CSC242: Introduction to Artificial Intelligence

Lecture 2.4

Please put away all electronic devices

Announcements

- Unit 2 Exam: One week from today
- Unit 2 Project due that night 1159PM
  - Don't wait to be finished

Factored Representation

- Splits a state into variables (or attributes) that can have values
- Factored states can be more or less similar (unlike atomic states)
- Can also represent uncertainty (don't know value of some attribute)

Constraint Satisfaction Problem (CSP)

- X: Set of variables $\{ X_1, ..., X_n \}$
- D: Set of domains $\{ D_1, ..., D_n \}$
  - Each D_i : set of values $\{ v_1, ..., v_k \}$
- C: Set of constraints $\{ C_1, ..., C_m \}$
- Solution: Assign to each X_i a value from D_i such that all the C_i are satisfied

Propositional Logic

- Programming language for knowledge
- Factored representation of state
  - Propositions, connectives, sentences
- Model (possible world) = assignment of true or false to propositions
- Entailment ("follows from"): $\alpha \models \beta$
  - Every model of α is a model of β

Propositional Inference

- Computing whether $\alpha \models \beta$
- Model Checking
  - Intractable (but see AIMA 7.6)
- Inference rules: Soundness, Completeness
- Derivation: $\alpha \vdash \beta$
  - Searching for proofs is an alternative to enumerating models
    - May be faster in practice
- Resolution is a sound and complete inference rule
  - Works on clauses (CNF), requires refutation proof

Inference Using Resolution

- Convert sentences (KB) to CNF (set of clauses)
- Apply resolution inference rule to pairs of clauses with complementary literals
- Add resulting clause to set of clauses
- Until...

Proof by Contradiction

- $\alpha \models \beta$ if and only if $(\alpha \wedge \neg\beta)$ is unsatisfiable
- If negation of goal is inconsistent with our knowledge
- Then the goal itself is entailed by our knowledge

Resolution Refutation

- Convert $(KB \wedge \neg\beta)$ to CNF
- Apply resolution rule until:
  - No new clauses can be added
    - $KB \not\models \beta$
  - Two clauses resolve to yield the <u>empty</u> clause (contradiction)

    $KB \models \beta$

1. $Flu \Rightarrow Sneezing$
2. $Cold \Rightarrow Congested$
3. $Congested \Rightarrow Coughing$

1. $Flu \Rightarrow Sneezing$
2. $Cold \Rightarrow Congested$
3. $Congested \Rightarrow Coughing$
4. $Flu \vee Cold$
5. $\neg Sneezing$

1. $Flu \Rightarrow Sneezing$

2. $Cold \Rightarrow Congested$

3. $Congested \Rightarrow Coughing$

4. $Flu \lor Cold$

5. $\neg Sneezing$

To prove: $Coughing$

1. $Flu \Rightarrow Sneezing$ | $\neg Flu \vee Sneezing$

2. $Cold \Rightarrow Congested$ | $\neg Cold \vee Congested$

3. $Congested \Rightarrow Coughing$ | $\neg Congested \vee Coughing$

4. $Flu \vee Cold$ | $Flu \vee Cold$

5. $\neg Sneezing$ | $\neg Sneezing$

To prove: $Coughing$

1. $\neg Flu \vee Sneezing$
2. $\neg Cold \vee Congested$
3. $\neg Congested \vee Coughing$
4. $Flu \vee Cold$
5. $\neg Sneezing$

To prove: $Coughing$

1. $\neg Flu \lor Sneezing$
2. $\neg Cold \lor Congested$
3. $\neg Congested \lor Coughing$
4. $Flu \lor Cold$
5. $\neg Sneezing$
6. $\neg Coughing$

1. $\neg Flu \vee Sneezing$
2. $\neg Cold \vee Congested$
3. $\neg Congested \vee Coughing$
4. $Flu \vee Cold$
5. $\neg Sneezing$
6. $\neg Coughing$

