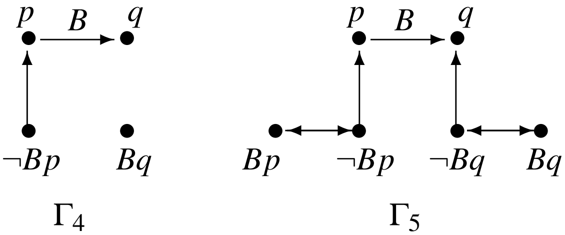
Let $AF = (\{p, q\}, \{(p, q)\})$ and $\sigma \in \{$(co)mplete, (st)able, (pr)eferred, (gr)ounded $\}$.

- $\Gamma_1 = (AF, \{Bp, B(p \rightarrow q)\})$ has the $\sigma$ belief extension $E_1 = \{p, Bp, \neg Bq\}$.
- $\Gamma_2 = (AF, \{\neg Bp\})$ has the $\sigma$ belief extension $E_2 = \{\neg Bp, q\}$.
- $\Gamma_3 = (AF, \{Bq, \neg B(p \rightarrow q)\})$ has the $\sigma$ belief extension $E_3 = \{p, q, Bq\}$.



# Example (2)

- $\Gamma_4 = (AF, \{Bq, B(p \to q)\})$ has the $\sigma$ belief extension $E_4 = \{\neg Bp, Bq, q\}$.
- $\Gamma_5 = (AF, \{Bp, Bq, B(p \to q)\})$ has the grounded belief extension $E_5 = \varnothing$; four stable (or preferred) belief extensions $E_6 = \{p, Bp, \neg Bq\}$, $E_7 = \{p, Bp, Bq\}$, $E_8 = \{q, \neg Bp, Bq\}$, and $E_9 = \{\neg Bp, \neg Bq\}$; and five complete belief extensions $E_5, E_6, E_7, E_8, E_9$.



# Properties (1)

- Since $co$, $pr$, $gr$ are universal, $\Gamma = (AF, S)$ has a $\sigma$ belief extension if $AF = (A, R)$ has a $\sigma$ extension for $\sigma \in \{co, pr, gr\}$.
- When $AF = (A, R)$ has a stable extension, $\Gamma = (AF, S)$ may not have a stable extension; and when $AF = (A, R)$ has no stable extension, $\Gamma = (AF, S)$ may have a stable belief extension.

**Example**

(1) $AF = (\{p, q\}, \{(p, q), (q, q)\})$ has the stable extension $\{p\}$, while $AFB = (AF, \{\neg Bp\})$ has no stable belief extension.
(2) $AF = (\{p\}, \{(p, p)\})$ has no stable extension, while $AFB = (AF, \{\neg Bp\})$ has the stable belief extension $\{\neg Bp\}$.

# Properties (2)

An AFB $\Gamma = (A, R, S)$ is rational if $cl_{AT}(S)$ is consistent, i.e., a rational AFB represents an agent who has a consistent belief over AF.

Let $\Gamma = (A, R, S)$ be a rational AFB and $\sigma \in \{co, st, pr, gr\}$.

- $cl_{AT}(S)_A \subseteq E$ holds for any $\sigma$ belief extension $E$ of $\Gamma$, where $cl_{AT}(S)_A = cl_{AT}(S) \cap \{ Bp, \neg Bp \mid p \in A \}$,
- If $B(p \leftrightarrow q)$ is in $cl_{AT}(S)$, there is no $\sigma$ belief extension $E$ such that $\{Bp, Bq\} \subseteq E$.
- If $B(p \rightarrow p)$ is in $cl_{AT}(S)$, there is no $\sigma$ belief extension $E$ such that $Bp \in E$.

# Dialogue

Consider dialogues between two agents $a$ and $b$. Belief of each agent is represented by $B_a$ and $B_b$, respectively.
An argument $p$ made by an agent $a$ is represented by $p_a$.

**dialogue**

A dialogue between two agents $a$ and $b$ is defined as a pair $\Delta = (\Gamma_a, \Gamma_b)$ where $\Gamma_a = (AF, S_a)$ and $\Gamma_b = (AF, S_b)$ are AFBs.

**(in)sincere agent**

Let $\Gamma_a = (AF, S_a)$ be an AFB with $AF = (A, R)$.
The agent $a$ is sincere if $p_a \in A$ implies $B_a p_a \in S_a$; otherwise, $a$ is insincere.

A sincere agent makes an argument only if she believes it.

