Lecture 1 – Course Introduction

1. Practical Matters

2. What is the course about?

3. Detailed Outline

   (a) Part I – The Frame Problem: Pre- Features and Fluents

   (b) Part II – The Frame Problem: Features and Fluents

4. Course Goals

5. Tips on taking the course

Practical Matters

1. Course Material

2. Reading

3. Take Home Exercises

4. Examination

5. Who to contact

What is the course about?

1. Temporal Reasoning in AI
   (a) Different types of temporal logics
   (b) Preferential Entailment and Circumscription
2. Reasoning about Action, Change, and Planning
   (a) Existing formalisms
   (b) The Frame Problem
   (c) Features and Fluents

Temporal Reasoning

Temporal reasoning involves reasoning about change over time and the formalization of dynamically changing worlds.

A Temporal Reasoning Problem:

- A set of partial state descriptions.
- A set of action occurrences.
- Background knowledge – How actions cause change, definitional knowledge true in all states.

A Scenario Description:

`obs1` `The turkey is alive and the gun is not loaded at time 0.`

`act1` `The gun is loaded from time 2 to time 4.`

`act2` `The gun is fired from time 5 to time 6.`

`law1` **If the gun is loaded during the interval from time** t_1 **to time** t_2**, then the gun will be loaded at time** t_2**.**

`law2` **If the gun is fired during the interval from time** t_1 **to time** t_2**, then, if the gun is loaded at time** t_1 **then the gun is not loaded at time** t_2 **and the turkey is not alive at time** t_2**.**

The informal scenario description includes statements about the systems properties (`law1`,`law2`), actions that may be performed (`act1`,`act2`), and constraints on system states (`obs1`).

When translated into a set of logical formula, the reasoning task is to derive additional statements about the course of events.

Formal Scenario Description:

obs1	[0] $a \wedge \neg l$		
act1	[2,4] $Load$		
act2	[5,6] $Fire$		
law1	$[t_1, t_2]$ $Load \leadsto$	$[t_1, t_2]$ $l := T$	
law2	$[t_1, t_2]$ $Fire \leadsto$	$([t_1]$ $l \rightarrow$	
		$[t_1, t_2](\neg a \wedge \neg l) := T)$.	

Expanded Action Schemata:

obs1	[0] $a \wedge \neg l$
scd1	[2,4] $l := T$
scd2	[5] $l \rightarrow$ [5,6]$(\neg a \wedge \neg l) := T$.

Intended Completion:

[0,2] $a \wedge \neg l$
[3] a
[4,5] $a \wedge l$
[6,∞) $\neg a \wedge \neg l$

Various Reasoning Tasks

- **Prediction** – Given
  - An initial state description
  - A sequence of actions taking place in or after the initial state.

  Predict the results of performing these actions, or describe certain details of the resulting state.

- **Postdiction** – Given
  - An resulting state description
  - A sequence of actions.

  Identify the actions or state that could have led to the resulting state description.

- **Plan Synthesis** – Given
  - An initial state description s_0
  - An resulting state description s_r.

  Generate constructively, a sequence of actions that given the initial state will lead to the resulting state.

Why reason about Action and Change?

Much of AI involves the study of the rational behaviour of intelligent agents. We are intelligent because we solve problems intelligently. Consequently, our focus is on types of reasoning required to solve problems in an intelligent manner.

Problem Solving $\approx$ Planning

Informally, a *Plan* may be considered as a *detailed synthesis of a program of actions*.

Informally, reasoning about plans is a process in which a set of actions is inferred and formulated through an understanding of the world.

An agent's objective is acheived when this set of actions, which forms a plan is performed successfully.

Formalizing Reasoning about Actions and Plans

Requires Knowledge Representation of:

- **the world** – A problem domain can be regarded as a world formalized as a set S of potential world models.

- **world models** – A world model represents a state of the world at an instant in time.

- **events, actions, activities** – A set A of allowable actions can be defined in term's of the agents problem solving activities.