1 & 4: 7. $Cold \vee Sneezing$

1. $\neg Flu \lor Sneezing$
2. $\neg Cold \lor Congested$
3. $\neg Congested \lor Coughing$
4. $Flu \lor Cold$
5. $\neg Sneezing$
6. $\neg Coughing$

1 & 4: 7. $Cold \lor Sneezing$

7 & 5: 8. $Cold$

1. $\neg Flu \vee Sneezing$
2. $\neg Cold \vee Congested$
3. $\neg Congested \vee Coughing$
4. $Flu \vee Cold$
5. $\neg Sneezing$
6. $\neg Coughing$

1 & 4: 7. $Cold \vee Sneezing$

7 & 5: 8. $Cold$

2 & 3: 9. $\neg Cold \vee Coughing$

9 & 6: 10. $\neg Cold$

1. $\neg Flu \vee Sneezing$
2. $\neg Cold \vee Congested$
3. $\neg Congested \vee Coughing$
4. $Flu \vee Cold$
5. $\neg Sneezing$
6. $\neg Coughing$

1 & 4: 7. $Cold \vee Sneezing$

7 & 5: 8. $Cold$

2 & 3: 9. $\neg Cold \vee Coughing$

9 & 6: 10. $\neg Cold$

8 & 10: $\varnothing$

1. $\neg Flu \lor Sneezing$
2. $\neg Cold \lor Congested$
3. $\neg Congested \lor Coughing$
4. $Flu \lor Cold$
5. $\neg Sneezing$
6. $\neg Coughing$

Unsatisfiable

1 & 4: 7. $Cold \lor Sneezing$

7 & 5: 8. $Cold$

2 & 3: 9. $\neg Cold \lor Coughing$

9 & 6: 10. $\neg Cold$

8 & 10: $\varnothing$

1. $\neg Flu \vee Sneezing$
2. $\neg Cold \vee Congested$
3. $\neg Congested \vee Coughing$
4. $Flu \vee Cold$
5. $\neg Sneezing$
6. $\neg Coughing$

Satisfiable

Unsatisfiable

1 & 4: 7. $Cold \vee Sneezing$

7 & 5: 8. $Cold$

2 & 3: 9. $\neg Cold \vee Coughing$

9 & 6: 10. $\neg Cold$

8 & 10: $\varnothing$

1. $\neg Flu \vee Sneezing$
2. $\neg Cold \vee Congested$
3. $\neg Congested \vee Coughing$
4. $Flu \vee Cold$
5. $\neg Sneezing$
6. $\neg Coughing$

Satisfiable

Unsatisfiable

1 & 4: 7. $Cold \vee Sneezing$

7 & 5: 8. $Cold$

2 & 3: 9. $\neg Cold \vee Coughing$

9 & 6: 10. $\neg Cold$

8 & 10: $\varnothing$

1. $\neg Flu \lor Sneezing$
2. $\neg Cold \lor Congested$
3. $\neg Congested \lor Coughing$
4. $Flu \lor Cold$
5. $\neg Sneezing$
6. $\neg Coughing$

Satisfiable

Unsatisfiable

Proven: $Coughing$

1 & 4: 7. $Cold \lor Sneezing$

7 & 5: 8. $Cold$

2 & 3: 9. $\neg Cold \lor Coughing$

9 & 6: 10. $\neg Cold$

8 & 10: $\varnothing$

Resolution Refutation

- Convert $(KB \wedge \neg\beta)$ to CNF
- Apply resolution rule until:
    - No new clauses can be added
        - $KB \not\models \beta$
    - Two clauses resolve to yield the <u>empty</u> clause (contradiction)

        $KB \models \beta$

PL Pros

- Declarative: based on a truth relation between sentences and possible worlds
- Expressive: can represent partial information (e.g., disjunction, negation)
- Compositional: the meaning of a sentence is a function of the meanings of its parts

PL Cons

- Model checking takes exponential time
  - Theorem proving may help
- Lacks the expressive power to concisely describe complex environments (many objects, relationships between them)

