Lecture 31

CS 6380 Artificial Intelligence

Knowledge Representation: Reasoning Systems & Default Information

Department of Computer Science and Engineering

Indian Institute of Technology Madras (IITM), India



General-Purpose Ontologies

- Must work in *any* domain, once domain-specific axioms are added
- Must unify areas: a circuit-repair robot reasons about wiring, layout and time together
- Four ways they have been built:
  - **Experts writing axioms**: CYC
  - **Importing databases**: DBpedia, from Wikipedia infoboxes
  - **Reading text**: TextRunner, from web pages
  - **Crowdsourcing**: Open Mind, from volunteers

"*Every ontology is a treaty*" (Tom Gruber, 2004)

Categories as Objects

- Most reasoning is about categories: you want *a basketball*, not BB_9
- Either a predicate $\mathrm{Basketball}(\mathrm{b})$, or reify the category as an object $\mathrm{Basketballs}$

  $\mathrm{b} \in \mathrm{Basketballs}$ $\mathrm{Basketballs} \subset \mathrm{Balls}$
- Inheritance: Food is edible, Fruit ⊂ Food, Apples ⊂ Fruit, so every apple is edible
- Subclass links form a taxonomy: Linnaeus for species, Dewey Decimal for books
- Our vocabulary:

  $\mathrm{Alfred} \in \mathrm{IITMStudents}$, $\mathrm{IITMStudents} \subset \mathrm{Students}$, $\mathrm{Students} \subset \mathrm{Persons}$

What We Can Say About Categories

Kind of fact	Example
Membership	$BB_9 \in \text{Basketballs}$
Subclass	$\text{Basketballs} \subset \text{Balls}$
Property of every member	$(x \in \text{Basketballs}) \Rightarrow \text{Spherical}(x)$
Recognizing members	$\text{Orange}(x) \wedge \text{Round}(x) \wedge \text{Diameter}(x) = 9.5 \wedge x \in \text{Balls} \Rightarrow x \in \text{Basketballs}$
Property of the category	$\text{Dogs} \in \text{DomesticatedSpecies}$

- DomesticatedSpecies is a category of categories
- Real rules have exceptions: a punctured basketball isn't spherical

Disjoint, Exhaustive, Partition

- Saying Males ⊂ Animals and Females ⊂ Animals doesn't say no animal is both

  Disjoint({Animals, Vegetables})
  ExhaustiveDecomposition({Americans, Canadians, Mexicans}, NorthAmericans)
  Partition({Males, Females}, Animals)

- Not a partition of North Americans: some people hold two citizenships
- Definitions:

  $\text{Disjoint}(s) \Leftrightarrow (\forall c_1, c_2 \;\; c_1 \in s \wedge c_2 \in s \wedge c_1 \neq c_2 \Rightarrow \text{Intersection}(c_1, c_2) = \{\})$
  $\text{ExhaustiveDecomposition}(s, c) \Leftrightarrow (\forall i \;\; i \in c \Leftrightarrow \exists c_2 \;\; c_2 \in s \wedge i \in c_2)$
  $\text{Partition}(s, c) \Leftrightarrow \text{Disjoint}(s) \wedge \text{ExhaustiveDecomposition}(s, c)$

Definitions and Natural Kinds

- Some categories have strict definitions:

  $x \in \text{Bachelors} \Leftrightarrow \text{Unmarried}(x) \wedge x \in \text{Adults} \wedge x \in \text{Males}$

- Most don't. Natural kinds such as tomatoes vary in colour, size and shape
- Describe typical members instead:

  $\text{Typical}(c) \subseteq c \qquad x \in \text{Typical}(\text{Tomatoes}) \Rightarrow \text{Red}(x) \wedge \text{Round}(x)$

- **Wittgenstein:** what single definition covers chess, tag, solitaire and dodgeball?
- **Quine:** is the Pope a bachelor?