- **plans** – A sequence P of actions forms a plan if the agent's objective can be acheived by performing P.

The Situation Calculus

The *Situation Calculus* [McCarthy and Hayes 1969] was the earliest proposal to formalize reasoning about action and plans and still the most predominant.

Extension of classical predicate logic that introduces the notion of *situation*.

Green [1969] was the first to implement the deductive procedures associated with reasoning about actions and plans using the situation calculus.

Used resolution theorem proving to generate a constructive proof procedure to synthesize a plan of action for robots.

Definition 1 (World Model) *A world model $s \in S$ is a logical description of a state of the world at a given instant of time.s is generally represented as a set of formulas of a first-order language $L : s = \{l_1, l_2, \ldots l_m\}$, in which the l_i must be* wffs *in L.*

Definition 2 (Action) *An action $f \in A$ is a mapping from state to state. f can be defined as a binary relation R upon $S : \langle s_i, s_j \rangle \in R$, in which $s_i, s_j \in S$. It is said that s_j is directly reachable to s_i, written as $s_i \to s_j$.*

Definition 3 (World History) *Let M be a sequence of states as follows: $M = [s_0, s_1, \ldots s_{n+1}]$, in which $s_i \in S$, then M is referred to as a world history iff for any pair of adjacent states $\langle s_i, s_{i+1} \rangle$, s_i is a history of s_{i+1},* i.e., *there exists an action $f_i \in A$ such that $f_i : \langle s_i, s_{i+1} \rangle \in R, (0 \leq i \leq n)$.*

Definition 4 (Plan) *Let* c_I *and* c_G *be an initial condition and a goal condition respectively, then an action sequence* $P = [f_0, f_1, \ldots f_n]$ *forms a plan if* P *produces a world history* $M = [s_0, s_1, \ldots s_{n+1}]$ *such that* $s_0 \supset c_I$ *and* $s_{n+1} \supset c_G$*. It is said that* s_{n+1} *is reachable to* s_0*, written as* $s_0 \Longrightarrow s_{n+1}$*.*

Definition 5 (Plan Acheivement) *Let* S *be a set of potential world states,* c_I *an initial condition, and* c_G *a goal condition. A plan can be acheived by proving the following theorem:*

$$(\forall s_0) s_0 \supset c_I \vdash (\exists s_n)[(s_0 \Longrightarrow s_n) \wedge (s_n \supset c_G)]$$

which states that for an initial state $s_0 \in S$ *in which* c_I *is logically true, there exists a state* $s_n \in S$ *such that* s_n *is reachable to* s_0 *and* c_G *logically holds in* s_n*.*

The Classical Planning Problem

The problem domain is only concerned with a single agent in a state-based model of the world.

1. At any instant of time, the world is regarded as being in one of a potentially infinite number of states.

2. Various properties remain unchanged in each of these states.

3. The world changes from one state to the next only by the occurrence of an action.

This rules out: concurrency, multi-agents, surprises, continuous processes and time, etc.

Three Fundamental Problems

Reasoning about Actions and Plans gives rise to three fundamental problems in knowledge representation:

1. **Frame Problem** – The problem of describing and inferring all properties that remain unchanged when actions are performed.

2. **Qualification Problem** – The problem of specifying the preconditions under which actions take place. Specifiying the effects of actions is usually subject to qualification.

3. **Ramification Problem** – The problem of specifying the consequences of actions. What consequences should be represented explicitly and how do we assure that indirect effects are properly represented.

“The very idea of using logic as a KR language and the Situation Calculus, itself, stem from John McCarthy’s 1959 “Programs with Common Sense” . . . The frame problem also has a long (for AI) history . If AI can be said to have a classic problem in the way that Hilbert’s tenth problem was for mathematics, then the frame problem is it.” (R. Reiter)

The Frame Problem

The World: 4 blocks – $\{b_1, b_2, b_3, b_4\}$. 3 locations – $\{l_1, l_2, l_3\}$.

