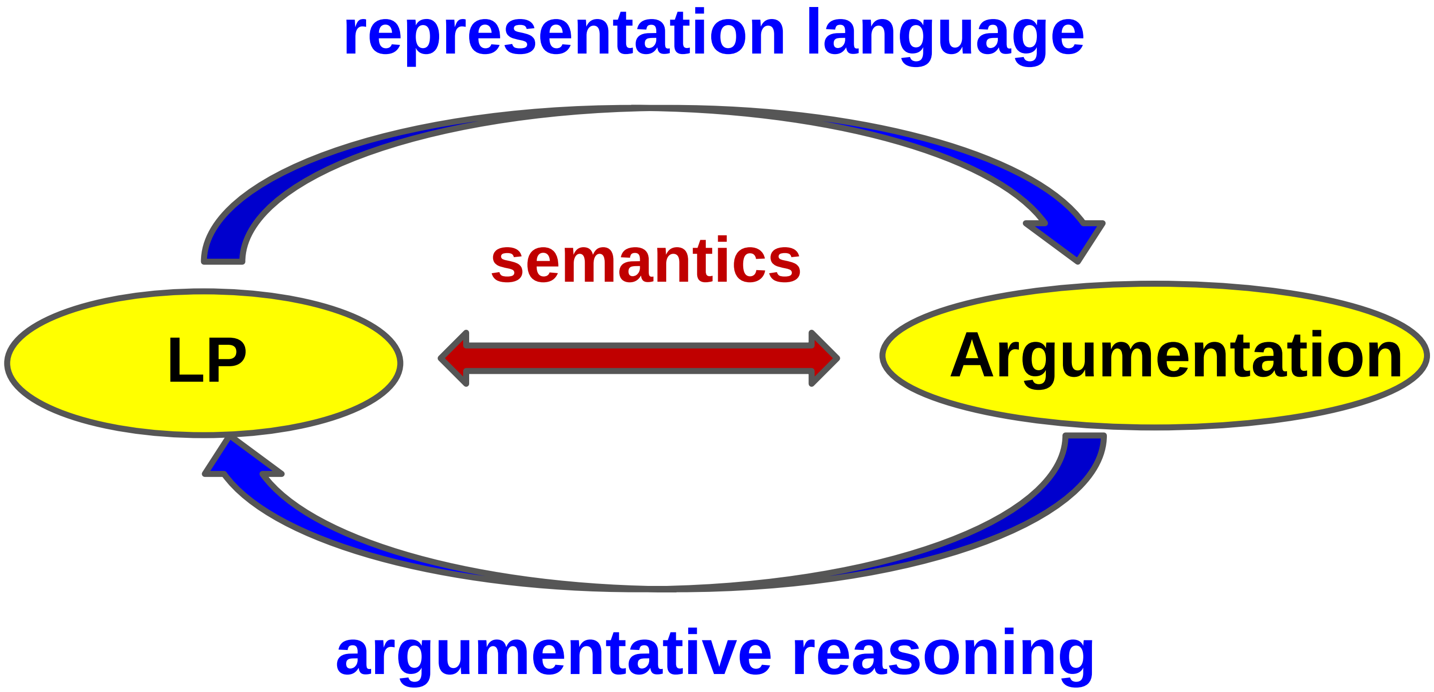
DEBATE GAMES IN LOGIC PROGRAMMING

CHIAKI SAKAMA

WAKAYAMA UNIVERSITY, JAPAN

WFLP / WLP 2013, Kiel, Germany, September 2013

LOGIC PROGRAMMING AND ARGUMENTATION



(formal dialogue, legal reasoning, etc)

DEBATE GAME [SAKAMA, COMMA 2012]

- A **debate game** provides an abstract model of debates between two players.
  - each player has background knowledge as an **argumentation framework**
  - each player revises its argumentation framework by new arguments provided by the opponent player
  - a player may claim **inaccurate** or even **false** arguments as a tactic to win a debate.
- We realize debate games in **logic programming**.

EXTENDED LOGIC PROGRAM

- A **program** consists of rules of the form:

  $L_0 \leftarrow L_1, \ldots, L_m, not\ L_{m+1}, \ldots, not\ L_n$

  where each L_i is called an **objective literal**, and $not\ L_i$ is called a **default literal**.