Parts and Wholes

- PartOf gives a part hierarchy, like the subset hierarchy

  PartOf(IITM, Chennai) PartOf(Chennai, TamilNadu)
  PartOf(TamilNadu, India) PartOf(India, Asia)

- It is transitive and reflexive:

  $\text{PartOf}(x, y) \wedge \text{PartOf}(y, z) \Rightarrow \text{PartOf}(x, z)$ $\text{PartOf}(x, x)$

- So we conclude PartOf(IITM, Asia)

- But parts are not members: a category has no mass, a composite object does

Composite Objects and Bunches

- A biped has exactly two legs attached to a body:

$$\begin{aligned}&\text{Biped}(a) \Rightarrow \exists l_1, l_2, b \;\; \text{Leg}(l_1) \wedge \text{Leg}(l_2) \wedge \text{Body}(b) \wedge \\ &\quad \text{PartOf}(l_1, a) \wedge \text{PartOf}(l_2, a) \wedge \text{PartOf}(b, a) \wedge \\ &\quad \text{Attached}(l_1, b) \wedge \text{Attached}(l_2, b) \wedge \\ &\quad l_1 \neq l_2 \wedge [\forall l_3 \;\; \text{Leg}(l_3) \wedge \text{PartOf}(l_3, a) \Rightarrow (l_3 = l_1 \vee l_3 = l_2)]\end{aligned}$$

- "The apples in this bag weigh 2 kg": a *set* has no weight
- $\text{BunchOf}(\{\text{Apple}_1, \text{Apple}_2, \text{Apple}_3\})$: an object whose parts are the apples
- Logical minimization: BunchOf(s) is the smallest object with every element of s as a part

Measurements

- A measure is an abstract object, named with a units function

  $\text{Length}(L_1) = \text{Inches}(1.5) = \text{Centimeters}(3.81)$
  $\text{Centimeters}(2.54 \times d) = \text{Inches}(d)$
  $\text{Price}(\text{MasalaDosa}_7) = ₹(80) \qquad d \in \text{Days} \Rightarrow \text{Duration}(d) = \text{Hours}(24)$

- $₹(80)$ is not a banknote: there is only one object named ₹(80)
- Some measures have no scale (difficulty, taste, beauty), but they can still be ordered

  $\text{Difficulty}(e_1) > \text{Difficulty}(e_2) \Rightarrow \text{ExpectedScore}(e_1) < \text{ExpectedScore}(e_2)$

- This is the idea behind qualitative physics

Things and Stuff

- Count nouns (aardvark, theorem) vs. mass nouns (butter, water, energy)
- Cut an aardvark in half: not two aardvarks. Cut butter in half: still butter

$$b \in \text{Butter} \land \text{PartOf}(p, b) \Rightarrow p \in \text{Butter}$$
$$b \in \text{Butter} \Rightarrow \text{MeltingPoint}(b, \text{Centigrade}(30))$$

- Intrinsic properties survive cutting: density, flavour, colour
- Extrinsic properties don't: weight, length, shape
- PoundOfButter is not stuff: halve a pound and you don't get two pounds

Events: Beyond Situation Calculus

- Situation calculus: discrete, instantaneous actions, one at a time
  - It can't describe filling a bathtub, or brushing your teeth while it fills
- Event calculus: based on points in time; reifies fluents and events

$$\mathrm{T}(\mathrm{At}(\mathrm{Alfred}, \mathrm{Chennai}), \mathrm{t})$$
$$\mathrm{E}_1 \in \mathrm{Flyings} \wedge \mathrm{Flyer}(\mathrm{E}_1, \mathrm{Alfred}) \wedge \mathrm{Origin}(\mathrm{E}_1, \mathrm{Chennai}) \wedge \mathrm{Destination}(\mathrm{E}_1, \mathrm{Delhi})$$
$$\mathrm{Happens}(\mathrm{E}_1, \mathrm{i}) \qquad \mathrm{i} = (\mathrm{t}_1, \mathrm{t}_2)$$

- Because E_1 is an object, we can add anything: $\mathrm{Bumpy}(\mathrm{E}_1)$

The Event Calculus Predicates

Predicate	Meaning
T(f, t)	Fluent f is true at time t
Happens(e, i)	Event e happens over time interval i
Initiates(e, f, t)	Event e makes fluent f start to hold at time t
Terminates(e, f, t)	Event e makes fluent f stop holding at time t
Clipped(f, i)	Fluent f stops being true at some point during i
Restored(f, i)	Fluent f becomes true at some point during i