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

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

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

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

$$\ldots$$

$$B_{4,4} \Leftrightarrow (P_{3,4} \vee P_{4,3})$$

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

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

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

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

$$\ldots$$

$$B_{4,4} \Leftrightarrow (P_{3,4} \vee P_{4,3})$$

"Rooms adjacent to pits are breezy"

- Rooms adjacent to pits are breezy
- Socrates is a person
  All people are mortal
- Anybody's grandmother is either their mother's or their father's mother

Logic 2.0

- Define a language based on propositional logic that will allow us to say all these things
- Define entailment ("follows from")
- Find inference rules that will allow us to compute the consequences of our knowledge (entailments)

- Rooms adjacent to pits are breezy
- Socrates is a person
  All people are mortal
- Anybody's grandmother is either their mother's or their father's mother

- Rooms adjacent to pits are breezy
- Socrates is a person
  All people are mortal
- Anybody's grandmother is either their mother's or their father's mother

- Rooms adjacent to pits are breezy
- Socrates is a person
  All people are mortal
- Anybody's grandmother is either their mother's or their father's mother

- Rooms adjacent to pits are breezy
- Socrates is a person
  All people are mortal
- Anybody's grandmother is either their mother's or their father's mother

- Rooms adjacent to pits are breezy
- Socrates is a person
  All people are mortal
- Anybody's grandmother is either their mother's or their father's mother

Ontology

ontology |änˈtäləjē|
noun
1 the branch of metaphysics dealing with the nature of being.

2 a set of concepts and categories in a subject area or domain that shows their properties and the relations between them

ORIGIN early 18th cent.: from modern Latin ***ontologia***, from Greek ***ōn***, ***ont-*** ***'being'*** + **-logy**.

Ontology

- **Objects**: people, houses, numbers, theories, Socrates, colors, wars, ...
- **Relations**:
  - Unary (Properties): breezy, mortal, red, round, bogus, prime, ...
  - n-ary: brother of, bigger than, inside, part of, has color, occurred after, owns, above
- **Functions**: "single-valued" relations: mother of, father of, best friend, one more than, ...

“Socrates is a person.”

- Objects: Socrates
- Property (unary relation): being a person

"Rooms adjacent to the wumpus are smelly."

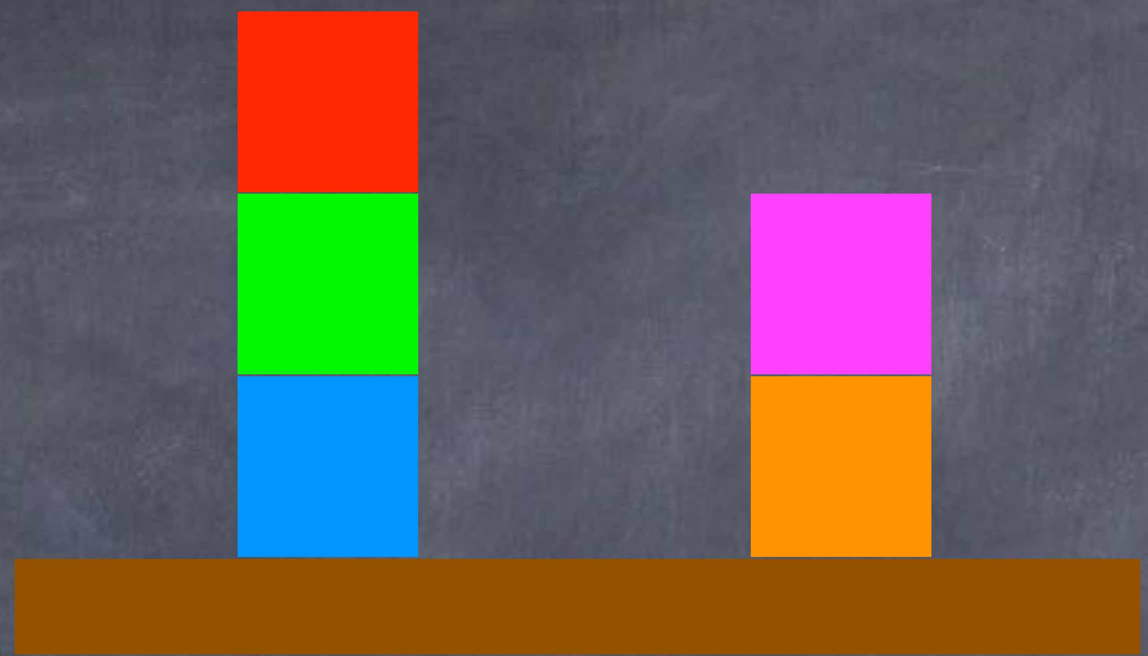
- Objects: wumpus, rooms
- Property (unary relation): smelly
- Relation (binary): adjacent to

