Inference & Logical Agents

Based on slides from AIMA, Marie desJardin, Charles R. Dyer, Richard Lathrop

1

Bookkeeping

- Notes
  - HW4 out **tonight**—this will be the last homework (because projects)
  - **NO CLASS NEXT THURSDAY**—*please* work on projects!
- Today's class
  - A little about higher-order logics
  - More on knowledge-based agents
    - Situations – reasoning over time
  - Inference in knowledge bases, 5 ways
- And a PSA (next)

2

Early Voting: Ends Thursday 10/31

https://elections.maryland.gov/voting/early_voting.html

Early Voting

2024 Presidential General Election

You can vote in person during early voting. Early voting will be held from Thursday, October 24, 2024 through Thursday, October 31, 2024. Early voting centers will be open from 7 am to 8 pm.

Who can vote early?

Any person that is registered to vote can vote during early voting. Any person that is eligible to register to vote can vote during early voting.

Registered voters have always been able to vote during early voting, but now individuals who are eligible but not yet registered can register and vote.

To register and vote during early voting, go to an early voting center in the county where you live and bring a document that proves where you live. This document can be your MVA-issued license, ID card, change of address card, your paycheck, bank statement, utility bill, or other government document with your name and new address. You will be able to register to vote and vote.

When can I vote early?

For the 2024 Presidential General Election, early voting will be available from Thursday, October 24, 2024 through Thursday, October31, 2024 (including Saturday and Sunday) from 7 am to 8 pm.

3

Election Day, 11/5: Find Your Polling Place

https://voterservices.elections.maryland.gov/VoterSearch

You can use this tool to:

- Review your voter registration record.
- Request a duplicate voter registration card.
- Find out where to vote.
- Find out your voting districts.
- Find out your local board of elections.
- See the status of your mail-in or provisional ballot.
- View your sample ballot.

Search

First Name:
Provide your first name as it appears on your voter registration record.

Last Name:
Provide your last name as it appears on your voter registration record.

Date Of Birth:
Provide your date of birth as it appears on your voter registration record.
Month: MM **Day:** DD **Year:** YYYY

Zip Code:
Provide the 5 digit zip code of the Maryland residential address provided on your voter registration record.

I'm not a robot
reCAPTCHA
Privacy - Terms

Web Site Privacy Policy

Search Clear

4

Eligible but Not Registered? Same Day Registration!

When may I apply to register to vote?

You can use Maryland's Online Voter Registration System (OLVR) or submit a voter registration application to your local board of elections or the State Board of Elections.

You can also register to vote during early voting or on election day. Go to an early voting center in the county where you live during early voting, or go to your assigned election day polling place and bring a document that proves where you live.

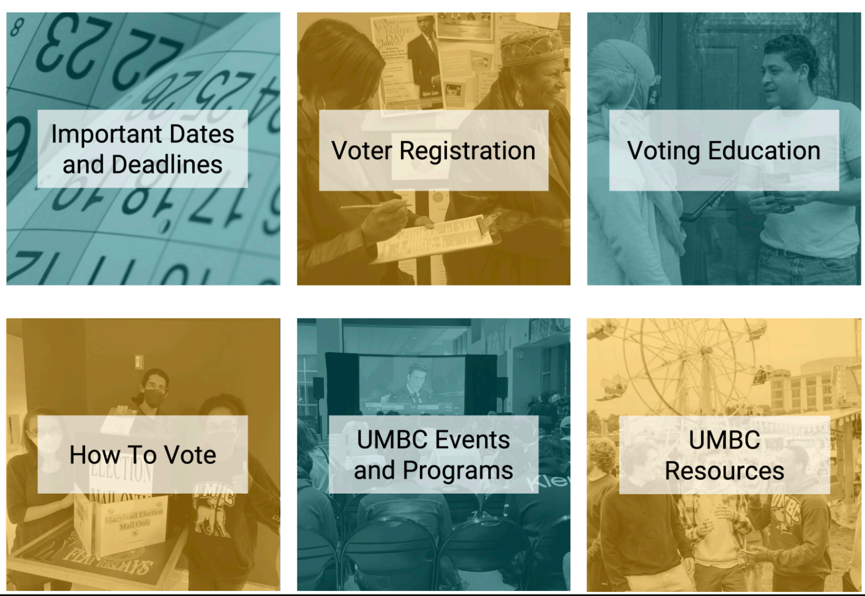
To prove where you live, bring your MVA-issued license, ID card, or change of address card, or your paycheck, bank statement, utility bill, or other government document with your name and new address.

https://elections.maryland.gov/voter_registration/index.html

5

https://civiclife.umbc.edu/learning-engagement/election-resources/

Welcome to UMBC's election resource page. Click the images below to access a variety of resources for engaging with the election. If you have any questions, please contact us at civiclife@umbc.edu.



6

Translating English to FOL

- **No purple mushroom is poisonous.**
  - $\neg\exists x$ purple(x) $\wedge$ mushroom(x) $\wedge$ poisonous(x)
  - $\forall x$ (mushroom(x) $\wedge$ purple(x)) $\Rightarrow$ $\neg$poisonous(x)

Equivalent

- **There are exactly two purple mushrooms.**
  - $\exists x \, \exists y$ mushroom(x) $\wedge$ purple(x) $\wedge$ mushroom(y) $\wedge$ purple(y) ^ $\neg$(x=y) $\wedge$ $\forall z$ (mushroom(z) $\wedge$ purple(z)) $\Rightarrow$ ((x=z) $\vee$ (y=z))
- **Mary is not tall.**
  - $\neg$tall(Mary)
- **X is above Y iff X is on directly on top of Y or there is a pile of one or more other objects directly on top of one another starting with X and ending with Y.**
  - $\forall x \, \forall y$ above(x,y) $\leftrightarrow$ (on(x,y) $\vee$ $\exists z$ (on(x,z) $\wedge$ above(z,y)))

7

Semantics of FOL

- **Domain M:** the set of all objects in the world (of interest)
- **Interpretation I:** includes
  - Assign each constant to an object in M
  - Define each function of n arguments as a mapping M^n => M
  - Define each predicate of n arguments as a mapping M^n => {T, F}
  - Therefore, every ground predicate with any instantiation will have a truth value
  - In general there is an infinite number of interpretations because |M| is infinite
- **Define logical connectives:** ~, ^, v, =>, <=> as in PL
- **Define semantics of ($\forall$x) and ($\exists$x)**
  - ($\forall$x) P(x) is true iff P(x) is true under all interpretations
  - ($\exists$x) P(x) is true iff P(x) is true under some interpretation

8

Terminology

- **Model**: an interpretation of a set of sentences such that every sentence is True
- A sentence is
  - **Satisfiable** if it is true under some interpretation
  - **Valid** if it is true under all possible interpretations
  - **Inconsistent** if there does not exist any interpretation under which the sentence is true
- **Logical consequence**: S ⊨ X if all models of S are also models of X

9

Axioms, Definitions and Theorems

- **Axioms** are facts and rules that attempt to capture all of the (important) facts and concepts about a domain; axioms can be used to prove **theorems**
  - Mathematicians don't want any unnecessary (dependent) axioms –ones that can be derived from other axioms
  - Dependent axioms can make reasoning faster, however
  - Choosing a good set of axioms for a domain is a kind of design problem
- A definition of a predicate is of the form "p(X) ↔ …" and can be decomposed into two parts
  - **Necessary** description: "p(x) ⇒ …"
  - **Sufficient** description "… ⇒ p(x)"
  - Some concepts don't have complete definitions (e.g., person(x))

10

Necessary and Sufficient

- p is **necessary** for q
  - $\neg p \Rightarrow \neg q$ ("no p, no q!")
- p is **sufficient** for q
  - $p \Rightarrow q$ ("p is all we need to know!")
- Note that $\neg p \Rightarrow \neg q$ is equivalent to $q \Rightarrow p$
- So if p is necessary and sufficient for q, then p iff q.

