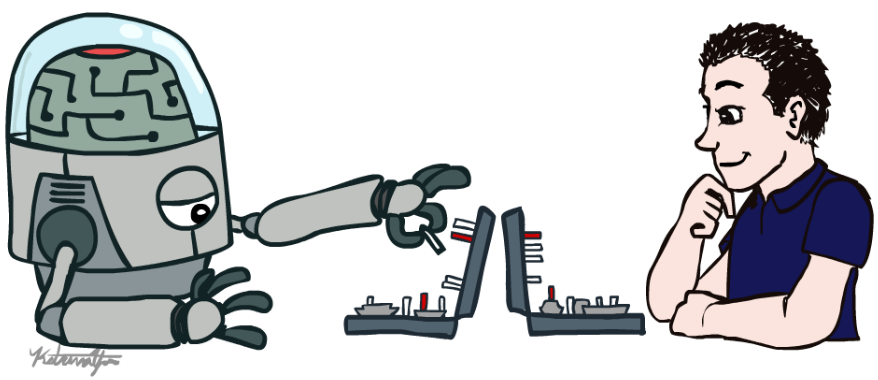
CSE 3521:
Introduction to Artificial Intelligence



[Many slides are adapted from the UC Berkeley. CS188 Intro to AI at UC Berkeley and previous CSE 3521 course at OSU.]

THE OHIO STATE UNIVERSITY

Simple Wumpus Knowledge Base

- For simplicity, only deal with the pits
- Choose vocabulary
  - Let $P_{i,j}$ be True if there is a pit in [*i*,*j*]
  - Let $B_{i,j}$ be True if there is a breeze in [*i*,*j*]
- KB sentences
  - **FACT:** "There is no pit in [1,1]"

    R_1: $\neg P_{1,1}$
  - **RULE:** "There is breeze in adjacent neighbor of pit"

    R_2: $B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$

    R_3: $B_{2,1} \Leftrightarrow (P_{1,1} \vee P_{2,2} \vee P_{3,1})$

Need rule for each square!

Wumpus Environment

- Given knowledge base
- Include percepts as move through environment (online)
- Need to “deduce what to do”
- Derive chains of conclusions that lead to the desired goal
    - Use inference rules

$$\frac{\alpha}{\beta}$$ Inference rule: “α derives β”

Knowing α is true, then β must also be true

Inference Rules for Prop. Logic

- Modus Ponens
  - From implication and premise of implication, can infer conclusion

$$\frac{\alpha \Rightarrow \beta, \quad \alpha}{\beta}$$

Inference Rules for Prop. Logic

- And-Elimination
  - From conjunction, can infer any of the conjuncts

$$\frac{\alpha_1 \land \alpha_2 \land \cdots \land \alpha_n}{\alpha_i}$$

Inference Rules for Prop. Logic

- And-Introduction
  - From list of sentences, can infer their conjunction

$$\frac{\alpha_1, \alpha_2, \ldots, \alpha_n}{\alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_n}$$

Inference Rules for Prop. Logic

- Or-Introduction
  - From sentence, can infer its disjunction with anything else

$$\frac{\alpha_i}{\alpha_1 \lor \alpha_2 \lor \cdots \lor \alpha_n}$$

Inference Rules for Prop. Logic

- Double-Negation Elimination
  - From doubly negated sentence, can infer a positive sentence

$$\frac{\neg\neg\alpha}{\alpha}$$

Inference Rules for Prop. Logic

- Unit Resolution
  - From disjunction, if one of the disjuncts is false, can infer the other is true

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

Inference Rules for Prop. Logic

- <u>Resolution</u>
  - Most difficult because β cannot be both true and false
  - One of the other disjuncts must be true in one of the premises
    - (implication is transitive)

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

$\neg\beta$ **OR** β

$\dfrac{\alpha \vee \beta, \neg\beta}{\alpha}$ **OR** $\dfrac{\neg\beta \vee \gamma, \beta}{\gamma}$

α **OR** γ

α	β	γ	$\alpha \vee \beta$	$\neg\beta \vee \gamma$	$\alpha \vee \gamma$
T	T	T	T	T	T
T	T	F	T	F	T
T	F	T	T	T	T
T	F	F	T	T	T
F	T	T	T	T	T
F	T	F	T	F	T
F	F	T	F	T	F
F	F	F	F	T	F

TASK: Find the Wumpus

Can we infer that the Wumpus is in cell (1,3), given our percepts and environment rules?

