Reasoning about Epistemic States of Agents by Modal Logic Programming

Linh Anh Nguyen

Institute of Informatics, University of Warsaw
ul. Banacha 2, 02-097 Warsaw, Poland
nguyen@mimuw.edu.pl

CLIMA VI, 27-06-2005

Motivation

In multi-agent systems, agents should be able to collaborate or compete with each other. For this aim, an agent should have knowledge about other agents in the system and be able to reason about their epistemic states.

It is not that an agent can have all information it wants or can reason exactly as the others, but at least it can simulate epistemic states of the other agents, using some assumptions.

We study the wise men puzzle by modal logic programming (using the direct approach).

Overview

- The modal logic $KD4I_g5_a$ for reasoning about belief and common belief

- The wise men puzzle

- A formalization of the wise men puzzle in $KD4I_g5_a$

- An SLD-resolution calculus for modal logic programming in $KD4I_g5_a$

- An SLD-refutation for the wise men puzzle

- Conclusions

The multimodal logic $KD4I_g5_a$

Let $\Box_i\varphi$ stand for agent i (or group of agents i) believes in φ.

For reasoning about belief and common belief we can adopt the axioms:

- (D) : $\Box_i\varphi \rightarrow \neg\Box_i\neg\varphi$ (consistence),
- (4) : $\Box_i\varphi \rightarrow \Box_i\Box_i\varphi$ (positive introspection),
- (I_g) : $\Box_i\varphi \rightarrow \Box_j\varphi$ if i is a supergroup of j,
- (5_a) : $\neg\Box_i\varphi \rightarrow \Box_i\neg\Box_i\varphi$ if i is a single agent (negative introspection).

The multimodal logic $KD4I_g5_a$ (2)

Let n be the number of single agents.
Then $m = 2^n - 1$ is the number of non-empty groups of agents.
Let g be a bijection that maps $1 \leq i \leq m$ to a non-empty group of agents.

The logic $KD4I_g5_a$ can be defined as

$$KD4I_g5_a = K_{(m)} + (D) + (4) + (I_g) + (5_a)$$

where $K_{(m)}$ is the smallest normal multimodal logic with m operators $\Box_i$.

We use the standard Kripke semantics, with fixed-domain and rigid terms.

The Wise Men Puzzle

The puzzle is a famous benchmark introduced by McCarthy for AI.
It can be stated as follows:

A king wishes to know whether his three advisors (A, B, C) are as wise as they claim to be. Three chairs are lined up, all facing the same direction, with one behind the other. The wise men are instructed to sit down in the order A, B, C. Each of the men can see the backs of the men sitting before them (e.g. C can see A and B). The king informs the wise men that he has three cards, all of which are either black or white, at least one of which is white. He places one card, face up, behind each of the three wise men, explaining that each wise man must determine the color of his own card. Each wise man must announce the color of his own card as soon as he knows what it is. All know that this will happen. The room is silent; then, after a while, wise man A says "My card is white!".

The Wise Men Puzzle (2)

The wise men puzzle has been previously studied in several works.

By modal logics:
- McCarthy (1979) directly used possible worlds.
- Konolige (1984) focused on limited reasoning.
- Nonnengart (1994) used semi-functional translation
  for modal logic programming.
- Baldoni (2000) used a prefixed tableau system.

Other approaches:
- Elgot-Drapkin (1991) used step-logics.
- Cimatti and Serafini (1994), Attardi and Simi (1994)
  studied reasoning in belief-contexts.

A formalization of the wise men puzzle in $KD4I_g5_a$

We formalize the puzzle by a modal logic program in $KD4I_g5_a$. The goal is whether wise man A believes that his card is white: $\leftarrow \Box_a white(a)$.