Initial State Axioms (s_0)

$$\begin{aligned} On(b_1, l_1), Clear(b_1) \\ On(b_2, l_2) \\ On(b_3, b_2), Clear(b_3) \\ On(b_4, l_3), Clear(b_4) \end{aligned} \tag{1}$$

Action Effect Axioms

Stack block x on top of block y:

$$Clear(x) \wedge Clear(y) \supset On(x, y)_{do(stack(x,y))} \tag{2}$$

Unstack block x to location l:

$$Clear(x) \wedge Clear(l) \supset At(x, l)_{do(unstack(x,l))} \tag{3}$$

Frame Axioms

$$
\begin{aligned}
On(z,z_1)\wedge\neg Is(x,z)&\supset On(z,z_1)_{do(stack(x,y))}\\
At(z,l)\wedge\neg Is(x,z)&\supset At(z,l)_{do(stack(x,y))}\\
Clear(z)\wedge\neg Is(z,y)&\supset Clear(z)_{do(stack(x,y))}
\end{aligned}
\tag{4}
$$

$$
\begin{aligned}
On(z,z_1)\wedge\neg Is(x,z)&\supset On(z,z_1)_{do(unstack(x,l))}\\
At(z,l_1)\wedge\neg Is(x,z)&\supset At(z,l_1)_{do(unstack(x,l))}\\
Clear(z)\wedge\neg Is(x,z)&\supset Clear(z)_{do(unstack(x,l))}
\end{aligned}
\tag{5}
$$

- **Epistemologically inadequate** – Given m possible actions and n properties, as many as $m\times n$ frame axioms may be needed.

  No flexibility – Frame axioms have exceptions. Unforeseen events may change a property.

- **Heuristically inadequate** – Most proof steps will focus on what properties remain true after the occurrence of an action.

Qualification Problem

Suppose that α is an action, P_α denotes its preconditions, and C_α its effects.

$P_\alpha = p_1 \wedge p_2 \wedge \ldots \wedge p_m$ and $C_\alpha = \{c_1, c_2 \ldots c_n\}$

where p_i and c_i are facts and relations regarding states.

The performance of an action α can be defined as follows:

$(\forall s_i)(s_i \supset P_\alpha) \vdash (\exists s_{i+1})[(s_i \rightarrow s_{i+1}) \wedge (s_{i+1} \supset C_\alpha)]$

- **Epistemologically inadequate** – It is generally not feasible to define P_α by explicitly listing all p_i that may prevent the action from being performed. For example, consider all the factors which could prevent a car from starting.

Ramification Problem

To record the consequences $C_\alpha = \{c_1, c_2 \ldots c_n\}$ of the action α, all the c_i have to be explicitly listed to represent intended changes to the world caused by α.

- **Epistemologically inadequate** – For any given action α, there are potentially an infinite number of possible results that might occur, depending upon the details of the situation in which the action occurs.

  For example, if I move my TV from one place to another, there are a number of nontrivial changes and indirect effects, such as the things on top of the TV move with it, and a limitless number of trivial changes such as the rug moving slightly, etc.

Part I – Pre- Features and Fluents

McCarthy[1959] – “Programs with Common Sense”. The Advice Taker.
McCarthy and Hayes[1969] – The Situation Calculus. Proposal for describing actions over time.

Problem: Frame Problem – Knowing what is true in a situation and what action has taken place does not imply knowing what is true in the resulting situation. Know what changes, but not what remains the same or persists.

Solution: Monotonic Frame Axioms – Express that only those properties that are explicitly changed by an action, change when that action is performed.

Frame axioms were an initial attempt at capturing a "law of inertia":

$$
\begin{aligned}
&\forall t, act, preconds, f. \\
&\quad [(\neg Cause(preconds, act, not(f)))\vee \\
&\qquad (\neg Occurs(t, act))\vee \\
&\qquad (\neg Holds(t, preconds))]\supset \\
&\qquad [Holds(t, f)\supset Holds(t+1, f)]
\end{aligned} \tag{6}
$$

If either

1. There is no causal rule yielding $not(f)$, or
2. the action causing $not(f)$ does not occur, or
3. the actions preconditions aren't satisfied

then the property f persists.