- A special event Start sets up the initial state

Event Calculus Axioms

- A fluent holds if some event started it and nothing has stopped it since:

$$\text{Happens}(e, (t_1, t_2)) \land \text{Initiates}(e, f, t_1) \land \neg\text{Clipped}(f, (t_1, t)) \land t_1 < t \Rightarrow \text{T}(f, t)$$
$$\text{Happens}(e, (t_1, t_2)) \land \text{Terminates}(e, f, t_1) \land \neg\text{Restored}(f, (t_1, t)) \land t_1 < t \Rightarrow \neg\text{T}(f, t)$$

- Domain axioms play the role of successor-state axioms. In the wumpus world:

$$\text{Initiates}(e, \text{HaveArrow}(a), t) \Leftrightarrow e = \text{Start}$$
$$\text{Terminates}(e, \text{HaveArrow}(a), t) \Leftrightarrow e \in \text{Shootings}(a)$$

- Extensions handle simultaneous, exogenous and continuous events

Processes: Liquid Events

- Discrete events have structure: a trip interrupted halfway is a different trip
- Processes (liquid events): every sub-interval is also an instance
  - Any 20 minutes of a flight is still a flying

$$(e \in \text{Processes}) \wedge \text{Happens}(e, (t_1, t_4)) \wedge (t_1 < t_2 < t_3 < t_4) \Rightarrow \text{Happens}(e, (t_2, t_3))$$

- The same split we saw for objects:
  - In space: things vs. stuff In time: discrete events vs. processes
- So processes are temporal substances; butter is a spatial substance

Time Intervals

$\text{Partition}(\{\text{Moments}, \text{ExtendedIntervals}\}, \text{Intervals})$
$i \in \text{Moments} \Leftrightarrow \text{Duration}(i) = \text{Seconds}(0)$

- Begin and End pick out moments; Time places a moment on a time scale

  $\text{Interval}(i) \Rightarrow \text{Duration}(i) = (\text{Time}(\text{End}(i)) - \text{Time}(\text{Begin}(i)))$
  $\text{Time}(\text{Begin}(\text{AD1900})) = \text{Seconds}(0)$
  $\text{Time}(\text{Begin}(\text{AD2001})) = \text{Date}(0, 0, 0, 1, \text{Jan}, 2001)$

- The zero point is arbitrary; here it is midnight GMT on 1 January 1900

Allen's Interval Relations

Meet(i, j)

$End(i) = Begin(j)$



Overlap(i, j)

$Begin(i) < Begin(j) < End(i) < End(j)$



Before(i, j)

$End(i) < Begin(j)$



Begins(i, j)

$Begin(i) = Begin(j)$



After(i, j)

Before(j, i)



Finishes(i, j)

$End(i) = End(j)$



During(i, j)

$Begin(j) < Begin(i) < End(i) < End(j)$



Equals(i, j)

Begin and End both equal



Dark bar: interval i. Light bar: interval j. (Allen, 1983)

Interval Relations in Use

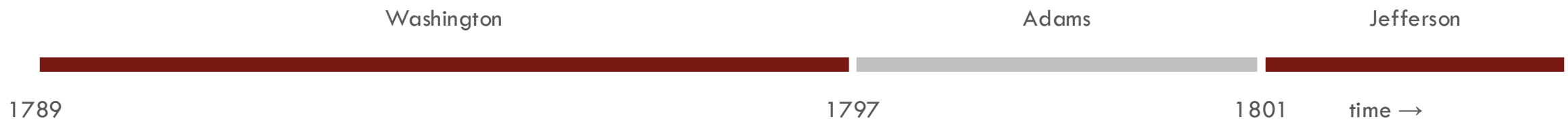
$\text{Meets}(\text{ReignOf}(\text{GeorgeVI}), \text{ReignOf}(\text{ElizabethII}))$
$\text{Overlap}(\text{Fifties}, \text{ReignOf}(\text{Elvis}))$
$\text{Begin}(\text{Fifties}) = \text{Begin}(\text{AD1950}) \qquad \text{End}(\text{Fifties}) = \text{End}(\text{AD1959})$

- Careful: $\text{Overlap}(i, j)$ is not symmetric here; i must begin first
- Interval reasoning appears in scheduling, planning and medical records
  - "The fever began during the course of antibiotics"
- Try it: which relation holds between your B.Tech and your time at IITM?

