

Propagação de confiança em sistemas multiagentes

Lucas Leite

Federal University of Ceará

lgoncalves@lia.ufc.br

November 19, 2015

Overview

- 1 Introdução
 - Argumentation
- 2 Definições
 - Lógica
 - Confiança
 - Argumentação
- 3 Propostas
 - Tópicos
 - Desconfiança
 - Propagação de confiança
- 4 Conclusão

Introdução

- Tomada de decisão

Introdução

- Tomada de decisão
- Argumentação

Introdução

- Tomada de decisão
- Argumentação
- Confiança

Introdução

- Entendimento e formalização de confiança:
- Raciocínio sobre confiança:

Introdução

- Entendimento e formalização de confiança:
 - “O que significa o agente a confiar em um agente b ?”
- Raciocínio sobre confiança:

Introdução

- Entendimento e formalização de confiança:
 - “O que significa o agente a confiar em um agente b ?”
- Raciocínio sobre confiança:
 - “ a deve confiar ou não em uma afirmação de b ?”
 - Modelos estatísticos
 - Modelos lógicos

Introdução

- Amgoud:
 - Acreditar ou não em uma dada informação.
 - Confiar ou não em uma fonte de informação.
 - Aceitar ou não uma informação ou oferta recebida de uma fonte.

Introdução

- Amgoud:
 - Acreditar ou não em uma dada informação.
 - Confiar ou não em uma fonte de informação.
 - Aceitar ou não uma informação ou oferta recebida de uma fonte.
- Propostas:
 - Tratar tópicos.
 - Introduzir a desconfiança.
 - Propagação de confiança ou de desconfiança.

Sintaxe

Sintaxe

- ATOM: conjunto de proposições atômicas denotadas por p, q, r, \dots
- AGENTE: conjunto não-vazio de agentes denotados por i, j, k, \dots
- $\phi ::= p \mid \neg\phi \mid \phi \vee \phi \mid Bel_i\phi \mid Inf_{j,i}\phi$

Axiomas

Axiomas - Crenças

- (K) $Bel_i(\phi \rightarrow \psi) \rightarrow (Bel_i\phi \rightarrow Bel_i\psi)$
- (D) $\neg(Bel_i\phi \wedge Bel_i\neg\phi)$
- (Nec) Se $\vdash \phi$, então $\vdash Bel_i\phi$

Axiomas

Axiomas - Informação

- (EQV) Se $\vdash \phi \leftrightarrow \psi$, então $\vdash \text{Inf}_{j,i}\phi \leftrightarrow \text{Inf}_{j,i}\psi$
- (CONJ) $\text{Inf}_{j,i}\phi \wedge \text{Inf}_{j,i}\psi \rightarrow \text{Inf}_{j,i}(\phi \wedge \psi)$
- (OBS) $\text{Inf}_{j,i}\phi \rightarrow \text{Bel}_i\text{Inf}_{j,i}\phi$
- (OBS') $\neg\text{Inf}_{j,i}\phi \rightarrow \text{Bel}_i\neg\text{Inf}_{j,i}\phi$

Confiança

Sinceridade

$$\mathit{ConfSinc}(i, j, \phi) ::= \mathit{Bel}_i(\mathit{Inf}_{j,i}\phi \rightarrow \mathit{Bel}_j\phi)$$

Validade

$$\mathit{ConfVal}(i, j, \phi) ::= \mathit{Bel}_i(\mathit{Inf}_{j,i}\phi \rightarrow \phi)$$

Completude

$$\mathit{ConfComple}(i, j, \phi) ::= \mathit{Bel}_i(\phi \rightarrow \mathit{Inf}_{j,i}\phi)$$

Confiança

Cooperatividade

$$\mathit{ConfCoop}(i, j, \phi) ::= \mathit{Bel}_i(\mathit{Bel}_j\phi \rightarrow \mathit{Inf}_{j,i}\phi)$$

Competência

$$\mathit{ConfCompet}(i, j, \phi) ::= \mathit{Bel}_i(\mathit{Bel}_j\phi \rightarrow \phi)$$

Vigilância

$$\mathit{ConfVig}(i, j, \phi) ::= \mathit{Bel}_i(\phi \rightarrow \mathit{Bel}_j\phi)$$

Confiança

Relação entre propriedades

- (V)
 $\vdash \text{ConfSinc}(i, j, \phi) \wedge \text{ConfCompet}(i, j, \phi) \rightarrow \text{ConfVal}(i, j, \phi)$
- (C)
 $\vdash \text{ConfVig}(i, j, \phi) \wedge \text{ConfCoop}(i, j, \phi) \rightarrow \text{ConfComple}(i, j, \phi)$

Confiança

Impacto da informação

- (E1) $\vdash \text{ConfSinc}(i, j, \phi) \rightarrow (\text{Inf}_{j,i}\phi \rightarrow \text{Bel}_i\text{Bel}_j\phi)$
- (E2) $\vdash \text{ConfVal}(i, j, \phi) \rightarrow (\text{Inf}_{j,i}\phi \rightarrow \text{Bel}_i\phi)$
- (E3) $\vdash \text{ConfCoop}(i, j, \phi) \rightarrow (\neg\text{Inf}_{j,i}\phi \rightarrow \text{Bel}_i\neg\text{Bel}_j\phi)$
- (E4) $\vdash \text{ConfComple}(i, j, \phi) \rightarrow (\neg\text{Inf}_{j,i}\phi \rightarrow \text{Bel}_i\neg\phi)$

Argumentos

Argumentos

- $\langle H, h \rangle$
 - $K \cup H \vdash h$;
 - $K \cup H$ é consistente;
 - $\nexists H' \subset H$ tal que $K \cup H' \vdash h$

Framework de Argumentação

Framework de Argumentação

- $AF = \langle \mathcal{A}, Def \rangle$
 - \mathcal{A} é um conjunto de argumentos
 - $Def \subseteq \mathcal{A} \times \mathcal{A}$

Relação de ataque

Ataque sobre premissa

- Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos. Dizemos que A realiza um ataque sobre premissa em B se e somente se $\exists h'' \in H'$ tal que $h = Bel_i \phi$ e $h'' = Bel_i \neg \phi$.