# Static belief extension

Given $AF = (A, R)$, attacks over beliefs and the attack axiom are modified as:

$$R_B = R \cup \{(\neg B_i p_j, p_j), (\neg B_i p_j, B_i p_j), (B_i p_j, \neg B_i p_j)\},$$

$$\textbf{(AT)} \quad B_i p_j \wedge B_i(p_j \rightarrow q_k) \supset \neg B_i q_k$$

where $p_j, q_k \in A$ and $i, j, k \in \{a, b\}$.

## static belief extension

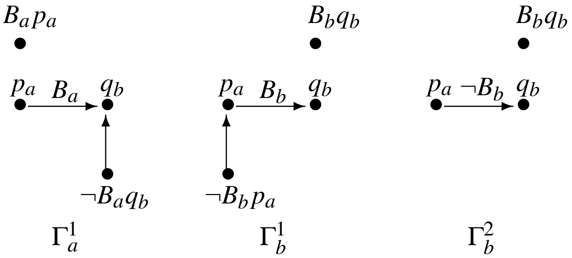
Let $\Delta = (\Gamma_a, \Gamma_b)$ be a dialogue where $\Gamma_a = (AF, S_a)$ and $\Gamma_b = (AF, S_b)$. A pair $(E, F)$ is a static $\sigma$ belief extension (or $\sigma$-SBE for short) of $\Delta$ if

- $E$ is a $\sigma$ extension of $AF = (X, Y)$ where
  $X = A \cup cl_{AT}(S_a)_A$;
  $Y = ((X \times X) \cap R_B) \setminus \{(p \rightarrow q) \mid \neg B_a(p \rightarrow q) \in cl_{AT}(S_a)_R\}$.
- $F$ is a $\sigma$ extension of $AF = (X, Y)$ where
  $X = A \cup cl_{AT}(S_b)_A$;
  $Y = ((X \times X) \cap R_B) \setminus \{(p \rightarrow q) \mid \neg B_b(p \rightarrow q) \in cl_{AT}(S_b)_R\}$.

# Example (1)

Let $AF = (\{p_a, q_b\}, \{(p_a, q_b)\})$ and $\sigma \in \{co, st, pr, gr\}$.

- $\Delta_1 = (\Gamma_a^1, \Gamma_b^1)$ where $\Gamma_a^1 = (AF, \{B_a(p_a \to q_b), B_a p_a\})$ and $\Gamma_b^1 = (AF, \{B_b(p_a \to q_b), B_b q_b\})$ has the $\sigma$-SBE $(\{p_a, B_a p_a, \neg B_a q_b\}, \{q_b, B_b q_b, \neg B_b p_a\})$.
- $\Delta_2 = (\Gamma_a^1, \Gamma_b^2)$ where $\Gamma_b^2 = (AF, \{\neg B_b(p_a \to q_b), B_b q_b\})$ has the $\sigma$-SBE $(\{p_a, B_a p_a, \neg B_a q_b\}, \{p_a, q_b, B_b q_b\})$.



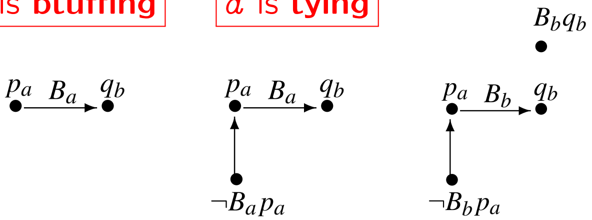
# Example (2)

Suppose that the agent $a$ is **insincere**.

- $\Delta_3 = (\Gamma_a^3, \Gamma_b^1)$ where $\Gamma_a^3 = (AF, \{B_a(p_a \to q_b)\})$ and $\Gamma_b^1 = (AF, \{B_b(p_a \to q_b), B_b q_b\})$ has the $\sigma$-SBE $(\{p_a\}, \{q_b, B_b q_b, \neg B_b p_a\})$.
- $\Delta_4 = (\Gamma_a^4, \Gamma_b^1)$ where $\Gamma_a^4 = (AF, \{\neg B_a p_a, B_a(p_a \to q_b)\})$ has the $\sigma$-SBE $(\{\neg B_a p_a, q_b\}, \{q_b, B_b q_b, \neg B_b p_a\})$.

$a$ is **bluffing**

$a$ is **lying**




$\Gamma_a^3$ $\Gamma_a^4$ $\Gamma_b^1$

# Characterizing dynamic aspect

$B_a^t p$ (resp. $B_a^t(p \rightarrow q)$) means that $a$ believes $p$ (resp. $p \rightarrow q$) at time $t$ where $t \geq 0$ is an integer representing discrete time steps. Let $T$ be a set of integers.