A = agent

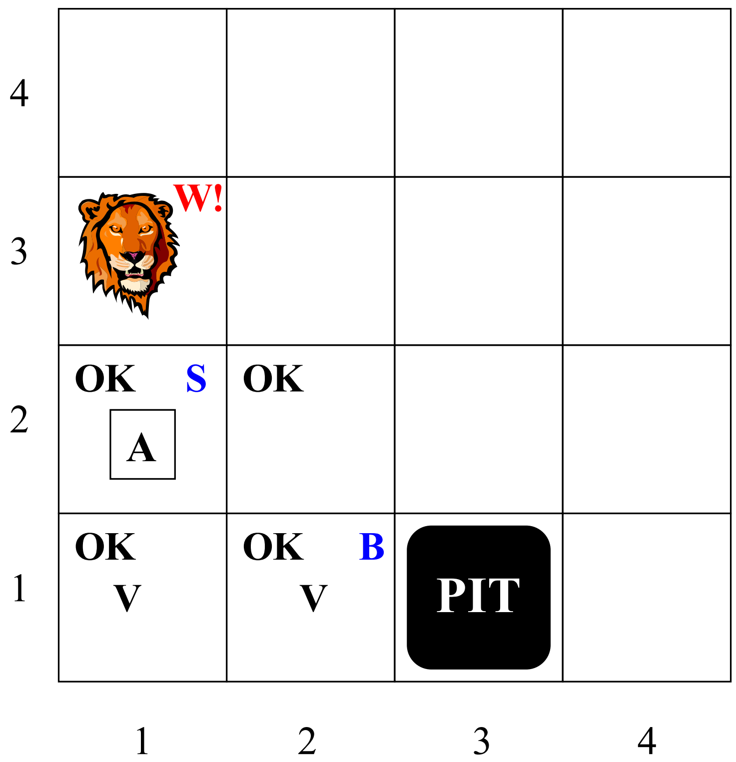
B = breeze

OK = safe square

S = stench

V = visited

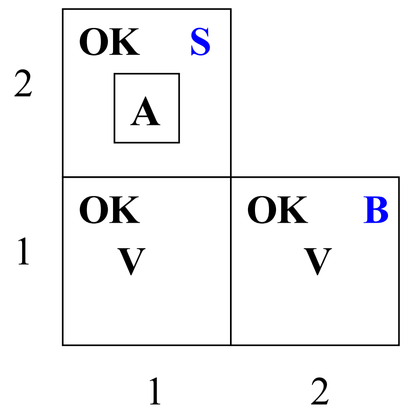
W = wumpus



Wumpus Knowledge Base

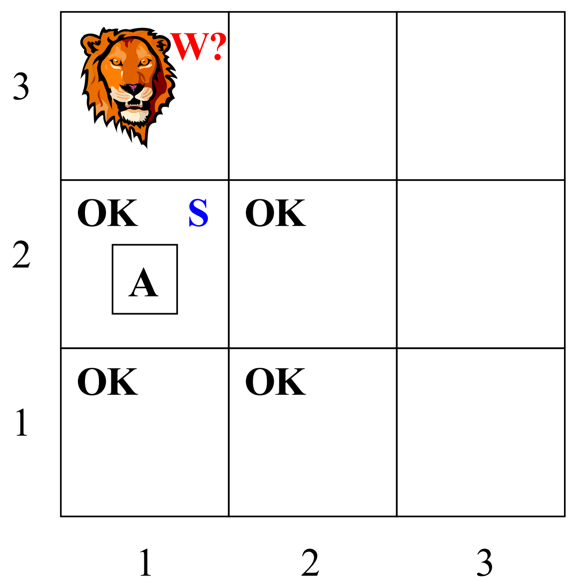
- Percept sentences (facts) "at this point"