11

More on Definitions

- Examples: define father(x, y) by parent(x, y) and male(x)
  - parent(x, y) is a **necessary (but not sufficient**) description of father(x, y)
    - father(x, y) ⇒ parent(x, y)
  - parent(x, y) ^ male(x) ^ age(x, 35) is a **sufficient (but not necessary**) description of father(x, y):
    - father(x, y) ⟵ parent(x, y) ^ male(x) ^ age(x, 35)
  - parent(x, y) ^ male(x) is a **necessary and sufficient** description of father(x, y)
    - parent(x, y) ^ male(x) ⟷ father(x, y)

12

Converting FOL to CNF

- Eliminate biconditionals and implications
- Move ¬ inwards
- **Standardize variables apart** by renaming them: each quantifier should use a different variable
- **Skolemize:** each existential variable is replaced by a Skolem constant or Skolem function of the enclosing universally quantified variables.
  - For instance, ∃x Rich(x) becomes Rich(G1) where G1 is a new Skolem constant
  - "Everyone has a heart" [∀x Person(x) ⇒ ∃y Heart(y)∧Has(x,y)] becomes ∀x Person(x) ⇒ Heart(H(x)) ∧ Has(x, H(x)), where H is a new symbol (Skolem function)
- **Drop universal quantifiers**
  - For instance, ∀ x Person(x) becomes Person(x).
- Distribute ∧ over ∨

13

Summary: First Order Logic (FOL)

- Uses the same logical symbols as Propositional Logic (PL)
- Adds: variables, quantification, predicates and functions
  - Names of terms: **constants**, **variables**, **predicates**, **functions**
- Existential and universal quantifiers can be used to create rules
- Need to be able to translate English to and from FOL
- Some extensions…

14

Higher-Order Logic

- FOL only allows to quantify over variables, and variables can only range over objects.
- HOL allows us to quantify over relations
- Example: (quantify over functions)
  - "two functions are equal iff they produce the same value for all arguments"
  - $\forall f\ \forall g\ (f = g) \leftrightarrow (\forall x\ f(x) = g(x))$
- Example: (quantify over predicates)
  - $\forall r$ transitive(r) $\Rightarrow (\forall xyz)\ r(x,y) \wedge r(y,z) \Rightarrow r(x,z))$
- More expressive, but undecidable.

15

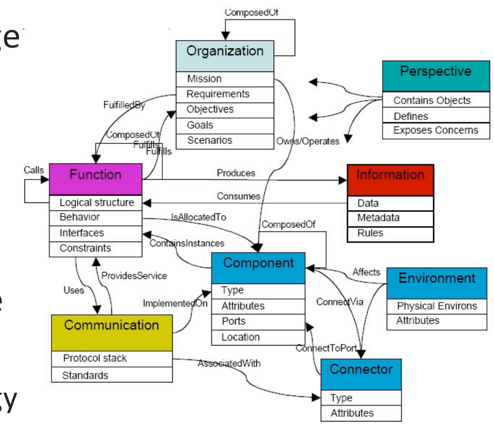
Expressing Uniqueness

- Sometimes we want to say that there is a single, unique object that satisfies a certain condition
- "There exists a unique x such that king(x) is true"
  - $\exists x\ king(x) \wedge \forall y\ (king(y) \Rightarrow x=y)$
  - $\exists x\ king(x) \wedge \neg\exists y\ (king(y) \wedge x\neq y)$
- Iota operator: "ι x P(x)" means "the unique x such that p(x) is true"
  - "The unique ruler of Freedonia is dead"
  - dead(ι x ruler(freedonia,x))

16

Knowledge bases/ontologies

- Ontology: the study of what there is—an inventory of what exists
- **An** ontology: a hierarchical categorization system for things in the world
- A formally represented corpus of knowledge
  - Defined by some *grammar*
  - Incorporates *rules* (implicitly or explicitly)
  - Not divided into tables: more like a *graph*
  - Often hierarchical
- Usually incorporate background knowledge (not purely domain-specific)
  - Although many are **in** a domain, such as biology



17

17

Ontological knowledge

- Assertions

d:

R)

  - (implies (isa ?X Cup) (isa ?X Container))

- Combine them to draw conclusions:
  - (isa ?WHAT Mug-17) → ??????

We'll do more on ontologies when we get to knowledge representation.

This is a formal representation space than can underpin questions like "What type of thing is a mug?" or "Who is Dr. M's brother?

18

Notational differences

- Different symbols for and, or, not, implies, ...
  - ∀ ∃ ⇒ ⇔ ∧ ∨ ¬ • ⊃
  - p v (q ^ r)
  - p + (q * r)
  - etc
- Prolog
  - cat(X) :- furry(X), meows (X), has(X, claws)
- Lispy notations

```
(forall ?x (implies (and (furry ?x)
                         (meows ?x)
                         (has ?x claws))
                    (cat ?x)))
```

19

Exercise: FOL translation

1. Everything is bitter or sweet.
2. Either everything is bitter or everything is sweet.
3. There is somebody who is loved by everyone.
4. Nobody is loved by no one.
5. If someone is noisy, everybody is annoyed
6. Frogs are green.
7. Frogs are not green.
8. No frog is green.
9. Some frogs are not green.
10. A mechanic likes Bob.
11. A mechanic likes herself.
12. Every mechanic likes Bob.
13. Some mechanic likes every nurse.
14. There is a mechanic who is liked by every nurse.

Exercises: disi.unitn.it/~bernardi/Courses/LSNL/Slides/fl1.pdf

20

Exercise: FOL translation

1. $\forall x\ (\text{bitter}(x) \lor \text{sweet}(x))$
2. $\forall x\ (\text{bitter}(x)) \lor \forall x\ (\text{sweet}(x))$
3. $\exists x\ \forall y\ (\text{loves}(y,x))$
4. $\neg\exists x\ \neg\exists y\ (\text{loves}(y,x))$
5. $\exists x\ (\text{noisy}(x)) \Rightarrow \forall y(\text{annoyed}(y))$
6. $\forall x\ (\text{frog}(x) \Rightarrow \text{green}(x))$
7. $\forall x\ (\text{frog}(x) \Rightarrow \neg\text{green}(x))$
8. $\neg\exists x\ (\text{frog}(x) \land \text{green}(x))$
9. $\exists x\ (\text{frog}(x) \land \neg\text{green}(x))$
10. $\exists x\ (\text{mech.}(x) \land \text{likes}(x, \text{Bob}))$
11. $\exists x\ (\text{mech.}(x) \land \text{likes}(x, x))$
12. $\forall x\ (\text{mech.}(x) \Rightarrow \text{likes}(x, \text{Bob}))$
13. $\exists x\ \forall y\ (\text{mech}(x) \land \text{nurse}(y) \Rightarrow \text{likes}(x, y))$
14. $\exists x(\text{mech}(x) \land \forall y(\text{nurse}(y) \Rightarrow \text{likes}(y, x))$

Exercises: disi.unitn.it/~bernardi/Courses/LSNL/Slides/fl1.pdf

21

A Note on Common Sense Reasoning – example adapted from Lenat

- You are told: John drove to the grocery store and bought a pound of noodles, a pound of ground beef, and two pounds of tomatoes.
  - Is John 3 years old?
  - Is John a child?
  - What will John do with the purchases?
  - Did John have any money?
  - Does John have less money after going to the store?
  - Did John buy at least two tomatoes?
  - Were the tomatoes made in the supermarket?
  - Did John buy any meat?
  - Is John a vegetarian?
  - Will the tomatoes fit in John’s car?
- Can Propositional Logic support these inferences?

22

A Note on Common Sense Reasoning

- There are a number of inferences and conclusions that we can draw that depend on *background knowledge*
- We refer to this background knowledge as “common sense”
- Can be represented as a set of statements in a knowledge base
  - Only adults can drive
  - Tomatoes weigh less than 2 pounds
  - Purchasing involves spending money
  - When you spend money, you no longer have it [etc.]
- Given these statements, we can **infer** the answers to the questions

23

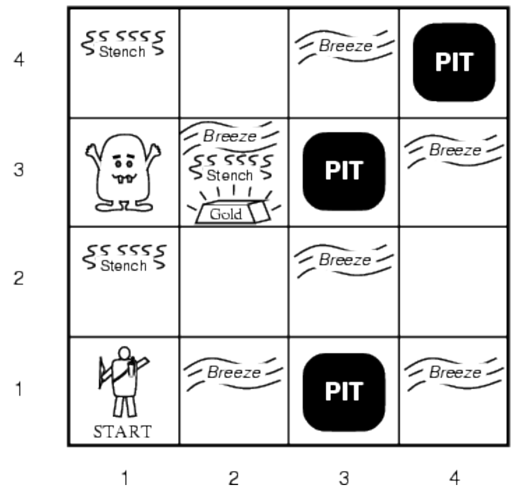
Logical Agents for Wumpus World

- Three (non-exclusive) agent architectures:
  - **Reflex** agents
    - Have rules that classify situations, specifying how to react to each possible situation
  - **Model-based** agents
    - Construct an internal model of their world
  - **Goal-based** agents
    - Form goals and try to achieve them

24

A Typical Wumpus World

- The agent always starts in the field [1,1].
- The task of the agent is to find the gold, return to the field [1,1] and climb out of the cave.