**belief change axiom**

$$\textbf{(BC)} \qquad B_a^t p \wedge B_a^t(p \rightarrow q) \supset \neg B_a^{t+1} q \quad (t \in T)$$

**inertia rule**

$$\textbf{(IR)} \qquad \frac{B_a^t \alpha : B_a^{t+1} \alpha}{B_a^{t+1} \alpha} \quad \text{and} \quad \frac{\neg B_a^t \alpha : \neg B_a^{t+1} \alpha}{\neg B_a^{t+1} \alpha} \quad (t \in T)$$

where $\alpha$ is either an argument $p$ or an attack $p \rightarrow q$.

(**IR**) are normal default rules in **default logic** meaning that if $(\neg)B_a^t \alpha$ is the case and $(\neg)B_a^{t+1} \alpha$ is consistently assumed then conclude $(\neg)B_a^{t+1} \alpha$.

# Closure

Given $AF = (A, R)$, define

$$\mathcal{B}^T_{AF} = \{\, B^t_i p,\ \neg B^t_i p \mid p \in A \text{ and } t \in T \,\} \cup \{\, B^t_i(p \to q),\ \neg\, B^t_i(p \to q) \mid (p, q) \in R \text{ and } t \in T \}.$$

## $cl_D(S)$

Given $S \subseteq \mathcal{B}^T_{AF}$, define $cl_D(S) \subseteq \mathcal{B}^T_{AF}$ as the smallest set of belief atoms satisfying the following conditions:

1. $S \subseteq cl_D(S)$.
2. If $B^t_a p \in cl_D(S)$ and $B^t_a(p \to q) \in cl_D(S)$ then $\neg B^{t+1}_a q \in cl_D(S)$.
3. If $B^{t+1}_a q \in cl_D(S)$ and $B^t_a(p \to q) \in cl_D(S)$ then $\neg B^t_a p \in cl_D(S)$.
4. If $B^t_a p \in cl_D(S)$ and $B^{t+1}_a q \in cl_D(S)$ then $\neg B^t_a(p \to q) \in cl_D(S)$.
5. If $B^t_a \alpha \in cl_D(S)$ and $\{B^{t+1}_a \alpha\} \cup cl_D(S)$ is consistent, then $B^{t+1}_a \alpha \in cl_D(S)$.
6. If $\neg B^t_a \alpha \in cl_D(S)$ and $\{\neg B^{t+1}_a \alpha\} \cup cl_D(S)$ is consistent, then $\neg B^{t+1}_a \alpha \in cl_D(S)$.

# Dynamic belief extension

Given $AF = (A, R)$, define

$$R_D = R \cup \{ (\neg B_i^t p_j, p_j), (\neg B_i^t p_j, B_i^t p_j), (B_i^t p_j, \neg B_i^t p_j) \mid p_j \in A,\ i, j \in \{a, b\}, \text{ and } t \in T\},$$

$$cl_D(S)_A = cl_D(S) \cap \{ B_i^t p, \neg B_i^t p \mid p \in A, i \in \{a, b\}, t \in T \},$$

$$cl_D(S)_R = cl_D(S) \cap \{ B_i^t (p \to q), \neg B_i^t (p \to q) \mid (p \to q) \in R, i \in \{a, b\}, t \in T \}.$$

## dynamic belief extension

Let $\Delta = (\Gamma_a, \Gamma_b)$ be a dialogue where $\Gamma_a = (AF, S_a)$ and $\Gamma_b = (AF, S_b)$. A pair $(E, F)$ is a dynamic $\sigma$ belief extension (or $\sigma$-DBE for short) of $\Delta$ if

- $E$ is a $\sigma$ extension of $AF = (X, Y)$ where
 $X = A \cup cl_D(S_a)_A$;
 $Y = ((X \times X) \cap R_D) \setminus \{(p \to q) \mid \neg B_a^t (p \to q) \in cl_D(S_a)_R\}$.
- $F$ is a $\sigma$ extension of $AF = (X, Y)$ where
 $X = A \cup cl_D(S_b)_A$;
 $Y = ((X \times X) \cap R_D) \setminus \{(p \to q) \mid \neg B_b^t (p \to q) \in cl_D(S_b)_R\}$.