%If Y sits behinds X then X's card is white if Y considers this as possible.

$$\Box_{abc}(white(a) \leftarrow \Diamond_b white(a))$$
$$\Box_{abc}(white(a) \leftarrow \Diamond_c white(a))$$
$$\Box_{abc}(white(b) \leftarrow \Diamond_c white(b))$$

%If Y sits behinds X and X's card is black then Y know it.

$$\Box_{abc}(\Box_b black(a) \leftarrow black(a))$$
$$\Box_{abc}(\Box_c black(a) \leftarrow black(a))$$
$$\Box_{abc}(\Box_c black(b) \leftarrow black(b))$$

%At least one of the wise men has a white card.

$$\Box_{abc}(white(a) \leftarrow black(b), black(c))$$
$$\Box_{abc}(white(b) \leftarrow black(c), black(a))$$
$$\Box_{abc}(white(c) \leftarrow black(a), black(b))$$

%Each of B and C does not know the color of his own card. In particular,
%each of the men considers that it is possible that his own card is black.

$$\Box_{abc}\Diamond_b black(b)$$
$$\Box_{abc}\Diamond_c black(c)$$

The MProlog Language

We use
- $\triangle$ to denote a modality (a sequence of modal operators),
- $\boxdot$ to denote a universal modality,
- E to denote a classical atom,
- A, B to denote simple atoms of the form E, $\Box_i E$, or $\Diamond_i E$,
- $\boxdot(\varphi \leftarrow \psi_1, \ldots, \psi_n)$ to denote the formula $\forall(\boxdot(\varphi \vee \neg\psi_1 \ldots \vee \neg\psi_n))$.

A program clause is a formula of the form $\boxdot(A \leftarrow B_1, \ldots, B_n)$, where $n \geq 0$. $\boxdot$ is called the *modal context*, A the *head*, and $B_1, \ldots, B_n$ the *body* of the program clause.

An MProlog program is a finite set of program clauses.

An MProlog goal atom is a formula of the form $\boxdot E$ or $\boxdot\Diamond_i E$.

An MProlog goal is a formula written in the clausal form $\leftarrow \alpha_1, \ldots, \alpha_k$, where each α_i is an MProlog goal atom.

The MProlog Language (2)

In $KD4I_g5_a$, if $g(i)$ is a singleton then $\nabla_i\nabla'_i\varphi \equiv \nabla'_i\varphi$ for any modal operators ∇_i and ∇'_i with the same modal index i.

Hence we can assume that modal contexts of program clauses and goal atoms in $KD4I_g5_a$-MProlog do not contain $\nabla_i\nabla'_i$ if $g(i)$ is a singleton.

Let P be an $KD4I_g5_a$-MProlog program and $G = \leftarrow \alpha_1, \ldots, \alpha_k$ be an $KD4I_g5_a$-MProlog goal. An answer θ for $P \cup \{G\}$ is a substitution whose domain is a set of variables of G. We say that θ is a correct answer in $KD4I_g5_a$ for $P \cup \{G\}$ if θ is an answer for $P \cup \{G\}$ and

$$P \models_{KD4I_g5_a} \forall((\alpha_1 \wedge \ldots \wedge \alpha_k)\theta)$$

An SLD-Resolution Calculus for MProlog in $KD4I_g5_a$

We use the direct approach (i.e. without translation to classical logic).

Features:

- labeled existential modal operators
- a normal form for modalities in $KD4I_g5_a$
- rules used as meta-clauses
- an ordering of modal operators

From now on, for abbreviation we use L to denote $KD4I_g5_a$.

Labeled Modal Operators

The forward labeled form of $\Diamond_i E$ is $\langle E \rangle_i E$.

The intuition is that: To satisfy a formula $\Diamond_i E$ at a world w we connect w to a new world u via R_i and add E to the content of u. The world u can be denoted by the composition of w and $\langle E \rangle_i$.

The backward labeled form of $\Diamond_i E$ is $\langle X \rangle_i E$, where X is a fresh variable.

We cannot replace $\leftarrow \Diamond_i(A \wedge B)$ by $\leftarrow \Diamond_i A, \Diamond_i B$.
But we can break $\leftarrow \langle S \rangle_i (A \wedge B)$ into $\leftarrow \langle S \rangle_i A, \langle S \rangle_i B$.

Definitions

- A modality is in labeled form if it does not contain unlabeled existential modal operators. A modality is in *L*-normal form if it does not contain any subsequence of the form $\nabla_i \nabla'_i$ for i being a single agent.