$\neg S_{1,1}$	$\neg B_{1,1}$
$\neg S_{2,1}$	$B_{2,1}$
$S_{1,2}$	$\neg B_{1,2}$



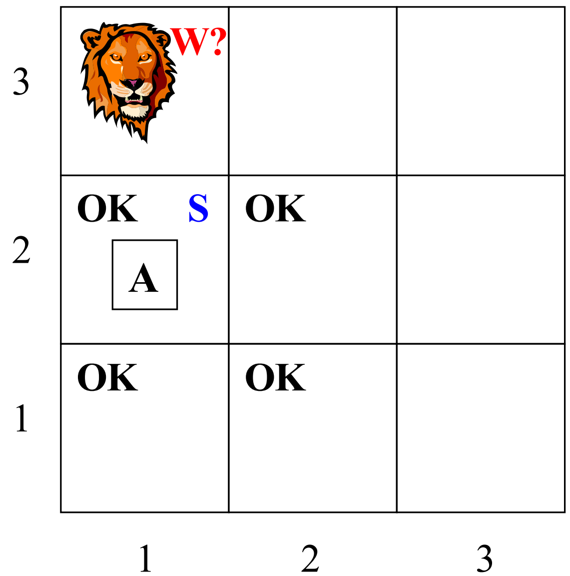
Environment Rules

$$R_1: \quad \neg S_{1,1} \Rightarrow \neg W_{1,1} \wedge \neg W_{1,2} \wedge \neg W_{2,1}$$



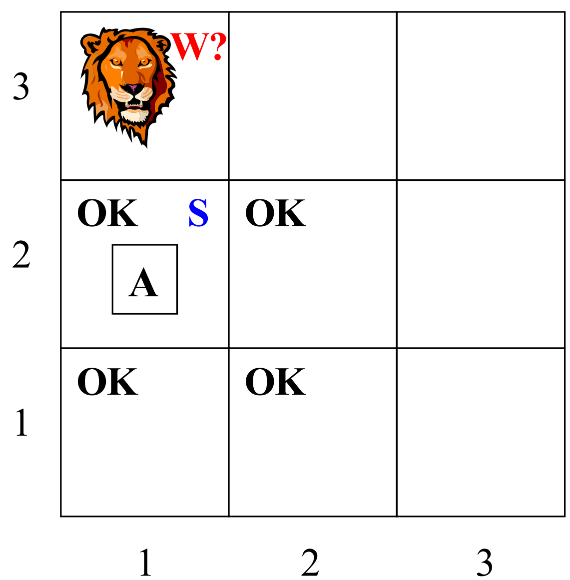
Environment Rules

$$R_2:\ \neg S_{2,1} \Rightarrow \neg W_{1,1} \wedge \neg W_{2,1} \wedge \neg W_{2,2} \wedge \neg W_{3,1}$$



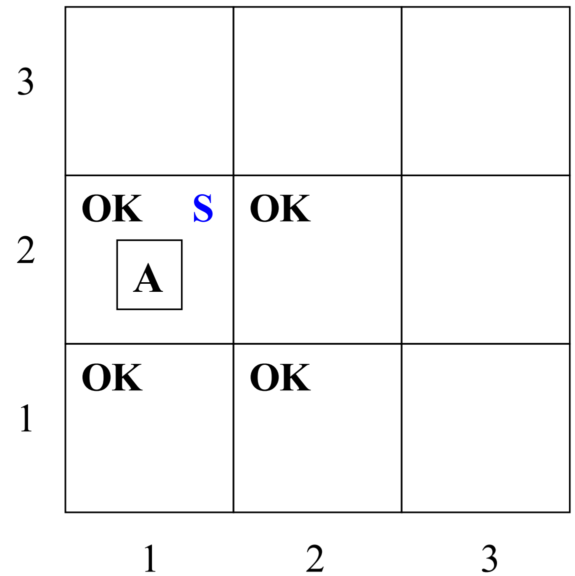
Environment Rules

$$R_3: \quad \neg S_{1,2} \Rightarrow \neg W_{1,1} \wedge \neg W_{1,2} \wedge \neg W_{2,2} \wedge \neg W_{1,3}$$



Environment Rules

R_4: $S_{1,2} \Rightarrow W_{1,3} \vee W_{1,2} \vee W_{2,2} \vee W_{1,1}$



Conclude $W_{1,3}$?