- For a rule *r* of the above form, $head(r)=L_0$ and $body(r)=\{ L_1, \ldots, L_m, not\ L_{m+1}, \ldots, not\ L_n \}$

- A program is **consistent** if it has a (consistent) **answer set**.

ARGUMENTS ASSOCIATED WITH LP

- An **argument associated with a program** *P* is a finite sequence $A=[\ r_1\ ;\ \ldots\ ;\ r_n\ ]$ of rules $r_i \in P$ **(1≦*i*≦n)** s.t.
  - $\forall i\ \forall L \in body(r_i)\ \exists\ \mathrm{r_k}$ (k>i) in *A* s.t. $head(r_k)=L$.
  - $r_i \neq r_j$ implies $head(r_i) \neq head(r_j)$
- For r_i in *A*, $head(r_i)$ is called a **conclusion**, and *not L* in $body(r_i)$ is called an **assumption**.
- An argument *A* with a conclusion *L* is a **minimal** argument for *L* if there is no subargument (i.e., subsequence of A which is an argument) of *A* with the conclusion *L*.
- An argument is **minimal** if it is minimal for some *L*.

EXAMPLE

P: $p \leftarrow q$
$\neg p \leftarrow not\ q$
$q \leftarrow$
$r \leftarrow s$

- The minimal argument for p is $A_1 = [\ p \leftarrow q\ ;\ q \leftarrow\ ]$.
- The minimal argument for $\neg p$ is $A_2 = [\ \neg p \leftarrow not\ q\ ]$.
- The minimal argument for q is $A_3 = [\ q \leftarrow\ ]$.
- r and s have no minimal argument.

UNDERCUT, REBUT, ATTACK, DEFEAT

Let A_1 and A_2 be two arguments.

- A_1 **undercuts** A_2 if there is an objective literal ***L*** such that ***L*** is a conclusion of A_1 and ***not L*** is an assumption of A_2.
- A_1 **rebuts** A_2 if there is an objective literal ***L*** such that ***L*** is a conclusion of A_1 and $\neg$***L*** is a conclusion of A_2.
- A_1 **attacks** A_2 if A_1 undercuts or rebuts A_2.
- A_1 **defeats** A_2 if A_1 undercuts A_2,
  or A_1 rebuts A_2 and A_2 does not undercut A_1.
- An argument is **coherent** if it does not attack itself.
- Given a program ***P***, the set of minimal and coherent arguments associated with ***P*** is written as *Args*(*P*).

KNOWLEDGE BASE AND REVISION

- A player has a knowledge base $K=(P, O)$ where P is a consistent program representing the player's belief and O is a set of rules brought by another player.
A player is identified with its knowledge base.

- Let $K=(P,O)$ be a player and A an argument.
The **revision** of K with A is defined as:

$$rev(K, A) = (P \setminus R, O \cup A)$$

where $R=\{ r \mid A$ undercuts $[r]$ and A is not defeated by any argument associated with $P \cup O \cup A \}$.

- The result of i-th revision of K is written as $K^i=(P^i, O^i)$ $(i \geqq 0)$ where $K^0=(P, \{ \})$.

EXAMPLE

Let K^0=(**P**, { }) with

P: $p \leftarrow not\ q$

$\neg t \leftarrow not\ s$

$r \leftarrow$

- Given A_1 =[$q \leftarrow not\ r$], K^1=(**P**, { $q \leftarrow not\ r$ }) because A_1 undercuts [$p \leftarrow not\ q$] but [$r \leftarrow$] defeats A_1 .
- Given A_2 =[$s \leftarrow$], K^2=(**P** \ { $\neg t \leftarrow not\ s$ }, { $q \leftarrow not\ r$, $s \leftarrow$ }) because A_2 undercuts [$\neg t \leftarrow not\ s$] and A_2 is not defeated by any argument associated with **P** $\cup$ { $q \leftarrow not\ r$, $s \leftarrow$ }.