"One plus two equals three."

- Objects: one, two, three
- Relation (binary): equals
- Function: plus

"Evil King John ruled England in 1200."

- Objects: John, England, 1200
- Properties (unary relations): evil, king
- Relation (binary): rules



- Objects: ■, ■, ■, ■, ■
- Relations: being on, being above, being clear, being on the table
- Functions: “the block on top of me”

Ontology (Domain of Discourse, Conceptualization)

- **Objects**: people, houses, numbers, theories, Socrates, colors, wars, ...
- **Relations**:
  - Unary (Properties): breezy, mortal, red, round, bogus, prime, ...
  - n-ary: brother of, bigger than, inside, part of, has color, occurred after, owns, above
- **Functions**: "single-valued" relations: mother of, father of, best friend, one more than, ...

A Programming Language for Knowledge

- Syntax:
  - What counts as a well-formed statement, formula, sentence, or program
- Semantics:
  - What these statements, formulas, sentences, or programs mean

Constant Symbols

- Symbols denoting objects in the world
- *Socrates, George, Fido, Dogbert, ...*

denote |diˈnōt|

verb [trans.]
be a sign of; indicate *: this mark denotes purity and quality.*
• (often **be denoted**) stand as a name or symbol for *:*
the level of output per firm, denoted by X.

Relation (Predicate) Symbols

- Symbols denoting relations
- $Mortal(\cdot)$, $Smelly(\cdot)$, $Breezy(\cdot)$, $On(\cdot,\cdot)$, $Above(\cdot,\cdot)$, $Equals(\cdot,\cdot)$ a.k.a. "$\cdot = \cdot$", …
- Arity: number of arguments

Function Symbols

- Symbols denoting functions
- $mother(\cdot)$, $father(\cdot)$, $oneMoreThan(\cdot)$, $hat(\cdot)$, $plus(\cdot,\cdot)$ a.k.a. "$\cdot + \cdot$", …
- Arity: number of arguments

Symbols

- Constant symbols: $Socrates$, $George$
- Relation symbols: $Mortal(\cdot)$, $Above(\cdot,\cdot)$
- Function symbols: $mother(\cdot)$, $plus(\cdot,\cdot)$

Term

- A logical expression that denotes (refers to) an object
- Constant symbol; or
- Function symbol and tuple of terms of appropriate arity

$Socrates$ $\qquad\qquad$ $mother(George)$

$plus(1,2)$ a.k.a. "$1+2$" $\qquad\qquad$ $mother(father(George))$

Atomic Sentence

- States a fact
- Predicate (relation) symbol and tuple of terms of appropriate arity

$$Mortal(Socrates)$$

Atomic Sentence

- States a fact
- Predicate (relation) symbol and tuple of terms of appropriate arity

$$Mortal(Socrates)$$

$$On(A,\ B)$$

$$Brother(Richard,\ John)$$

$$Married(father(Richard),\ mother(John))$$

Connectives

- Connect sentences into larger sentences that can also be true or false
- Negation (not): $\neg$
- Conjunction (and): $\wedge$
- Disjunction (or): $\vee$
- Implication (if-then): $\Rightarrow$
- Biconditional (iff): $\Leftrightarrow$

Connectives

$\neg On(A,\ B)$

$King(Richard) \lor King(John)$

$\neg King(Richard) \Rightarrow King(John)$

Logic 2.0 (Syntax)