25

A Simple Reflex Agent

- Rules to **map percepts into observations**:
  - ∀b,g,u,c,l Percept([Stench, b, g, u, c], l) ⇒ Smelly(l)
  - ∀s,g,u,c,l Percept([s, Breeze, g, u, c], l) ⇒ Breezy(l)
  - ∀s,b,u,c,l Percept([s, b, Glitter, u, c], l) ⇒ AtGold(l)
- Rules to **select an action given observations**:
  - ∀l AtGold(l) ⇒ Action(Grab, l)

26

A Simple Reflex Agent

- Some difficulties:
- Climb?
  - There is no percept that indicates the agent should climb out – **position and holding gold are not part of the percept sequence**
- Loops?
  - The percept will be repeated when you return to a square, which should cause the same response (unless we maintain some **internal model of the world**)

27

KB-Agents Summary

Wumpus percepts:
[Stench, Breeze, Glitter, Bump, Scream]

- Logical agents
  - Reflex: rules map directly from percepts → beliefs or percepts → actions

    ∀b,g,u,c,t Percept([Stench, b, g, u, c],) ⇒ Smelly(l)

    ∀t AtGold(l) ⇒ Action(Grab, l)
  - Model-based: construct a *model* (set of t/f beliefs about sentences) as they learn; map from models → actions

    Action(Grab, l) ⇒ HaveGold(l) ← the consequent gets asserted into the KB

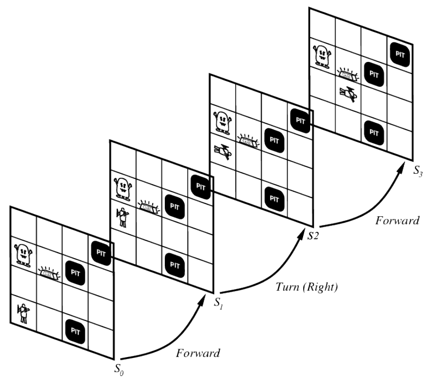
    HaveGold(l) ⇒ Action(RetraceSteps, l)
  - Goal-based: form goals, then try to accomplish them
  - Encoded as a rule:

    (∀s) Holding(Gold,s) ⇒ GoalLocation([1,1],s)

28

Representing Change

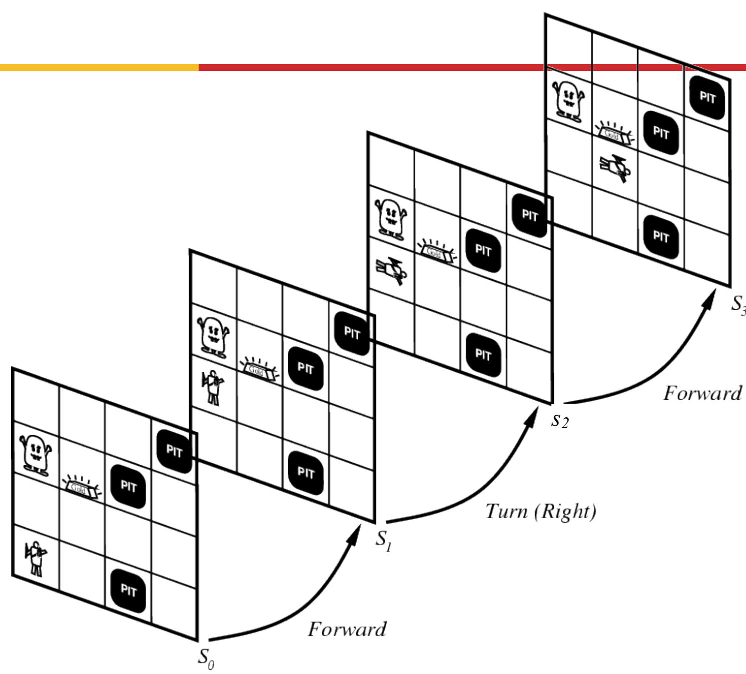
- Representing change in the world in logic can be tricky.
- One way is just to change the KB
  - Add and delete sentences from the KB to reflect changes
  - How do we remember the past, or reason about changes?
- **Situation calculus** is another way
- A **situation** is a snapshot of the world at some instant in time
- When the agent performs an action A in situation S1, the result is a new situation S2.



29

Situations

- Situations over time.
  - (We would not have this level of full knowledge.)



30

Situation Calculus

- A **situation** is:
  - A snapshot of the world
  - At an interval of time
  - During which nothing changes
- Every true or false statement is made wrt. a situation
  - Add **situation variables** to every predicate.
  - *at(Agent,1,1)* becomes *at(Agent,1,1,s0)*:
    at(Agent,1,1) **is true in** situation (i.e., state) s0.

31

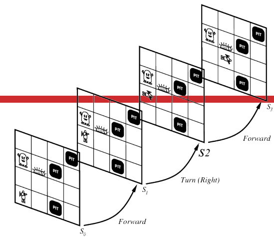
Situation Calculus

- Alternatively, add a special 2nd-order predicate, **holds(f,s),** that means "f is true in situation s." E.g., holds(at(Agent,1,1),s0)
- Or: add a new function, **result(a,s),** that maps a situation s into a new situation as a result of performing action a. For example, result(forward, s) is a function that returns the successor state (situation) to s
- Example: The action *agent-walks-to-location-y* could be represented by

$$(\forall x)(\forall y)(\forall s)\ (at(Agent,x,s) \wedge \neg onbox(s)) \Rightarrow at(Agent,y,result(walk(y),s))$$

32

Situations Summary



- Representing a dynamic world
  - Situations ($s_0 \ldots s_n$): the world in situation 0-n
    Teaching(DrM,s_0) — today, 1:00, whenNotSick, …
  - Add ‘situation’ argument to statements
    AtGold(t,s_0)
  - Or, add a ‘holds’ predicate that says ‘sentence is true in this situation’
    holds(At[2,1], s_1)
  - Or, add a result(action, situation) function that takes an action and situation, and returns a new situation
    results(Action(goNorth), s_0) → s_1

33

Deducing Hidden Properties

- From the perceptual information we obtain in situations, we can **infer properties of locations**

  l = location, s = situation

  $\forall l,s$ at(Agent,l,s) $\wedge$ Breeze(s) $\Rightarrow$ Breezy(l)

  $\forall l,s$ at(Agent,l,s) $\wedge$ Stench(s) $\Rightarrow$ Smelly(l)

- Neither Breezy nor Smelly need situation arguments because pits and Wumpuses do not move around

34

Deducing Hidden Properties II

- We need to write some rules that relate various aspects of a single world state (as opposed to across states)
- There are two main kinds of such rules:

35

Deducing Hidden Properties II

- We need to write some rules that relate various aspects of a single world state (as opposed to across states)
- There are two main kinds of such rules:

  - **Causal rules** reflect assumed direction of causality:

    $(\forall l1,l2,s)\ At(Wumpus,l1,s) \wedge Adjacent(l1,l2) \Rightarrow Smelly(l2)$

    $(\forall\ l1,l2,s)\ At(Pit,l1,s) \wedge Adjacent(l1,l2) \Rightarrow Breezy(l2)$
- Systems that reason with causal rules are called **model-based reasoning** systems

36

Deducing Hidden Properties II

- We need to write some rules that relate various aspects of a single world state (as opposed to across states)
- There are two main kinds of such rules:
  - **Diagnostic rules** infer the presence of **hidden properties** directly from the percept-derived information. We have already seen two:

    $(\forall\ l,s)\ At(Agent,l,s) \wedge Breeze(s) \Rightarrow Breezy(l)$

    $(\forall\ l,s)\ At(Agent,l,s) \wedge Stench(s) \Rightarrow Smelly(l)$

37

Frames: A Data Structure

- A frame divides knowledge into substructures by representing "stereotypical situations."
- Situations can be visual scenes, structures of physical objects, …
- Useful for representing commonsense knowledge.