- An atom is a formula of the form $\triangle E$, where $\triangle$ is a modality and E is a classical atom. It is in *L*-normal form if $\triangle$ is in *L*-normal form.

- An atom is in almost *L*-normal labeled form if it is of the form $\triangle A$ with $\triangle$ in *L*-normal labeled form and A being a simple atom (of the form E, $\Box_i E$, $\Diamond_i E$, or $\langle S \rangle_i E$).

Rules as Meta-Clauses

Our SLD-resolution calculus for $L = KD4I_g5_a$ contains the following rules. Both sides of each rule are atoms in almost L-normal labeled form.

rNF_L (reverse normalization) rules:

$\triangle\nabla_i E \leftarrow \triangle\langle X\rangle_i\nabla_i E$ if $g(i)$ is a singleton, ∇_i is of the form $\Box_i$ or $\langle E\rangle_i$, and X is a fresh atom variable

$rSat_L$ (reverse saturation) rules:

(1) $\triangle\Diamond_i E \leftarrow \triangle\langle X\rangle_i E$ for X being a fresh atom variable

(2) $\triangle\nabla_i\alpha \leftarrow \triangle\Box_j\alpha$ if $g(i) \subseteq g(j)$

(3) $\triangle\Diamond_i E \leftarrow \triangle\Diamond_j E$ if $g(i) \supset g(j)$

(4) $\triangle\Box_i\Box_i\alpha \leftarrow \triangle\Box_i\alpha$

(5) $\triangle\Box_i\alpha \leftarrow \triangle\langle X\rangle_i\Box_i\alpha$ if $g(i)$ is a singleton and X is a fresh atom variable

(6) $\triangle\nabla_i\Diamond_i E \leftarrow \triangle\Diamond_i E$ if $g(i)$ is a singleton

(7) $\triangle\Diamond_i E \leftarrow \triangle\langle X\rangle_j\Diamond_i E$ if $g(i) \supseteq g(j)$ and X is a fresh atom variable

Definitions

- To compare modal operators we define $\preceq_L$ to be the least reflexive and transitive relation between modal operators such that $\Diamond_i \preceq_L \langle S \rangle_i \preceq_L \Box_i$ and if $g(i) \subseteq g(j)$ then $\Box_i \preceq_L \Box_j$ and $\Diamond_j \preceq_L \Diamond_i$.

- The above order can be extended for comparing modalities and we can specify when an atom is an *L*-instance of another.

- Let $\boxdot$ and $\boxdot'$ be universal modalities. We say that $\boxdot'$ is an *L*-context instance of $\boxdot$ if $\boxdot\varphi \rightarrow \boxdot'\varphi$ is *L*-valid for any formula φ.

Resolving a Goal with a Program Clause

Let $G = \leftarrow \alpha_1, \ldots, \alpha_i, \ldots, \alpha_k$ be a goal and $\varphi = \boxdot(A \leftarrow B_1, \ldots, B_n)$ a program clause. Then G' is derived from G and φ in L using mgu θ, and called an *L-resolvent* of G and φ, if the following conditions hold:

- $\alpha_i = \triangle' A'$, with $\triangle'$ in L-normal labeled form, is called the *selected atom*.
- $\triangle'$ is an L-instance of $\boxdot'$ which is an L-context instance of $\boxdot$.
- θ is an mgu of A' and the forward labeled form of A.
- G' is the goal $\leftarrow (\alpha_1, \ldots, \alpha_{i-1}, \triangle' B_1, \ldots, \triangle' B_n, \alpha_{i+1}, \ldots, \alpha_k)\theta$.

SLD-Resolution for MProlog

SLD-derivation is defined using two kinds of steps:

- resolving a goal with a program clause,
- resolving a goal with an $rSat_L/rNF_L$ rule.

SLD-refutation and computed answer are defined in the usual way.