Fluents and Objects

- A physical object is a chunk of space-time: a generalized event
- USA began in 1776 with 13 states and now has 50; Population(USA) is a fluent
- President(USA) is one object made of different people at different times

  T(Equals(President(USA), GeorgeWashington), AD1790)
- Why Equals and not =? Logical identity cannot change over time



Mental Objects

- Agents that reason about knowledge: their own and other agents'
- "What is $\sqrt{1764}$?" "I don't know." "Think harder!"
  - Thinking harder can help
- "Is your mother sitting down right now?"
  - Thinking won't help, but asking her will: she knows
- Propositional attitudes: Believes, Knows, Wants, Intends, Informs
- They don't behave like ordinary predicates

Referential Opacity

- From Lecture 23: Chennai and Madras are two constants. The interpretation decides if they name one city

  $\text{Madras} = \text{Chennai}$
  $\text{Knows}(\text{Alfred}, \text{HasHotSummers}(\text{Madras}))$
  $\models \text{Knows}(\text{Alfred}, \text{HasHotSummers}(\text{Chennai}))$ yet Alfred may not know of the 1996 renaming

- FOL is referentially transparent: only the object matters, not its name
- Knowledge needs referential opacity: the name used matters
- The textbook's version: Lois knows Superman can fly. Does she know Clark Kent can?

Modal Logic and Possible Worlds

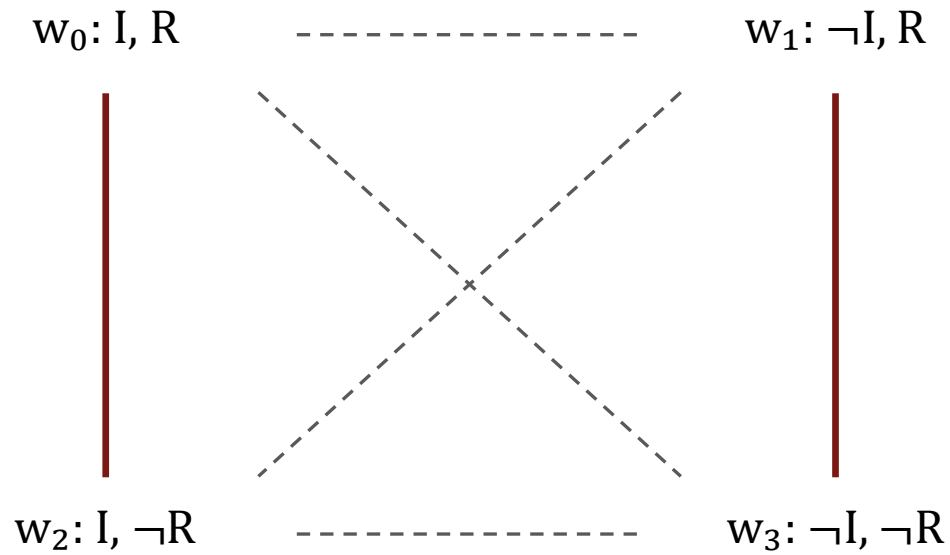
- Modal operators take sentences as arguments: $K_a\ P$ means "A knows P"
- A model is a set of possible worlds, linked by an accessibility relation for each agent
  - w_1 is accessible from w_0 if it is consistent with everything A knows in w_0
- $K_a\ P$ is true in w iff P is true in every world accessible from w
- Knowledge can be nested:

$$K_{Lois}\,[K_{Clark}\,\text{Identity}(\text{Superman}, \text{Clark}) \vee K_{Clark}\,\neg\text{Identity}(\text{Superman}, \text{Clark})]$$

  - Lois doesn't know whether Superman is Clark, but she knows that Clark knows

Possible Worlds: An Example

- I: Superman's secret identity is Clark Kent
- R: tomorrow's forecast says rain
- Superman knows I, but not R: from w_0 he can reach w_0 and w_2
- Lois knows neither: she can reach all four worlds
- In every world Lois can reach, Superman knows whether I. So Lois knows that Superman knows



Solid: Superman's accessibility. Dashed: Lois's (she can also use the solid links). Every world is accessible from itself; those links aren't drawn.