Slots	Fillers
publisher	Thomson
title	Expert Systems
author	Giarratano
edition	Third
year	1998
pages	600

Slot	Fillers
name	computer
specialization_of	a_kind_of machine
types	(desktop, laptop,mainframe,super) if-added: Procedure ADD_COMPUTER
speed	default: faster if-needed: Procedure FIND_SPEED
location	(home,office,mobile)
under_warranty	(yes, no)

intelligence.worldofcomputing.net/knowledge-representation/frames.html#.WCHhCNxBo8A

38

Representing Change: The Frame Problem

- **Frame axioms:** If property x doesn't change as a result of applying action a in state s, then it stays the same.
  - On (x, z, s) $\wedge$ Clear (x, s) $\Rightarrow$
    On (x, table, Result(Move(x, table), s)) $\wedge$
    $\neg$On(x, z, Result (Move (x, table), s))
  - On (y, z, s) $\wedge$ y$\neq$ x $\Rightarrow$ On (y, z, Result (Move (x, table), s))
- The proliferation of frame axioms becomes very cumbersome in complex domains

39

The Frame Problem II

- **Successor-state axiom:** General statement that characterizes **every way** in which a particular predicate can become true:
  - Either it can be **made true**, or it can **already be true and not be changed**:
  - On (x, table, Result(a,s)) $\leftrightarrow$
    [On (x, z, s) $\wedge$ Clear (x, s) $\wedge$ a = Move(x, table)] v
    [On (x, table, s) $\wedge$ a $\neq$ Move (x, z)]
- In complex worlds with longer chains of action, even these are too cumbersome
  - Planning systems use special-purpose inference to reason about the expected state of the world at any point in time during a multi-step plan

40

Qualification Problem

- Qualification problem: How can you possibly characterize every single effect of an action, or every single exception that might occur?
- When I put my bread into the toaster, and push the button, it will become toasted after two minutes, unless…
  - The toaster is broken, or…
  - The power is out, or…
  - I blow a fuse, or…
  - A neutron bomb explodes nearby and fries all electrical components, or…
  - A meteor strikes the earth, and the world we know it ceases to exist, or…

41

Ramification Problem

- How do you describe every effect of every action?
  - When I put my bread into the toaster, and push the button, the bread will become toasted after two minutes, and…
    - The crumbs that fall off the bread onto the bottom of the toaster over tray will also become toasted, and…
    - Some of the aforementioned crumbs will become burnt, and…
    - The outside molecules of the bread will become "toasted," and…
    - The inside molecules of the bread will remain more "breadlike," and…
    - The toasting process will release a small amount of humidity into the air because of evaporation, and…
    - The heating elements will become a tiny fraction more likely to burn out the next time I use the toaster, and…
    - The electricity meter in the house will move up slightly, and…

42

Knowledge Engineering!

- Modeling the “right” conditions and the “right” effects at the “right” level of abstraction is very difficult
- Knowledge engineering (creating and maintaining knowledge bases for intelligent reasoning) is a field
- Many researchers hope that automated knowledge acquisition and machine learning tools can fill the gap:
  - Our intelligent systems should be able to learn about the conditions and effects, just like we do.
  - Our intelligent systems should be able to learn when to pay attention to, or reason about, certain aspects of processes, depending on the context.

43

Preferences Among Actions

- A problem with the Wumpus world knowledge base: It’s hard to decide which action is best!
  - Ex: to decide between a forward and a grab, axioms describing when it is okay to move would have to mention glitter.
- This is not modular!
- We can solve this problem by **separating facts about actions from facts about goals.**
- This way our **agent can be reprogrammed just by asking it to achieve different goals.**

44

Preferences Among Actions

- The first step is to describe the desirability of actions independent of each other.
- In doing this we will use a simple scale: actions can be Great, Good, Medium, Risky, or Deadly.
- Obviously, the agent should always do the best action it can find:
  - $(\forall a,s)$ Great(a,s) $\Rightarrow$ Action(a,s)
  - $(\forall a,s)$ Good(a,s) $\wedge \neg(\exists b)$ Great(b,s) $\Rightarrow$ Action(a,s)
  - $(\forall a,s)$ Medium(a,s) $\wedge$ $(\neg(\exists b)$ Great(b,s) $\vee$ Good(b,s)) $\Rightarrow$ Action(a,s)

    ...

45

Preferences Among Actions

- We use this action quality scale in the following way.
- Until it finds the gold, the basic strategy for our agent is:
  - Great actions include picking up the gold when found and climbing out of the cave with the gold.
  - Good actions include moving to a square that’s OK and hasn't been visited yet.
  - Medium actions include moving to a square that is OK and has already been visited.
  - Risky actions include moving to a square that is not known to be deadly or OK.
  - Deadly actions are moving into a square that is known to have a pit or a Wumpus.

46

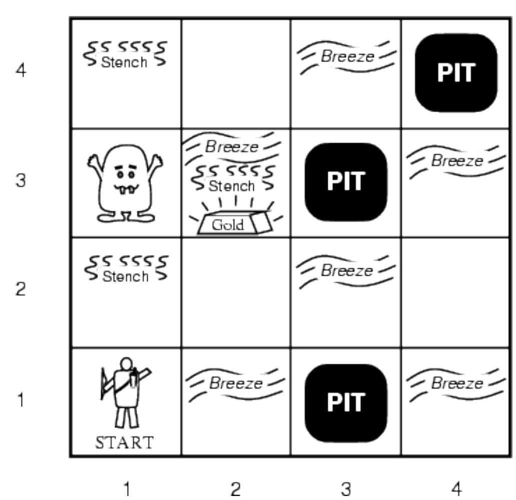
Goal-Based Agents

- Once the gold is found, it is necessary to change strategies. So now we need a new set of action values.
- We could encode this as a rule:
  - $(\forall s)$ Holding(Gold,s) $\Rightarrow$ GoalLocation([1,1]),s)
- We must now decide how the agent will work out a sequence of actions to accomplish the goal.
- Three possible approaches are:
  - **Inference**: good versus wasteful solutions
  - **Search**: make a problem with operators and set of states
  - **Planning**: coming soon!

47

An agent needs to make decisions!

- Where is there a pit?
- Where is there a wumpus?
- Should I fire my arrow?
- Where to explore next?
- Need to draw **conclusions** from knowledge in the knowledge base
- → Inference!



48

Logical Inference

Chapter 9

49

Review: English to FOL using quantifiers

3. There is somebody who is loved by everyone.
   $\exists x\ \forall y\ (loves(y,x))$ — this person also loves themselves
6. Frogs are green.
   $\forall x\ (frog(x) \Rightarrow green(x))$ — this is how we express rules
10. A mechanic likes Bob.
   $\exists x\ (mech.(x) \wedge likes(x, Bob))$ — express existence of something

- Usually:
  - Use AND with $\exists$ (so we don't have a false antecedent)
  - Use IMPLIES with $\forall$ (so we don't make too-strong claims)

50

Syntactic Ambiguity

- FOL provides many ways to represent the same thing.
- E.g., "Ball-5 is red."
  - **HasColor(Ball-5, Red):** Ball-5 and Red are objects related by HasColor.
  - **Red(Ball-5):** Red is a unary predicate applied to the Ball-5 object.
  - **HasProperty(Ball-5, Color, Red):** Ball-5, Color, and Red are objects related by HasProperty.
  - **ColorOf(Ball-5) = Red**: Ball-5 and Red are objects, and ColorOf() is a function.
  - **HasColor(Ball-5(), Red()):** Ball-5() and Red() are functions of zero arguments that both return an object, which objects are related by HasColor.
- This can GREATLY confuse a pattern-matching reasoner.