New Problem: Epistemological and Complexity – Frame axioms are too strong and there are too many.

Problem: What happens if there is only incomplete information? It appears we must defeasibly rule out:

1. unexpected causal connections,
2. unscheduled action occurrences, and
3. make defeasible assumptions about state features.

Solution: Nonmonotonic Reasoning.

Take your favorite temporal formalism, plug in a nonmonotonic logic, and use the new non-monotonic temporal formalism.

- Circumscription + (Situation Calculus)

- Default Logic +

- Autoepistemic Logic +

Include a general law of inertia:

$$\frac{Holds(sit, f) : \neg Ab(act, f, sit)}{Holds(result(act, sit), f)} \tag{7}$$

New Problem: – The Yale Shooting Problem.

The following is the Yale shooting scenario (below a and l are fluent symbols standing for $alive$ and $loaded$, respectively, and $Load$ and $Fire$ are action symbols).

obs1	[0] $a \wedge \neg l$	
ac1	[2,4] $Load$	
ac2	[5,6] $Fire$	
law1	$[t_1, t_2]$ $Load \leadsto$	$[t_1, t_2]$ $l := T$
law2	$[t_1, t_2]$ $Fire \leadsto$	$([t_1]$ $l \rightarrow$ $[t_1,t_2](\neg a \wedge \neg l) := T)$.

Question: Is the turkey alive or dead after firing the gun?

Solution: Temporal Reasoning is <u>not</u> general nonmonotonic reasoning. Special criteria are needed for preferring one model over another other than the standard set-theoretic comparison.

Essentially three types of proposed solutions:

1. Chronological Minimization

2. Causal Approaches

3. Explanation Closure

New Problem: Extreme confusion as regards the Frame and related problems.

1. Difficult to figure out what problem is being solved.

2. Difficult to compare formalisms. Tension between Situation Calculus based approaches and those using explicit temporal logic.

3. Clash of intuitions as regards intended conclusions for various scenario descriptions.

In Summary: Lack of an adequate methodology for analyzing, synthesizing and comparing approaches to reasoning about action and change

Solution?: Features and Fluents.

Sandewall proposes a general framework "Features and Fluents", which is a systematic approach to the representation of knowledge about dynamical systems.

Intention: Provide a new framework and methodology for analyzing, comparing and synthesizing logics of action and change.

Part II. – Features and Fluents

1. Introduction and methodology
2. Inhabited Dynamic Systems and Underlying Semantics
3. Epistemological and Ontological Constraints
4. Discrete Fluent Logic
5. Sandewall's Benchmark Scenario Descriptions
6. Various Logics of Preferential Entailment.
7. Verification of correctness technique.

In addition (If time permits):

- Recent work concerning PMON.
- TWEAK and Features and Fluents.
- Conclusions and Summary.

Course Goals and Philosophy

- Goals
  1. To study and understand a specific problem in KR in detail that is a current and important open problem.
  2. To learn about and use temporal logics (AI).
  3. To learn about and use non-monotonic reasoning techniques.
  4. To experience what AI research is:
     (a) How a problem arises,
     (b) How it evolves,
     (c) How we work at the fringe.
     (d) How to become a graduate student!

- Philosophy
  1. Jump right into the frying pan!
  2. Concerned with understanding of the problem not with memorizing techniques.
  3. Graduate course flavor: alot of reading, easy on the exercises.

Tips on Taking the Course

– **READ!, READ!, AND READ!**

– **THINK!, THINK!, AND THINK!**

– **Visit me, attend lectures and ask questions**

– I simply want to lift you to a level where you can discuss the problem area intelligently and where you could attend a specialized conference and understand most of the paper presentations. **The exam is only relevant for this reason.**