DEBATE GAME

- Let $K_1=(P_1,O_1)$ and $K_2=(P_2,O_2)$ be two players.
  - The **initial claim** by K_1 is: (in(X), _) where $X \in Args(P_1)$
    "the player K_1 claims the argument X"
  - A **counter-claim** by K_h is: (out(X), in(Y)) where
    $X \in Args(P_k \cup O_k)$ and $Y \in Args(P_h \cup O_h)$ ($k, h = 1,2;\ k \neq h$)
    "the argument X by the player k does not hold because the player h claims the argument Y".
- A **debate game** between two players is a sequence of claims: [(in(X_0), _), (out(X_0),in(Y_1)), (out(Y_1),in(X_1)),…] where
  - $X_i \in Args(P_1^i \cup O_1^i)$ and $Y_j \in Args(P_2^j \cup O_2^j)$ ($i, j \geqq 0$)
  - for each (out(U), in(V)), V defeats U.

EXAMPLE

A prosecutor has a knowledge base $K_P=(P_P, O_P)$ where

$P_P = \{\ guilty \leftarrow suspect, motive,$ $\quad O_P = \{\ \}$

$evidence \leftarrow witness, not\ \neg credible,$

$suspect \leftarrow, \quad motive \leftarrow, \quad witness \leftarrow\ \}$

A defense has a knowledge base $K_D=(P_D, O_D)$ where

$P_D = \{\ \neg guilty \leftarrow suspect, not\ evidence,$ $\quad O_D = \{\ \}$

$\neg credible \leftarrow witness, dark,$

$suspect \leftarrow, \quad dark \leftarrow\ \}$

P_P

guilty ← *suspect, motive*
evidence ← *witness, not* ¬*credible*
suspect ← *motive* ← *witness* ←

O_P

A debate game proceeds between **P**rosecutor and **D**efense:

P: (in(X), __) with X =[*guilty* ← *suspect, motive* ; *suspect* ←; *motive* ←]
(``The suspect is guilty because he has a motive for the crime.")

P_D^1

¬ guilty ← suspect, not evidence *¬ credible ← witness, dark* *suspect ←* *dark ←*

O_D^1

guilty ← suspect, motive *motive ←*

A debate game proceeds between **P**rosecutor and **D**efense:

P: (in(X), __) with X =[*guilty ← suspect, motive* ; *suspect ←*; *motive ←*]
(``The suspect is guilty because he has a motive for the crime.")

D: (out(X), in(Y)) with Y = [*¬ guilty ← suspect, not evidence* ; *suspect ←*]
(``The suspect is not guilty as there is no evidence.")

P_P^1

guilty ← *suspect, motive*
evidence ← *witness, not* ¬ *credible*
suspect ← *motive* ← *witness* ←

O_P^1

¬*guilty* ←
suspect, not evidence

A debate game proceeds between **P**rosecutor and **D**efense:

P: (in(X), __) with X =[*guilty* ← *suspect, motive* ; *suspect* ←; *motive* ←]
(``The suspect is guilty because he has a motive for the crime.")

D: (out(X), in(Y)) with Y = [¬ *guilty* ← *suspect, not evidence* ; *suspect* ←]
(``The suspect is not guilty as there is no evidence.")

P: (out(Y), in(Z)) with Z =[*evidence* ← *witness, not* ¬ *credible*;
witness ←]
(``There is an eyewitness who saw the suspect on the night of the crime.")

$P_D{}^2$

$\neg$ *guilty* ← *suspect, not evidence*
$\neg$ *credible* ← *witness, dark*
suspect ← *dark* ←

$O_D{}^2$

guilty ← *suspect, motive*
evidence ← *witness, not* $\neg$ *credible*
motive ← *witness* ←

A debate game proceeds between **P**rosecutor and **D**efense:

P: (in(X), __) with X =[*guilty* ← *suspect, motive* ; *suspect* ←; *motive* ←]
(``The suspect is guilty because he has a motive for the crime.")

D: (out(X), in(Y)) with Y = [$\neg$ *guilty* ← *suspect, not evidence* ; *suspect* ←]
(``The suspect is not guilty as there is no evidence.")

P: (out(Y), in(Z)) with Z =[*evidence* ← *witness, not* $\neg$ *credible*; *witness* ←]
(``There is an eyewitness who saw the suspect on the night of the crime.")