Soundness and Completeness of SLD-Resolution

Theorem:

The SLD-resolution calculus given for $KD4I_g5_a$ is sound and complete: Let P be an $KD4I_g5_a$-MProlog program and G an $KD4I_g5_a$-MProlog goal. Then every computed answer in $KD4I_g5_a$ of $P \cup \{G\}$ is a correct answer in $KD4I_g5_a$ of $P \cup \{G\}$. Conversely, for every correct answer θ in $KD4I_g5_a$ of $P \cup \{G\}$, there exists a computed answer γ in $KD4I_g5_a$ of $P \cup \{G\}$ which is more general than θ (i.e. $\theta = \gamma\delta$ for some substitution δ).

A Refutation for the Wise Men Puzzle

Goals Input clauses/rules MGUs

$\leftarrow \Box_a \, white(a)$

$\varphi_1 = \Box_{abc} (white(a) \leftarrow \Diamond_b \, white(a))$

A Refutation for the Wise Men Puzzle

Goals	Input clauses/rules	MGUs
$\leftarrow \Box_a \, white(a)$		
$\leftarrow \Box_a \Diamond_b \, white(a)$	φ_1	

$rSat_L(1) = \triangle \Diamond_i E \leftarrow \triangle \langle X \rangle_i E$ for X being a fresh atom variable

A Refutation for the Wise Men Puzzle

Goals	Input clauses/rules	MGUs
$\leftarrow \Box_a \, white(a)$		
$\leftarrow \Box_a \Diamond_b \, white(a)$	φ_1	
$\leftarrow \Box_a \langle X_2 \rangle_b \, white(a)$	$rSat_L(1)$	

$\varphi_2 = \Box_{abc} (white(a) \leftarrow \Diamond_c \, white(a))$

A Refutation for the Wise Men Puzzle

Goals	Input clauses/rules	MGUs
$\leftarrow \Box_a\, white(a)$		
$\leftarrow \Box_a \Diamond_b\, white(a)$	φ_1	
$\leftarrow \Box_a \langle X_2 \rangle_b\, white(a)$	$rSat_L(1)$	
$\leftarrow \Box_a \langle X_2 \rangle_b \Diamond_c\, white(a)$	φ_2	

$rSat_L(1) = \triangle \Diamond_i E \leftarrow \triangle \langle X \rangle_i E$ for X being a fresh atom variable

A Refutation for the Wise Men Puzzle

Goals	Input clauses/rules	MGUs
$\leftarrow \Box_a \, white(a)$		
$\leftarrow \Box_a \Diamond_b \, white(a)$	φ_1	
$\leftarrow \Box_a \langle X_2 \rangle_b \, white(a)$	$rSat_L(1)$	
$\leftarrow \Box_a \langle X_2 \rangle_b \Diamond_c \, white(a)$	φ_2	
$\leftarrow \Box_a \langle X_2 \rangle_b \langle X_4 \rangle_c \, white(a)$	$rSat_L(1)$	

$\varphi_7 = \Box_{abc} (white(a) \leftarrow black(b), black(c))$

A Refutation for the Wise Men Puzzle

Goals	Input clauses/rules	MGUs
$\leftarrow \Box_a\, white(a)$		
$\leftarrow \Box_a \Diamond_b\, white(a)$	φ_1	
$\leftarrow \Box_a \langle X_2 \rangle_b\, white(a)$	$rSat_L(1)$	
$\leftarrow \Box_a \langle X_2 \rangle_b \Diamond_c\, white(a)$	φ_2	
$\leftarrow \Box_a \langle X_2 \rangle_b \langle X_4 \rangle_c\, white(a)$	$rSat_L(1)$	
$\leftarrow \Box_a \langle X_2 \rangle_b \langle X_4 \rangle_c\, black(b), \Box_a \langle X_2 \rangle_b \langle X_4 \rangle_c\, black(c)$	φ_7	

$rSat_L(2) = \triangle \nabla_i \alpha \leftarrow \triangle \Box_j \alpha$ if $g(i) \subseteq g(j)$