51

More choices to be made

- "For every food, there is a person who eats that food."
- [Use: Food(x), Person(y), Eats(y, x)]
  - $\forall x\ \exists y\ Food(x) \Rightarrow (Person(y) \wedge Eats(y, x))$
  - $\forall x\ Food(x) \Rightarrow \exists y\ (Person(y) \wedge Eats(y, x))$
  - $\forall x\ \exists y\ \neg Food(x) \vee (Person(y) \wedge Eats(y, x))$
  - $\forall x\ \exists y\ (\neg Food(x) \vee Person(y)) \wedge (\neg Food(x) \vee Eats(y, x))$
  - $\forall x\ \exists y\ (Food(x) \Rightarrow Person(y)) \wedge (Food(x) \Rightarrow Eats(y, x))$
- Common Mistakes:
  - $\forall x\ \exists y\ (Food(x) \wedge Person(y)) \Rightarrow Eats(y, x)$
  - $\forall x\ \exists y\ Food(x) \wedge Person(y) \wedge Eats(y, x)$

52

And yet more…

- "Every person eats some food."
- [Use: Person (x), Food (y), Eats(x, y)]
  - $\forall x\ \exists y\ Person(x) \Rightarrow [\ Food(y) \wedge Eats(x, y)\ ]$
  - $\forall x\ Person(x) \Rightarrow \exists y\ [\ Food(y) \wedge Eats(x, y)\ ]$
  - $\forall x\ \exists y\ \neg Person(x) \vee [\ Food(y) \wedge Eats(x, y)\ ]$
  - $\forall x\ \exists y\ [\ \neg Person(x) \vee Food(y)\ ] \wedge [\ \neg Person(x) \vee Eats(x, y)\ ]$
- Common Mistakes:
  - $\forall x\ \exists y\ [\ Person(x) \wedge Food(y)\ ] \Rightarrow Eats(x, y)$
  - $\forall x\ \exists y\ Person(x) \wedge Food(y) \wedge Eats(x, y)$

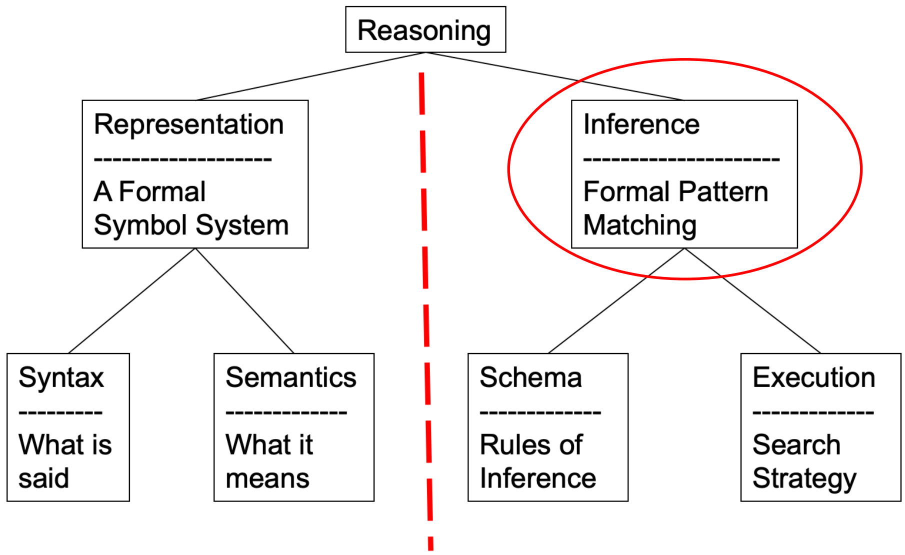
53

Syntactic Ambiguity—Partial Solution

- FOL can be **too** expressive, can offer **too many** choices
- Likely confusion, especially for teams of Knowledge Engineers
- Different team members can make different representation choices
  - E.g., represent "Ball43 is Red." as:
    - a predicate (= verb)? E.g., "Red(Ball43)" ?
    - an object (= noun)? E.g., "Red = Color(Ball43))" ?
    - a property (= adjective)? E.g., "HasProperty(Ball43, Red)" ?
- **Partial solution**
  - An upon-agreed ontology that settles these questions
  - Ontology = what exists in the world & how it is represented
  - The Knowledge Engineering teams agrees upon an ontology BEFORE they begin encoding knowledge

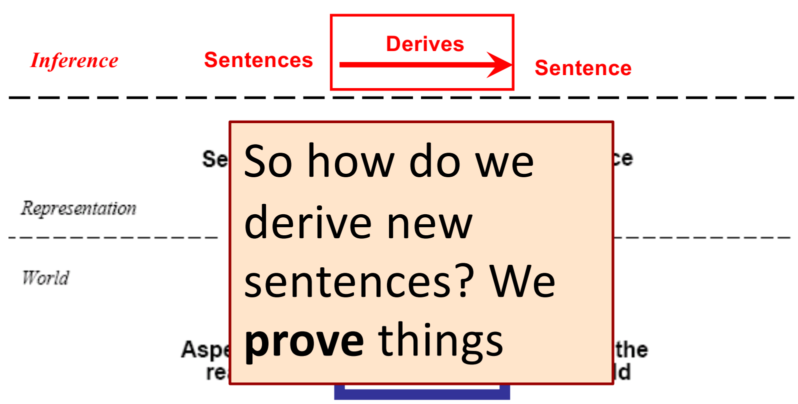
56

FOL (or FOPC) Ontology:
What kind of things exist in the world?
What do we need to describe and reason about?
Objects --- with their relations, functions, predicates, properties, and general rules.



57

Reminder: Schematic perspective



So how do we derive new sentences? We **prove** things

If KB is true in the real world,
*then any sentence α **derived** from KB*
by a sound inference procedure
is also true in the real world.

58

Proof methods

- Proof methods divide into (roughly) two kinds:
- Model checking:
  - Searching through truth assignments.
  - Improved backtracking: Davis-Putnam-Logemann-Loveland (DPLL)
  - Heuristic search in model space: Walksat
- Application of inference rules:
  - Legitimate (sound) generation of new sentences from old.
  - Forward & Backward chaining
  - Resolution — KB is in Conjunctive Normal Form (CNF)

59

Model Checking

- Given some knowledge base (KB), does sentence S hold?
- Basically **generate and test**:
  - Generate all the possible models
  - Consider the models M in which KB is TRUE
  - If $\forall M(S)$, then S is **provably true**
  - If $\forall M(\neg S)$, then S is **provably false**
  - Otherwise ($\exists M1\ S \wedge \exists M2\ \neg S$): S is **satisfiable** but neither provably true or provably false

What does model mean?

Model: an interpretation – or assignment of truth values to literals – of a set of sentences such that every sentence is *True*

60

Method 1: Inference by Enumeration

- Also called **Model Checking** or **Truth Table Enumeration**
- LET: KB = A ∨ C, B ∨ ¬C β = A ∨ B
- QUERY: KB ⊨ β ?

A	B	C
false	false	false
false	false	true
false	true	false
false	true	true
true	false	false
true	false	true
true	true	false
true	true	true

NOTE: The computer doesn't know the meaning of the proposition symbols

So, *all* logically distinct cases must be checked to prove that a sentence *can* be derived from KB

62

Inference by Enumeration

- LET: KB = A ∨ C, B ∨ ¬C β = A ∨ B
- QUERY: KB ⊨ β ?

A	B	C	AVC	BV ~C	KB
false	false	false	false	true	false
false	false	true	true	false	false
false	true	false	false	true	false
false	true	true	true	true	true
true	false	false	true	true	true
true	false	true	true	false	false
true	true	false	true	true	true
true	true	true	true	true	true

Rows where *all* of sentences in KB are true are the models of KB

63

Inference by Enumeration

- LET: KB = A ∨ C, B ∨ ¬C β = A ∨ B
- QUERY: KB ⊨ β ?

A	B	C	A∨C	B∨¬C	KB	A∨B	KB⇒β
false	false	false	false	true	false	false	true
false	false	true	true	false	false	false	true
false	true	false	false	true	false	true	true
false	true	true	true	true	true	true	true
true	false	false	true	true	true	true	true
true	false	true	true	false	false	true	true
true	true	false	true	true	true	true	true
true	true	true	true	true	true	true	true

β is entailed by `KB` if *all* models of `KB` are models of β, i.e., *all* rows where `KB` is true, β is also true

In other words: `KB` ⇒ β is valid

Inference by Enumeration

- Using inference by enumeration to build a complete truth table in order to tell if a sentence is entailed by KB is a complete inference algorithm for Propositional Logic
- But very slow: exponential time
- Imagine we had 5 literals… or 30, or hundreds, or millions

Method 2: Natural Deduction = Constructing a Proof

- A Proof is a sequence of inference steps that leads from α (i.e., KB) to β (i.e., query)
- This is a search problem!