Ataque sobre sinceridade

- Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos. A realiza um ataque sobre sinceridade em B se e somente se $h = Bel_i (Inf_{j,i} \phi \wedge \neg Bel_j \phi)$ e $ConfSinc(i, j, \psi) \in H'$.

Ataque sobre validade

- Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos. A realiza um ataque sobre validade em B se e somente se $h = Bel_i (Inf_{j,i} \phi \wedge \neg \phi)$ e $ConfVal(i, j, \psi) \in H'$.

Relação de ataque

Ataque sobre completude

- Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos. A realiza um ataque sobre completude em B se e somente se $h = Bel_i(\phi \wedge \neg Inf_{j,i}\phi)$ e $ConfComple(i, j, \psi) \in H'$.

Ataque sobre cooperatividade

- Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos. A realiza um ataque sobre cooperatividade em B se e somente se $h = Bel_i(Bel_j\phi \wedge \neg Inf_{j,i}\phi)$ e $ConfCoop(i, j, \psi) \in H'$.

Relação de ataque

Ataque sobre competência

- Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos. A realiza um ataque sobre competência em B se e somente se $h = Bel_i(Bel_j\phi \wedge \neg\phi)$ e $ConfCompet(i, j, \psi) \in H'$.

Ataque sobre vigilância

- Sejam $A = (H, h)$ e $B = (H', h')$ dois argumentos. A realiza um ataque sobre vigilância em B se e somente se $h = Bel_i(\phi \wedge \neg Bel_j\phi)$ e $ConfVig(i, j, \psi) \in H'$.

Semântica

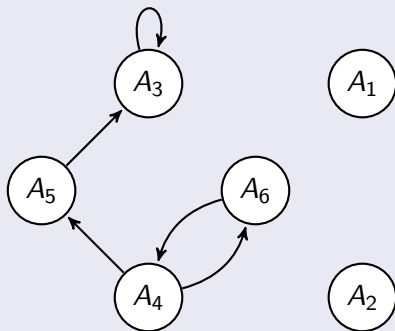
- Ponto de vista ou decisão
- Dados $AF = \langle \mathcal{A}, Def \rangle$ e $S \subseteq \mathcal{A}$:
 - $\forall a, b \in S, (a, b) \notin Def$;
 - $\forall a \in S$ e $\forall b \in \mathcal{A} \setminus S$ tal que $(b, a) \in Def$, $\exists c \in S$ tal que $(c, b) \in Def$.

Extensões

- Extensões admissíveis:
 - Um conjunto $S \subseteq \mathcal{A}$ é uma extensão preferida se e somente se é uma extensão admissível maximal (com respeito à inclusão em conjuntos).
 - Um conjunto $S \subseteq \mathcal{A}$ é uma extensão estável se e somente se é uma extensão preferida de AF que derrota qualquer argumento em $\mathcal{A} \setminus S$.
 - Um conjunto $S \subseteq \mathcal{A}$ é uma extensão completa se e somente se é livre de conflito e $S = \{A \mid A \in \mathcal{A}, S \text{ defende } A\}$.
 - Um conjunto $S \subseteq \mathcal{A}$ é uma extensão grounded se e somente se é a menor (com respeito à inclusão de conjuntos) extensão completa de AF .

Exemplo

Example



- AF tem como sua extensão estável o conjunto $\{A_1, A_2, A_5, A_6\}$

Tópicos

- Conjuntos de fórmulas
- $R(t, \phi)$
- Mudança nas definições de ataque

Desconfiança

- Não é apenas ausência de confiança
- $DesConfSinc(i, j, \phi) ::= Bel_i(Inf_{j,i}\phi \rightarrow \neg Bel_j\phi)$
- Mudança nas definições de ataque

Propagação de confiança

- O que seria necessário para propagar a confiança
- Propagação direta: $\text{Conf}(a,b) \wedge \text{Conf}(b,c) \rightarrow \text{Conf}(a,c)$
- Reciprocidade: $\text{Conf}(a,b) \rightarrow \text{Conf}(b,a)$
- Coimplicação: $\text{Conf}(a,b) \wedge \text{Conf}(c,b) \rightarrow \text{Conf}(a,c)$

Propagação de confiança

- Propagação direta:
 - a sabe que b confia em c ?
 - b é confiável em relação à recomendação de confiança?
 - a confia em b no mesmo contexto que b confia em c ?
- Reciprocidade:
 - b sabe que a confia nele?
 - b tem alguma razão para confiar em a com respeito a alguma propriedade?
- Coimplicação:
 - a sabe que c confia em b ?
 - a confia em b no mesmo contexto que c confia em b ?

Conclusão

- Confiança na recomendação.
- Confiança na capacidade de manter sigilo.

Obrigado!