- Does the Wumpus reside in square (1,3) ?
- In other words, can we infer $W_{1,3}$ from our knowledge base?

$$KB \vdash_i W_{1,3}$$

Conclude $w_{1,3}$ (Step #1)

- Modus Ponens $\dfrac{\alpha \Rightarrow \beta,\ \ \alpha}{\beta}$

  R_1: $\neg S_{1,1} \Rightarrow \neg W_{1,1} \wedge \neg W_{1,2} \wedge \neg W_{2,1}$
  Percept: $\neg S_{1,1}$

- Infer

  $\neg W_{1,1} \wedge \neg W_{1,2} \wedge \neg W_{2,1}$

Conclude $w_{1,3}$ (Step #2)

- And-Elimination $\dfrac{\alpha_1 \wedge \alpha_2 \wedge \ldots \wedge \alpha_n}{\alpha_i}$

$$\neg W_{1,1} \wedge \neg W_{1,2} \wedge \neg W_{2,1}$$

- Infer

$$\neg W_{1,1} \quad \neg W_{1,2} \quad \neg W_{2,1}$$

Conclude $w_{1,3}$ (Step #3)

- Modus Ponens $\frac{\alpha \Rightarrow \beta, \quad \alpha}{\beta}$

  R_2: $\neg S_{2,1} \Rightarrow \neg W_{1,1} \wedge \neg W_{2,1} \wedge \neg W_{2,2} \wedge \neg W_{3,1}$
  Percept: $\neg S_{2,1}$

- Infer

  $\neg W_{1,1} \wedge \neg W_{2,1} \wedge \neg W_{2,2} \wedge \neg W_{3,1}$

- And-Elimination $\frac{\alpha_1 \wedge \alpha_2 \wedge \ldots \wedge \alpha_n}{\alpha_i}$

  $\neg W_{1,1} \quad \neg W_{2,1} \quad \neg W_{2,2} \quad \neg W_{3,1}$

Conclude $W_{1,3}$ (Step #4)

- Modus Ponens $\dfrac{\alpha \Rightarrow \beta,\ \ \alpha}{\beta}$

  R_4: $S_{1,2} \Rightarrow W_{1,3} \vee W_{1,2} \vee W_{2,2} \vee W_{1,1}$

  Percept: $S_{1,2}$

- Infer

  $W_{1,3} \vee W_{1,2} \vee W_{2,2} \vee W_{1,1}$

Conclude $W_{1,3}$ (Step #5)

- Unit Resolution $\dfrac{\alpha \vee \beta, \quad \neg\beta}{\alpha}$

  $W_{1,3} \vee W_{1,2} \vee W_{2,2} \vee W_{1,1}$ from Step #4

  $\neg W_{1,1}$ from Step #2

- Infer

  $W_{1,3} \vee W_{1,2} \vee W_{2,2}$

Conclude $w_{1,3}$ (Step #6)

- Unit Resolution $\dfrac{\alpha \vee \beta, \ \ \neg\beta}{\alpha}$

  $W_{1,3} \vee W_{1,2} \vee W_{2,2}$ from Step #5

  $\neg W_{2,2}$ from Step #3

- Infer

  $W_{1,3} \vee W_{1,2}$

Conclude $W_{1,3}$ (Step #7)

- Unit Resolution

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

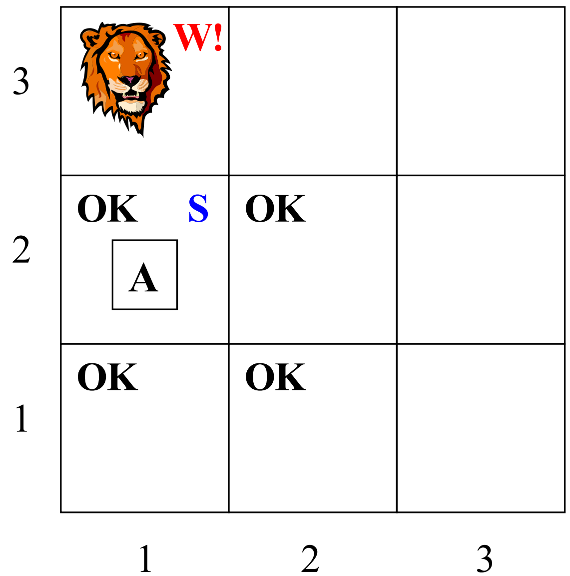
$W_{1,3} \vee W_{1,2}$ from Step #6

$\neg\, W_{1,2}$ from Step #2

- Infer

$W_{1,3} \rightarrow$ The wumpus is in cell 1,3!!!