KB:
$(P \land Q) \Rightarrow R$
$(S \land T) \Rightarrow Q$
S
T
P

Query:
R

67

Proof by Natural Deduction

KB:
$(P \land Q) \Rightarrow R$
$(S \land T) \Rightarrow Q$
S
T
P

Query:
R

1.	S	Premise (i.e., given sentence in KB)
2.	T	Premise
3.	$S \land T$	Conjunction(1, 2) (And-Introduction)
4.	$(S \land T) \Rightarrow Q$	Premise
5.	Q	Modus Ponens(3, 4)
6.	P	Premise
7.	$P \land Q$	Conjunction(5, 6)
8.	$(P \land Q) \Rightarrow R$	Premise
9.	R	Modus Ponens(7, 8)

This is the expected format of proofs in the homework!

68

Proof by Natural Deduction

- KB:
  - HaveLecture ⇔ (isTuesday ∨ isThursday)
  - ¬ HaveLecture
- Query:
  - ¬ isTuesday

Proof

Contraposition:
(P ⇒ Q) ≡ (¬Q ⇒ ¬P)

KB:

1. HaveLecture ⇔ (isTuesday ∨ isThursday)
2. ¬HaveLecture

3. (HaveLecture ⇒ (isTuesday ∨ isThursday)) ∧ ((isTuesday ∨ isThursday) ⇒ HaveLecture)	iff elimination to 1
4. (isTuesday ∨ isThursday) ⇒ HaveLecture	and-elimination to 3
5. ¬HaveLecture ⇒ ¬(isTuesday ∨ isThursday)	contraposition to 4
6. ¬(isTuesday ∨ isThursday)	Modus Ponens 2,5
7. ¬isTuesday ∧ ¬isThursday	de Morgan to 6
8. ¬isTuesday	and-elimination to 7

Automating FOL Inference with Generalized Modus Ponens

71

Automated Inference for FOL

- Automated inference using FOL is harder than PL
  - Variables can take on an infinite number of possible values
    - From their domains, anyway
    - This is a reason to do careful KR!
  - So, potentially infinite ways to apply Universal Elimination
- *Godel's Completeness Theorem* says that FOL entailment is only semidecidable*
  - If a sentence is **true** given a set of axioms, can prove it
  - If the sentence is **false**, then there is no guarantee that a procedure will ever determine this
  - **Inference may never halt**

*The "halting problem"

72

Generalized Modus Ponens (GMP)

- Apply modus ponens reasoning to generalized rules
- Combines And-Introduction, Universal-Elimination, and Modus Ponens
  - From *P(c)* and *Q(c)* and $(\forall x)(P(x) \wedge Q(x)) \Rightarrow R(x)$ derive *R(c)*
- General case: **Given**
  - **atomic sentences** $P_1, P_2, ..., P_N$
  - **implication sentence** $(Q_1 \wedge Q_2 \wedge ... \wedge Q_N) \Rightarrow R$
    - $Q_1, ..., Q_N$ and R are atomic sentences
  - **substitution** subst(θ, P_i) = subst(θ, Q_i) for i=1,...,N
  - **Derive new sentence: subst(θ, R)**

73

Generalized Modus Ponens (GMP)

- **Derive new sentence: subst(θ, R)**
- Substitutions
  - subst(θ, α) denotes the result of applying a **set of substitutions,** defined by θ, to the sentence α
  - A substitution list $\theta = \{v_1/t_1, v_2/t_2, ..., v_n/t_n\}$ means to replace all occurrences of variable symbol v_i by term t_i
  - Substitutions are made in left-to-right order in the list
  - subst({x/IceCream, y/Ziggy}, eats(y,x)) = eats(Ziggy, IceCream)

74

Review: Horn Clauses

- A Horn sentence or Horn clause has the form:

  $P1 \wedge P2 \wedge P3 \ldots \wedge Pn \Rightarrow Q$

- or alternatively

  $\neg P1 \vee \neg P2 \vee \neg P3 \ldots \vee \neg Pn \vee Q$

- where Ps and Q are non-negated atoms
- To get a proof for Horn sentences, apply Modus Ponens repeatedly until nothing can be done
- Horn clauses are a subset of the set of sentences representable in FOL

75

Horn Clauses II

- These are Horn clauses (special cases):
  - $P1 \wedge P2 \wedge \ldots Pn \Rightarrow Q$
  - $P1 \wedge P2 \wedge \ldots Pn \Rightarrow \text{false}$
  - $\text{true} \Rightarrow Q$
- These are not Horn clauses:
  - $p \vee q$ (all but one literal must be negated)
  - $(P \wedge Q) \Rightarrow (R \vee S)$ (non-literal after the implication)

76

Inference with Horn KBs

- If everything is Horn clauses, only 1 rule of inference needed
- Generalized Modus Ponens (GMP):

  Given P, Q, and $(P \land Q) \Rightarrow R$, conclude R

- Written as:

$$\frac{P, Q, (P \land Q) \Rightarrow R}{R}$$

77

Method 3: Forward Chaining

- Proofs start with the given axioms/premises in KB, deriving new sentences using GMP until the goal/query sentence is derived
- This defines a **forward-chaining** inference procedure because it moves “forward” from the KB to the goal [eventually]
- Forward chaining with Horn clause KB is **complete**
  - A formal system is called complete with respect to a particular property if every formula having the property can be derived using that system, i.e. is one of its theorems;
  - **Intuitively, a system is called complete if it can derive every formula that is true.**

78

Forward Chaining

- "Apply" any rule whose premises are satisfied in the KB
- Add its conclusion to the KB until query is derived

KB:
$$P \Rightarrow Q$$
$$L \wedge M \Rightarrow P$$
$$B \wedge L \Rightarrow M$$
$$A \wedge P \Rightarrow L$$
$$A \wedge B \Rightarrow L$$
$$A$$
$$B$$

query: Q

79

Forward Chaining

Given KB

1. $P \Rightarrow Q$
2. $L \wedge M \Rightarrow P$
3. $B \wedge L \Rightarrow M$
4. $A \wedge P \Rightarrow L$
5. $A \wedge B \Rightarrow L$
6. A
7. B
8. L GMP(5,6,7)
9. M GMP(3,7,8)
10. P GMP(2,8,9)
11. Q GMP(1,10)

80

Forward Chaining Example

- KB:
  - allergies(X) ⇒ sneeze(X)
  - cat(Y) ∧ allergic-to-cats(X) ⇒ allergies(X)
  - cat(Felix)
  - allergic-to-cats(Lise)
- Goal:
  - sneeze(Lise)

82

Forward Chaining

Knowledge Base

1. Allergies lead to sneezing.
   allergies(X) ⇒ sneeze(X)
2. Cats cause allergies if allergic to cats.
   cat(Y) ∧ allergic-cats(X) ⇒ allergies(X)
3. Felix is a cat.
   cat(Felix)
4. Lise is allergic to cats.
   allergic-cats(Lise)

sneeze(Lise) ← thing to infer truth of (query)

- Forward Chaining: apply rules

variable binding

cat(Y) ∧ allergic-cats(X) ⇒ allergies(X) ∧ cat(Felix)
⇒
cat(Felix) ∧ allergic-cats(X) ⇒ allergies(X) ∧ allergic-cats(Lise)
⇒
allergies(Lise) ∧ allergies(X) ⇒ sneeze(X)
⇒
sneeze(Lise) ✓

add new sentence to KB

83

Forward Chaining Exercise

- Consider the following KB:
  1. J ⇒ Q
  2. A ∧ I ⇒ J
  3. E ∧ F ⇒ I
  4. B ⇒ F
  5. A ∧ B ⇒ E
  6. A
  7. B

8. E (GMP 5,6,7)
9. F (GMP 4,7)
10. I (GMP 3,8,9)
11. J (GMP 2,6,10)
12. Q (GMP 1,11)

- Prove Q. (Remember, you'll just use GMP over and over!)
  - **A, B, (A ∧ B) ⇒ C, ∴ C**

84

Method 4: Backward Chaining

- Forward chaining problem: can generate a lot of irrelevant conclusions
  - Search forward, start state = KB, goal test = state contains query
- Backward chaining
  - Work backwards from goal to premises
  - Find all implications of the form (…) ⇒ query
  - Prove all the premises of one of these implications
  - Avoid loops: check if new subgoal is already on the goal stack
  - Avoid repeated work: check if new subgoal
    - Has already been proved true, or
    - Has already failed

85

Backward Chaining

- Backward-chaining deduction using GMP
  - Complete for KBs containing only Horn clauses.
- Proofs:
  - Start with the goal query
  - Find rules with that conclusion
  - Prove each of the antecedents in the implication
- Keep going until you reach premises!