# Example (1)

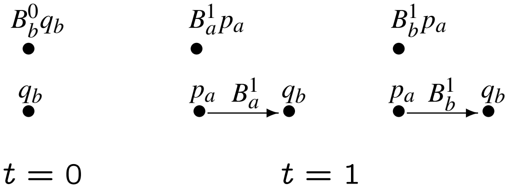
Consider a dialogue $\Delta = (\Gamma_a, \Gamma_b)$ with

$\Gamma_a = (AF, \{B_a^1(p_a \rightarrow q_b), B_a^1 p_a\})$ and

$\Gamma_b = (AF, \{B_b^1(p_a \rightarrow q_b), B_b^0 q_b, B_b^1 p_a\})$

where $AF = (\{p_a, q_b\}, \{(p_a, q_b)\})$.

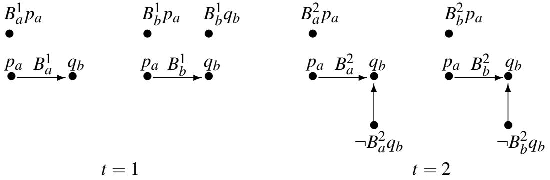
- At $t = 0$, $b$ makes an argument $q_b$ and she believes it.
- At $t = 1$, $a$ makes a counter-argument $p_a$ with the attack $p_a \rightarrow q_b$, and he believes them.
- At $t = 1$, $b$ also believes the argument $p_a$ and the attack $p_a \rightarrow q_b$.



# Example (2)

1. $B_a^1 p_a$ and $B_a^1(p_a \rightarrow q_b)$ imply $\neg B_a^2 q_b$ by (**BC**).
2. $B_b^1 p_a$ and $B_b^1(p_a \rightarrow q_b)$ imply $\neg B_b^2 q_b$ by (**BC**).
3. $B_b^0 q_b$ implies $B_b^1 q_b$ by (**IR**).
4. $B_a^1 p_a$ and $B_b^1 p_a$ respectively imply $B_a^2 p_a$ and $B_b^2 p_a$ by (**IR**).
5. $B_a^1(p_a \rightarrow q_b)$ and $B_b^1(p_a \rightarrow q_b)$ respectively imply $B_a^2(p_a \rightarrow q_b)$ and $B_b^2(p_a \rightarrow q_b)$ by (**IR**).
6. $B_b^1 q_b$ does **not** imply $B_b^2 q_b$ by (**IR**) and (2).

As a result, $\Delta$ has the $\sigma$-DBE $(E, F)$ such that
$E = \{p_a, B_a^1 p_a, B_a^2 p_a, \neg B_a^2 q_b\}$ and
$F = \{p_a, B_b^0 q_b, B_b^1 q_b, B_b^1 p_a, B_b^2 p_a, \neg B_b^2 q_b\}$.



# Deception

Consider a dialogue $\Delta = (\Gamma_a, \Gamma_b)$ with

$\Gamma_a = (AF, \{B_a^1(p_a \to q_b), \neg B_a^1 p_a\})$ and

$\Gamma_b = (AF, \{B_b^1(p_a \to q_b), B_b^0 q_b, B_b^1 p_a\})$.

The belief state of each agent is computed as follows.

1. $B_b^1 p_a$ and $B_b^1(p_a \to q_b)$ imply $\neg B_b^2 q_b$ by (**BC**).
2. $B_b^0 q_b$ implies $B_b^1 q_b$ by (**IR**).
3. $\neg B_a^1 p_a$ and $B_b^1 p_a$ respectively imply $\neg B_a^2 p_a$ and $B_b^2 p_a$ by (**IR**).
4. $B_a^1(p_a \to q_b)$ and $B_b^1(p_a \to q_b)$ respectively imply $B_a^2(p_a \to q_b)$ and $B_b^2(p_a \to q_b)$ by (**IR**).
5. $B_b^1 q_b$ does **not** imply $B_b^2 q_b$ by (**IR**) and (1).

As a result, $\Delta$ has the $\sigma$-DBE $(E, F)$ such that

$E = \{\neg B_a^1 p_a, \neg B_a^2 p_a, q_b\}$ and

$F = \{p_a, B_b^0 q_b, B_b^1 q_b, B_b^1 p_a, B_b^2 p_a, \neg B_b^2 q_b\}$.

As a result, $b$ accepts the argument $p_a$ and $a$ successfully deceives $b$ by lying.

# Final remarks

- The **AFB** is used for representing **belief states of players** and the **audience** of argumentation.
- In two-persons dialogue, AFB can distinguish belief states of **(in)sincere** players. Belief change of a player is represented by **dynamic belief extensions** that can also model **deceptive dialogues**.
- **Inner conflicts** of an agent are expressed using **nested beliefs**, and **self-deception** is realized by belief extensions of **AF with nested belief (AFNB)**.
- An interesting research issue is to represent and reason about argument and belief using **structured argumentation**.