D: (out(Z), in(W)) with W = [$\neg$ *credible* ← *witness, dark* ; *witness* ← ; *dark* ←] (``The testimony is incredible because it was dark at night.")

P_P^2

guilty ← *suspect, motive*
evidence ← *witness, not* ¬*credible*
suspect ←   *motive* ←   *witness* ←

O_P^2

¬*guilty* ← *suspect, not evidence*
¬*credible* ← *witness, dark*
dark ←

A debate game proceeds between **P**rosecutor and **D**efense:

P: (in(X), __) with X =[*guilty* ← *suspect, motive* ; *suspect* ←; *motive* ←] (``The suspect is guilty because he has a motive for the crime.")

D: (out(X), in(Y)) with Y = [¬*guilty* ← *suspect, not evidence* ; *suspect* ←] (``The suspect is not guilty as there is no evidence.")

P: (out(Y), in(Z)) with Z =[*evidence* ← *witness, not* ¬*credible*; *witness* ←] (``There is an eyewitness who saw the suspect on the night of the crime.")

D: (out(Z), in(W)) with W = [¬*credible* ← *witness, dark* ; *witness* ← ; *dark* ←] (``The testimony is incredible because it was dark at night.")

The prosecutor cannot make a counter-claim and the defense wins the game.

P_P^2

guilty ← suspect, motive
evidence ← witness, not ¬ credible
suspect ← *motive ←* *witness ←*
¬ dark ← light, not broken
light ← ***broken ←***

O_P^2

¬guilty ← suspect, not evidence
¬ credible ← witness, dark
dark ←

A debate game proceeds between **P**rosecutor and **D**efense:

P: (in(X), __) with X =[*guilty ← suspect, motive* ; *suspect ←*; *motive ←*] (``The suspect is guilty because he has a motive for the crime.")

D: (out(X), in(Y)) with Y = [*¬ guilty ← suspect, not evidence* ; *suspect ←*] (``The suspect is not guilty as there is no evidence.")

P: (out(Y), in(Z)) with Z =[*evidence ← witness, not ¬ credible*; *witness ←*] (``There is an eyewitness who saw the suspect on the night of the crime.")

D: (out(Z), in(W)) with W = [*¬ credible ← witness, dark* ; *witness ←* ; *dark ←*] (``The testimony is incredible because it was dark at night.")

P: (out(W), in(V)) with V =[***¬ dark ← light, not broken*** ; ***light←***] (`` It was not dark because the witness saw the suspect under the light of the victim's apartment.")

$P_P{}^2$

guilty ← suspect, motive
evidence ← witness, not ¬ credible
suspect ← *motive ←* *witness ←*
¬ dark ← light, not broken
light← ***broken←***

$O_P{}^2$

¬guilty ← suspect, not evidence
¬ credible ← witness, dark
dark ←

P: (out(W), in(V)) with V =[***¬ dark ← light, not broken*** ; ***light←***]
(`` It was not dark because the witness saw the suspect under the light of the victim's apartment.")

- The prosecutor claims the argument *V* but he/she does not believe its conclusion ***¬ dark***.
- In fact, ***¬ dark*** is included in no answer set of the program $\mathbf{P_P}^2 \cup \mathbf{Q}$ for any $\mathbf{Q} \subseteq \mathbf{O_P}^2$.

DISHONEST CLAIMS

Let Γ be a claim of either (in(U), _) or (out(V), in(U)) by a player *K*=(*P*,*O*).
Let U^S be an argument which consists of rules in the reduct of U wrt a set S.
The set of conclusions of U is written as concl(U).

- Γ is **credible** if concl(U)⊆S for every answer set S of P∪Q for some Q⊆O such that P∪Q is consistent and concl(U)=concl(U^S).
- Γ is **misleading** if concl(U)⊆S for every answer set S of P∪Q for some Q⊆O such that P∪Q is consistent and concl(U)≠concl(U^S).
- Γ is **incredible** if concl(U)⊆S for some (but not every) answer set S of P∪Q for any Q⊆O such that P∪Q is consistent.
- Γ is **incorrect** if concl(U)⊈S for any answer set S of P∪Q for any Q⊆O such that P∪Q is consistent and concl(U) ∪S is consistent for some answer set S of P∪Q for some Q ⊆O such that P∪Q is consistent.
- Γ is **false** if concl(U) ∪S is inconsistent for any answer set S of P∪Q for any Q⊆O such that P∪Q is consistent.