Avoid loops
- Is new subgoal already on goal stack?

Avoid repeated work: has subgoal
- Already been proved true?
- Already failed?

86

Backward Chaining Example

- KB:
  - allergies(X) ⇒ sneeze(X)
  - cat(Y) ∧ allergic-to-cats(X) ⇒ allergies(X)
  - cat(Felix)
  - allergic-to-cats(Lise)
- Goal:
  - sneeze(Lise)

88

Backward Chaining

Knowledge Base
1. Allergies lead to sneezing.
   allergies(X) ⇒ sneeze(X)
2. Cats cause allergies if allergic to cats.
   cat(Y) ∧ allergic-cats(X) ⇒ allergies(X)
3. Felix is a cat.
   cat(Felix)
4. Lise is allergic to cats.
   allergic-cats(Lise)

sneeze(Lise) ← query

- Backward Chaining: apply rules that end with the goal

variable binding

allergies(X) → sneeze(X) + sneeze(Lise)
new query: allergies(Lise)?

cat(Y) ∧ allergic-cats(X) → allergies(X) + allergies(Lise)
new query: cat(Y) ∧ allergic-cats(Lise)?

cat(Felix) + cat(Y) ∧ allergic-cats(Lise)
new sentence: cat(Felix) ∧ allergic-cats(Lise) ✓

89

Forward vs. Backward Chaining

- FC is data-driven
  - Automatic, unconscious processing
  - E.g., object recognition, routine decisions
  - May do lots of work that is irrelevant to the goal
- BC is goal-driven, appropriate for problem-solving
  - Where are my keys? How do I get to my next class?
  - Complexity of BC can be much less than linear in the size of the KB

91

Completeness of GMP

- GMP (using forward or backward chaining) is complete for KBs that contain only Horn clauses
- It is ***not* complete** for simple KBs that contain non-Horn clauses
- The following KB entails that S(A) is true:
  - $(\forall x)\ P(x) \Rightarrow Q(x)$
  - $(\forall x)\ \neg P(x) \Rightarrow R(x)$
  - $(\forall x)\ Q(x) \Rightarrow S(x)$
  - $(\forall x)\ R(x) \Rightarrow S(x)$
- If we want to conclude S(A), with GMP we cannot, since the second one is not a Horn clause
- It is equivalent to $P(x) \vee R(x)$

92

Automating FOL Inference with Resolution

93

93

Resolution Rule of Inference

- Resolution Rule of Inference:

$$\frac{\alpha \vee \beta, \neg\beta \vee \gamma}{\alpha \vee \gamma}$$

- Examples $\dfrac{A \vee B, \neg B}{A}$ $\dfrac{A \vee B \vee \neg C \vee D, \neg A \vee \neg E \vee F}{B \vee \neg C \vee D \vee \neg E \vee F}$

94

Resolution

- Take any two "clauses" where one contains some symbol, and the other contains its complement (negative)

  P ∨ **Q** ∨ R   **¬Q** ∨ S ∨ T

- Merge (resolve) them, by throwing away the symbol and its complement, to obtain their **resolvent** clause:

  P ∨ R ∨ S ∨ T

- If two clauses resolve and there's no symbol left, you have derived False, aka the empty clause

95

Method 5: Resolution Refutation

- Show KB $\models \alpha$ by proving that KB $\wedge \neg\alpha$ is unsatisifiable, i.e., deducing False from KB $\wedge \neg\alpha$
- Your algorithm can use all the logical equivalences to derive new sentences, plus:
- **Resolution rule:** a *single* inference rule
  - **Sound:** only derives entailed sentences
  - **Complete:** can derive any entailed sentence
    - Resolution is refutation complete: if KB $\models \beta$, then KB $\wedge \neg \beta \vdash$ False
  - But the sentences need to be preprocessed into CNF
  - But all sentences can be converted into this form

96

Resolution Refutation Algorithm

1. Add negation of query to KB
2. Pick 2 sentences that haven't been used before and can be used with the Resolution Rule of inference
3. If none, halt and answer that the query is NOT entailed by KB
4. Compute resolvent and add it to KB
5. If False in KB
   - Then halt and answer that the query IS entailed by KB
   - Else Goto 2

97

Review: Converting to CNF

- Replace all ⇔ using biconditional elimination
  - $\alpha \Leftrightarrow \beta \equiv (\alpha \Rightarrow \beta) \land (\beta \Rightarrow \alpha)$
- Replace all ⇒ using implication elimination
  - $\alpha \Rightarrow \beta \equiv \neg\alpha \lor \beta$
- Move all negations inward using
  - double-negation elimination
    - $\neg(\neg\alpha) \equiv \alpha$
  - de Morgan's rule
    - $\neg(\alpha \lor \beta) \equiv \neg\alpha \land \neg\beta$
    - $\neg(\alpha \land \beta) \equiv \neg\alpha \lor \neg\beta$
- Apply distributivity of ∨ over ∧
  - $\alpha \land (\beta \lor \gamma) \equiv (\alpha \land \beta) \lor (\alpha \land \gamma)$

Result: something with clauses made up of ORs, separated by ANDs:

$\underline{(\neg A \lor B \lor C)} \land \underline{(\neg B \lor A)} \land \underline{(\neg C \lor A)}$

99

Resolution Refutation Steps

- Given KB and β (query)
- Add ¬β to KB, and convert all sentences to CNF
- Show this leads to False (aka “empty clause”). Proof by contradiction
- Example KB:
  - $A \Leftrightarrow (B \lor C)$
  - $\neg A$
- Example query: $\neg B$

100

Review: Example Conversion to CNF

- Example: A $\Leftrightarrow$ (B $\vee$ C)
- Eliminate $\Leftrightarrow$ by replacing $\alpha \Leftrightarrow \beta$ with $(\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \alpha)$.
  - $= (A \Rightarrow (B \vee C)) \wedge ((B \vee C) \Rightarrow A)$
- 2. Eliminate $\Rightarrow$ by replacing $\alpha \Rightarrow \beta$ with $\neg\alpha \vee \beta$ and simplify.
  - $= (\neg A \vee B \vee C) \wedge (\neg(B \vee C) \vee A)$
- 3. Move $\neg$ inwards using de Morgan's rules and simplify.
  - $= (\neg A \vee B \vee C) \wedge ((\neg B \wedge \neg C) \vee A)$
- 4. Apply distributive law ($\wedge$ over $\vee$) and simplify.
  - $= (\neg A \vee B \vee C) \wedge (\neg B \vee A) \wedge (\neg C \vee A)$

101

Resolution Refutation Preprocessing

Example KB:
A ⇔ (B ∨ C)
¬A
Example query:
¬B

- Add ¬β to KB, and convert to CNF:
  - 1: ¬A V B V C
  - 2: ¬B V A
  - 3: ¬C V A
  - 4: ¬A
  - 5: B

CNF conversion of A ⇔ (B ∨ C)

- Want to reach goal: False (empty clause)

102

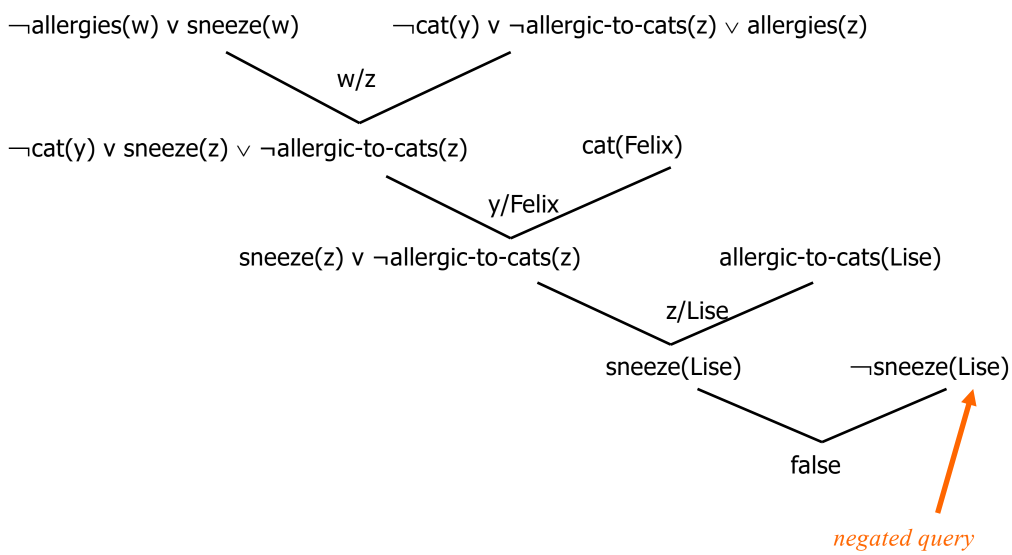
Resolution Refutation Example

- 1: ¬A ∨ B ∨ C
- 2: ¬B ∨ A
- 3: ¬C ∨ A
- 4: ¬A
- 5: B

- 6: A Resolve 2, 5
- 7: false/empty clause Resolve 6, 4

103

Refutation Resolution Proof Tree



104

Exercise: Did Curiosity Kill the Cat?