$\varphi_6 = \Box_{abc}\, (\Box_c\, black(b) \leftarrow black(b))$

A Refutation for the Wise Men Puzzle

Goals	Input clauses/rules	MGUs
$\leftarrow \Box_a\, white(a)$		
$\leftarrow \Box_a \Diamond_b\, white(a)$	φ_1	
$\leftarrow \Box_a \langle X_2 \rangle_b\, white(a)$	$rSat_L(1)$	
$\leftarrow \Box_a \langle X_2 \rangle_b \Diamond_c\, white(a)$	φ_2	
$\leftarrow \Box_a \langle X_2 \rangle_b \langle X_4 \rangle_c\, white(a)$	$rSat_L(1)$	
$\leftarrow \Box_a \langle X_2 \rangle_b \langle X_4 \rangle_c\, black(b), \Box_a \langle X_2 \rangle_b \langle X_4 \rangle_c\, black(c)$	φ_7	
$\leftarrow \Box_a \langle X_2 \rangle_b\, black(b), \Box_a \langle X_2 \rangle_b \langle X_4 \rangle_c\, black(c)$	$rSat_L(2)+\varphi_6$	

$\varphi_{10} = \Box_{abc} \Diamond_b\, black(b)$

A Refutation for the Wise Men Puzzle

Goals	Input clauses/rules	MGUs
$\leftarrow \Box_a\, white(a)$		
$\leftarrow \Box_a \Diamond_b\, white(a)$	φ_1	
$\leftarrow \Box_a \langle X_2 \rangle_b\, white(a)$	$rSat_L(1)$	
$\leftarrow \Box_a \langle X_2 \rangle_b \Diamond_c\, white(a)$	φ_2	
$\leftarrow \Box_a \langle X_2 \rangle_b \langle X_4 \rangle_c\, white(a)$	$rSat_L(1)$	
$\leftarrow \Box_a \langle X_2 \rangle_b \langle X_4 \rangle_c\, black(b), \Box_a \langle X_2 \rangle_b \langle X_4 \rangle_c\, black(c)$	φ_7	
$\leftarrow \Box_a \langle X_2 \rangle_b\, black(b), \Box_a \langle X_2 \rangle_b \langle X_4 \rangle_c\, black(c)$	$rSat_L(2)+\varphi_6$	
$\leftarrow \Box_a \langle black(b) \rangle_b \langle X_4 \rangle_c\, black(c)$	φ_{10}	$\{X_2/black(b)\}$

$\varphi_{11} = \Box_{abc} \Diamond_c\, black(c)$

A Refutation for the Wise Men Puzzle

Goals	Input clauses/rules	MGUs
$\leftarrow \Box_a \, white(a)$		
$\leftarrow \Box_a \Diamond_b \, white(a)$	φ_1	
$\leftarrow \Box_a \langle X_2 \rangle_b \, white(a)$	$rSat_L(1)$	
$\leftarrow \Box_a \langle X_2 \rangle_b \Diamond_c \, white(a)$	φ_2	
$\leftarrow \Box_a \langle X_2 \rangle_b \langle X_4 \rangle_c \, white(a)$	$rSat_L(1)$	
$\leftarrow \Box_a \langle X_2 \rangle_b \langle X_4 \rangle_c \, black(b), \Box_a \langle X_2 \rangle_b \langle X_4 \rangle_c \, black(c)$	φ_7	
$\leftarrow \Box_a \langle X_2 \rangle_b \, black(b), \Box_a \langle X_2 \rangle_b \langle X_4 \rangle_c \, black(c)$	$rSat_L(2) + \varphi_6$	
$\leftarrow \Box_a \langle black(b) \rangle_b \langle X_4 \rangle_c \, black(c)$	φ_{10}	$\{X_2 / black(b)\}$
$\Diamond$	φ_{11}	$\{X_4 / black(c)\}$

Conclusions

The main results of this work are:

- a sound and complete SLD-resolution calculus for the multimodal logic $KD4I_g5_a$,
- a formalization of the wise men puzzle by a modal logic program together with it refutation in $KD4I_g5_a$.

We did not consider temporal dimension, actions, and events. Thus the current version of MProlog is not yet an agent programming language like AgentSpeak(L), 3APL, and KARO.

To deal with the mentioned aspects, possible solutions are to adopt:
- CTL (like the BDI-architecture),
- (concurrent) dynamic logic (like the KARO system),
- discrete linear temporal logic.