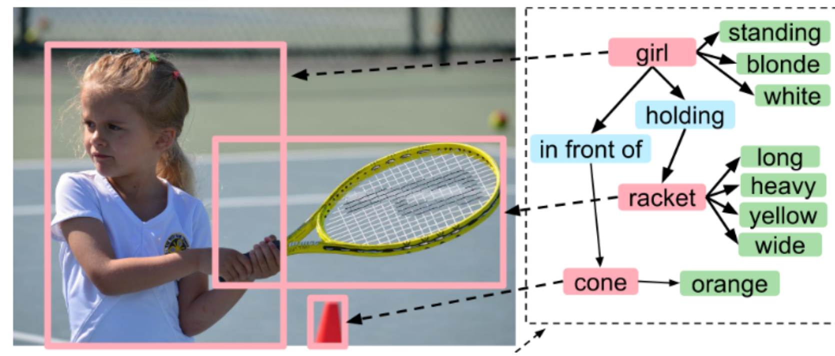
Wumpus in $W_{1,3}$



Propositional Logic

- Propositional logic commits to existence of facts about the world being represented
  - Simple syntax and semantics
- Proof methods
  - Truth table
  - Inference rules
    - Modus Ponens
    - And-Elimination
    - And/Or-Introduction
    - Double-Negation Elimination
    - Unit Resolution
    - Resolution
- Propositional logic quickly becomes impractical

First-Order Logic



- Also called first-order predicate calculus
  - FOL, FOPC
- Makes stronger commitments
  - World consists of objects
    - Things with identities
    - e.g., people, houses, colors, …
  - Objects have properties/relations that distinguish them from other objects
    - e.g., Properties: red, round, square, …
    - e.g., Relations: brother of, bigger than, inside, …
  - Have functional relations
    - Return the object with a certain relation to given "input" object
    - The "inverse" of a (binary) relation
    - e.g., father of, best friend

Examples of Facts as Objects and Properties or Relations

- “Squares neighboring the Wumpus are smelly”
  - Objects
    - Wumpus, squares
  - Property
    - Smelly
  - Relation
    - Neighboring

Syntax of FOL: Basic Elements

- Constant symbols for specific objects

  KingJohn, 2, *OSU*, …
- Variables

  x, *y*, *a*, *b*, …
- Predicate properties (unary) / relations (pairwise or more)

  Smart(), Brother(), Married(), >, …
- Functions (return objects)

  Sqrt() , *LeftTo(), FatherOf()*, …
- Connectives

  $\wedge$ $\vee$ $\neg$ $\Rightarrow$ $\Leftrightarrow$
- Quantifiers

  $\forall$ $\exists$
- Equality

  =

Atomic Sentences

- Collection of terms and relation(s) together to state facts
- Atomic sentence
  - *predicate*$(term_1, \ldots, term_n)$
  - Or $term_1 = term_n$
- Examples

  Brother(Richard, John)

  Married(FatherOf(Richard), MotherOf(John))

Complex Sentences

- Made from atomic sentences using logical connectives

  $\neg S, \quad S_1 \wedge S_2, \quad S_1 \vee S_2, \quad S_1 \Rightarrow S_2, \quad S_1 \Leftrightarrow S_2$

  - Examples:
    - $Older(John, 30) \Rightarrow \neg Younger(John, 30)$
    - $>(1,2) \vee \leq(1,2)$

Quantifiers

- Currently have logic that allows objects
- Now want to express properties of entire <u>collections of objects</u>
  - Rather than enumerate the objects by name
- Two standard quantifiers
  - Universal $\forall$
  - Existential $\exists$

Universal Qualification

- "For all …" (typically use implication $\Rightarrow$)
  - Allows for "rules" to be constructed
- $\forall$ *<variables> <sentence>*
  - Everyone at OSU is smart
    - $\forall x \ At(x, OSU) \Rightarrow Smart(x)$
- $\forall x \ P$ is equivalent to <u>conjunction</u> of all <u>instantiations</u> of *P*
    - $(At(John, OSU) \Rightarrow Smart(John))$
      $\wedge \ (At(Bob, OSU) \Rightarrow Smart(Bob))$
      $\wedge \ (At(Mary, OSU) \Rightarrow Smart(Mary)) \wedge \ \ldots$