- Jack owns a dog. Every dog owner is an animal lover. No animal lover kills an animal. Either Jack or Curiosity killed the cat, who is named Tuna. **Did Curiosity kill the cat?**
- These can be represented as follows:
  - A. $(\exists x)$ Dog(x) $\wedge$ Owns(Jack,x)
  - B. $(\forall x)$ $((\exists y)$ Dog(y) $\wedge$ Owns(x, y)) $\rightarrow$ AnimalLover(x)
  - C. $(\forall x)$ AnimalLover(x) $\rightarrow$ $((\forall y)$ Animal(y) $\rightarrow \neg$Kills(x,y))
  - D. Kills(Jack,Tuna) $\vee$ Kills(Curiosity,Tuna)
  - E. Cat(Tuna)
  - F. $(\forall x)$ Cat(x) $\rightarrow$ Animal(x)
  - G. Kills(Curiosity, Tuna) ← Query

105

Resolution Proof

CLAUSAL FORM CONVERSION:
1. Eliminate all $\leftrightarrow$ connectives
2. Eliminate all → connectives
3. Reduce the scope of each negation symbol to a single predicate
4. Standardize variables
5. Eliminate existential quantification with Skolem constants/functions
6. Remove universal quantifiers
7. Put into conjunctive normal form (conjunction of disjunctions)
8. Split conjuncts into separate clauses
9. Standardize variables again

- **Did Curiosity kill the cat?**
  - A. $(\exists x)$ Dog(x) $\wedge$ Owns(Jack,x)
  - B. $(\forall x)$ $((\exists y)$ Dog(y) $\wedge$ Owns(x, y)) $\rightarrow$ AnimalLover(x)
  - C. $(\forall x)$ AnimalLover(x) $\rightarrow$ $((\forall y)$ Animal(y) $\rightarrow \neg$Kills(x,y))
  - D. Kills(Jack,Tuna) $\vee$ Kills(Curiosity,Tuna)
  - E. Cat(Tuna)
  - F. $(\forall x)$ Cat(x) $\rightarrow$ Animal(x)
  - **G. Kills(Curiosity, Tuna) – goal: must be negated in your KB!**

RESOLUTION STEP: Given sentences

$P_1 \vee \ldots \vee P_n$ $\qquad$ $Q_1 \vee \ldots \vee Q_m$

- if P_j and $\neg Q_k$ **unify** with substitution list θ, then derive the resolvent sentence:
  subst(θ, $P_1 \vee \ldots \vee P_{j-1} \vee P_{j+1} \ldots P_n \vee Q_1 \vee \ldots Q_{k-1} \vee Q_{k+1} \vee \ldots \vee Q_m$)

106

Steps

- Convert to clause form
  - A1. (Dog(D)) ← D is a skolem constant
  - A2. (Owns(Jack,D))
  - B. (¬Dog(y), ¬Owns(x, y), AnimalLover(x))
  - C. (¬AnimalLover(a), ¬Animal(b), ¬Kills(a,b))
  - D. (Kills(Jack,Tuna), Kills(Curiosity,Tuna))
  - E. Cat(Tuna)
  - F. (¬Cat(z), Animal(z))
- Add the negation of query:
  - ¬G: (¬Kills(Curiosity, Tuna))

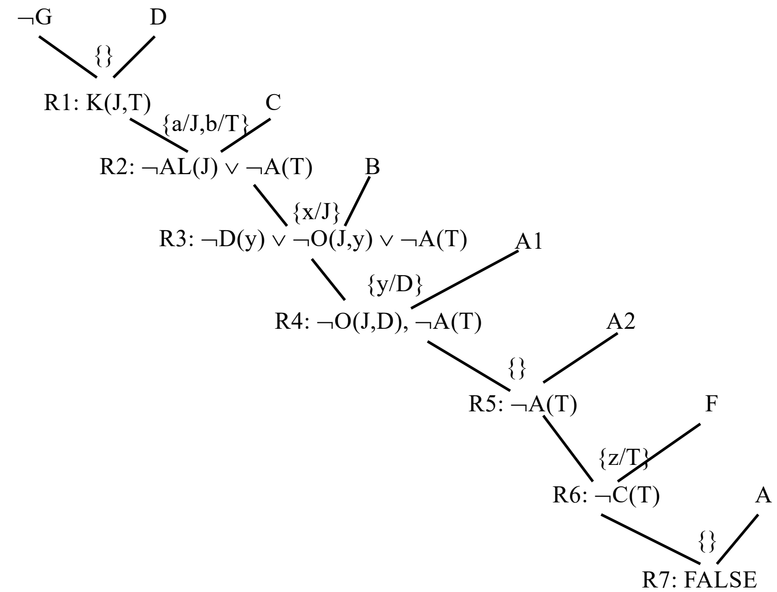
107

The Resolution Refutation Proof

- R1: ¬G, D, {} (Kills(Jack, Tuna))
- R2: R1, C, {a/Jack, b/Tuna} (~AnimalLover(Jack), ~Animal(Tuna))
- R3: R2, B, {x/Jack} (~Dog(y), ~Owns(Jack, y), ~Animal(Tuna))
- R4: R3, A1, {y/D} (~Owns(Jack, D), ~Animal(Tuna))
- R5: R4, A2, {} (~Animal(Tuna))
- R6: R5, F, {z/Tuna} (~Cat(Tuna))
- R7: R6, E, {} FALSE

108

The proof tree



109

Resolution Refutation

- Given a consistent set of axioms KB and goal sentence Q, show that KB $\models$ Q
- **Proof by contradiction:** Add ¬Q to KB and try to prove false.
  - i.e., (KB $\vdash$ Q) $\leftrightarrow$ (KB $\wedge$ ¬Q $\vdash$ False)
- Resolution is **refutation complete:** it can establish that a given sentence Q is entailed by KB, but can't (in general) be used to generate all logical consequences of a set of sentences
  - Also, it cannot be used to prove that Q is not entailed by KB.

110

Efficiency of Resolution Refutation

- Run time can be exponential in the worst case
  - Often much faster
- Factoring: if a new clause contains duplicates of the same symbol, delete the duplicates
  - $P \lor R \lor P \lor T \equiv P \lor R \lor T$
- If a clause contains a symbol and its complement, the clause is a tautology and is useless; it can be thrown away
  - a1: $(\neg A \lor B \lor C)$
  - a2: $(\neg B \lor A)$
  - Resolvent of a1 and a2 is: $B \lor C \lor \neg B$
  - Which is valid, so throw it away

111

Resolution Theorem Proving as Search

- Resolution can be thought of as the **bottom-up construction of a search tree**, where the leaves are the clauses produced by KB and the negation of the goal
- When a pair of clauses generates a new resolvent clause, add a new node to the tree with arcs directed from the resolvent to the two parent clauses
- **Resolution succeeds** when a node containing the **False** clause is produced, becoming the **root node** of the tree
- A strategy is **complete** if its use guarantees that the empty clause (i.e., false) can be derived whenever it is entailed

112

Summary

- Logical agents apply **inference** to a **knowledge base** to derive new information and make decisions
- Basic concepts of logic:
  - **syntax**: formal structure of **sentences**
  - **semantics: truth** of sentences wrt **models**
  - **entailment**: necessary truth of one sentence given another
  - **inference**: deriving sentences from other sentences
  - **soundness**: derivations produce only entailed sentences
  - **completeness**: derivations can produce all entailed sentences
- Resolution is complete for propositional logic.
  Forward and backward chaining are linear-time, complete for Horn clauses
- Propositional logic lacks expressive power

113

Next Time

- Final bits of reasoning via inference
- Knowledge Representation
- Project work – **bring computers**

114