DISHONEST CLAIMS (EXAMPLE)

- Given K=({ $p \leftarrow not\ q$ }, {}), the claim Γ=(in([$p \leftarrow not\ q$]), _) (``p holds because q does not hold'') is **credible**.
- Given K=({ $p \leftarrow not\ q,\quad p \leftarrow q,\quad q \leftarrow$ }, {}), the claim Γ=(in([$p \leftarrow not\ q$]), _) is **misleading.**
- Given K=({ $p \leftarrow not\ q,\quad q \leftarrow not\ p$ }, {}), the claim Γ=(in([$p \leftarrow not\ q$]), _) is **incredible**.
- Given K=({ $p \leftarrow not\ q,\quad q \leftarrow$ }, {}), the claim Γ=(in([$p \leftarrow not\ q$]), _) is **incorrect**.
- Given K=({ $p \leftarrow not\ \neg q,\quad \neg p \leftarrow$ }, {}), the claim Γ=(in([$p \leftarrow not\ \neg q$]), _) is **false**.
- A player is **honest** in a debate game if every claim made by the player is credible, otherwise, the player is **dishonest**.

PROPERTIES

A player K=(P,O) is **monotonic** if P contains no default literal.

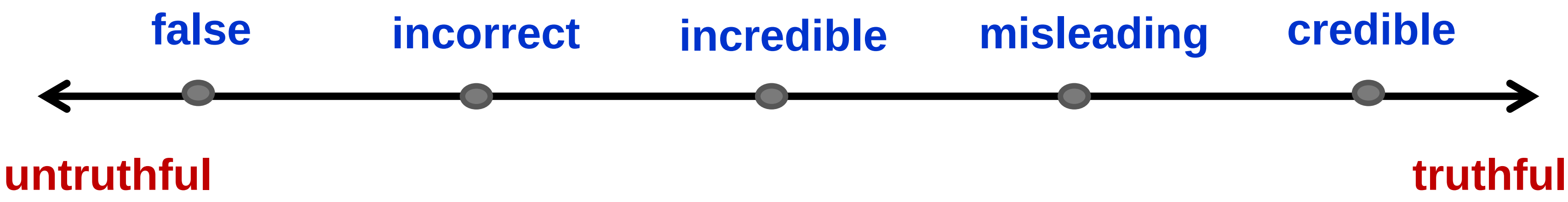
- Let Δ be a debate game between two ***monotonic*** players. Then, every claim in Δ is credible.

  The existence of dishonest claims is due to the nonmonotonic nature of a program.

- Let Δ be a debate game between two players $K_1=(P_1,O_1)$ and $K_2=(P_2,O_2)$. If K_1 (resp. K_2) is honest and $P_2 \subset P_1$ (resp. $P_1 \subset P_2$), then K_1 (resp. K_2) wins the game.

  If a player has information more than another player, he/she has no reason to behave dishonestly to win a debate.

DEGREE OF TRUTHFULNESS



- A player has an incentive to build a **dishonest** claim if he/she cannot build a honest counter-claim against the opponent.
- **Misleading** claims are useless for the purpose of winning a game, because a player can build a **credible** claim with the same conclusion.
- **Incredible** claims are preferred to **incorrect** claims, because a player credulously believes the conclusion of an incredible claim.
- **Incorrect** claims are preferred to **false** claims, because the conclusion of an incorrect claim is consistent with the player’s belief.
- The best-practice strategy: *credible > incredible > incorrect > false*

CONCLUSION

- We developed debate games using a non-abstract argumentation framework associated with LP, which contributes to a step toward integrating LP and formal argumentation.
- We showed an application of dishonest reasoning in argumentation-based LP, which contributes to modelling dishonest arguments of humans in daily life.
- Future work includes implementing a prototype system of debate games associated with LP.