Existential Quantification

- "There exists …" (typically use conjunction $\wedge$)
  - Makes a statement about <u>some</u> object (not all)
- $\exists$ *<variables> <sentences>*
  - Someone at OSU is smart
    - $\exists x\ At(x, OSU) \wedge Smart(x)$
- $\exists\ x\ P$ is equivalent to <u>disjunction</u> of all <u>instantiations</u> of P
    - $(At(John, OSU) \wedge Smart(John))$
      $\vee\ (At(Bob, OSU) \wedge Smart(Bob))$
      $(At(Mary, OSU) \wedge Smart(Mary))\ \vee\ \ldots$
- Uniqueness quantifier
  - $\exists!\ x$ says a <u>unique</u> object exists (i.e. there is exactly one)

Properties of Quantifiers

- Quantifier duality: Each can be expressed using the other
  - $\forall x$ *Person(x)* $\Rightarrow$ *Likes(x, IceCream)* "Everybody likes ice cream"
  - $\neg\exists x$ *Person(x)* $\wedge$ $\neg$*Likes(x, IceCream)* "Not exist anyone who does not like ice cream"
  - $\exists x$ *Person(x)* $\wedge$ *Likes(x, Broccoli)* "Someone likes broccoli"
  - $\neg\forall x$ *Person(x)* $\Rightarrow$ $\neg$*Likes(x, Broccoli)* "Not the case that everyone does not like broccoli"

Properties of Quantifiers

- Important relations
  - $\exists x \ P(x) = \neg \forall x \ \neg P(x)$
  - $\forall x \ P(x) = \neg \exists x \ \neg P(x)$
  - $P(x) \Rightarrow Q(x)$ is same as $\neg P(x) \vee Q(x)$
  - $\neg (P(x) \wedge Q(x))$ is same as $\neg P(x) \vee \neg Q(x)$

Proof: Check the Truth Table

P	Q	$P \Rightarrow Q$
False	False	**TRUE**
False	True	**TRUE**
True	False	**FALSE**
True	True	**TRUE**

P	Q	$\neg P$	$\neg P \vee Q$
False	False	True	**TRUE**
False	True	True	**TRUE**
True	False	False	**FALSE**
True	True	False	**TRUE**

$P(x) \Rightarrow Q(x)$

is same as

$\neg P(x) \vee Q(x)$

Proof : Check the Truth Table

P	Q	$P \wedge Q$	$\neg(P \wedge Q)$
False	False	False	**TRUE**
False	True	False	**TRUE**
True	False	False	**TRUE**
True	True	True	**FALSE**

P	Q	$\neg P$	$\neg Q$	$\neg P \vee \neg Q$
False	False	True	True	**TRUE**
False	True	True	False	**TRUE**
True	False	False	True	**TRUE**
True	True	False	False	**FALSE**

$\neg(P(x) \wedge Q(x))$

is same as

$\neg P(x) \vee \neg Q(x)$

Proof : Check the Truth Table

P	Q	$P \vee Q$	$\neg(P \vee Q)$
False	False	False	**TRUE**
False	True	True	**FALSE**
True	False	True	**FALSE**
True	True	True	**FALSE**

P	Q	$\neg P$	$\neg Q$	$\neg P \wedge \neg Q$
False	False	True	True	**TRUE**
False	True	True	False	**FALSE**
True	False	False	True	**FALSE**
True	True	False	False	**FALSE**

$\neg(P(x) \vee Q(x))$

is same as

$\neg P(x) \wedge \neg Q(x)$

Conversion Example

- $\forall x \; Person(x) \Rightarrow Likes(x, IceCream)$
  [use: $\forall x\, P(x) = \neg\exists x\, \neg P(x)$]
-
- $\neg\exists x\, \neg(Person(x) \Rightarrow Likes(x, IceCream))$
  [use: $P(x) \Rightarrow Q(x)$ is same as $\neg P(x) \vee Q(x)$]
- $\neg\exists x\, \neg(\neg Person(x) \vee Likes(x, IceCream))$
  [distribute negatives]
- $\neg\exists x\, Person(x) \wedge \neg Likes(x, IceCream)$
- $\neg\mathrm{x} == (\neg\mathrm{x})$