- Constant symbols
- Predicate (relation) symbols & arity
- Function symbols & arity
- Terms
- Atomic sentences
- Complex sentences (using connectives)

Predicate Logic

- Constant symbols
- Predicate (relation) symbols & arity
- Function symbols & arity
- Terms
- Atomic sentences
- Complex sentences (using connectives)

First-Order Predicate Logic

- Constant symbols
- Predicate (relation) symbols & arity
- Function symbols & arity
- Terms
- Atomic sentences
- Complex sentences (using connectives)

Propositional Logic
Possible World

- Assignment of true or false to all the atomic propositions
- A possible world satisfies a sentence if it makes the sentence true
  - "A model of the sentence"

Ontology

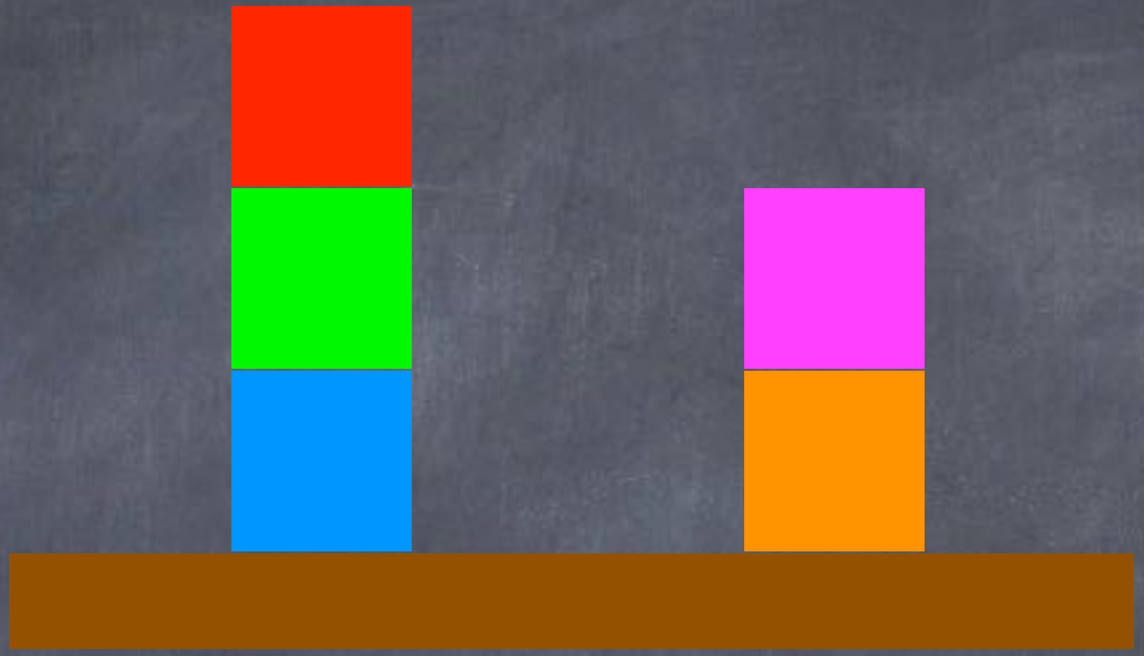
- **Objects**: people, houses, numbers, theories, Socrates, colors, wars, ...
- **Relations**:
  - Unary (Properties): breezy, mortal, red, round, bogus, prime, ...
  - n-ary: brother of, bigger than, inside, part of, has color, occurred after, owns, above
- **Functions**: "single-valued" relations: mother of, father of, best friend, one more than, ...

Interpretation

Language Elements		Domain Elements
Constant Symbols	$\longrightarrow$	Objects
Predicate Symbols	$\longrightarrow$	Relations
Function Symbols	$\longrightarrow$	Functions

Interpretation

Language Elements		Interpretation I Objects: Ω_I
Constant Symbols	σ	$I(\sigma) \in \Omega_I$
Predicate Symbols	π_n	$I(\pi_n) \subseteq \Omega_I^n$
Function Symbols	ϕ_n	$I(\phi_n) : \Omega_I^n \to \Omega_I$



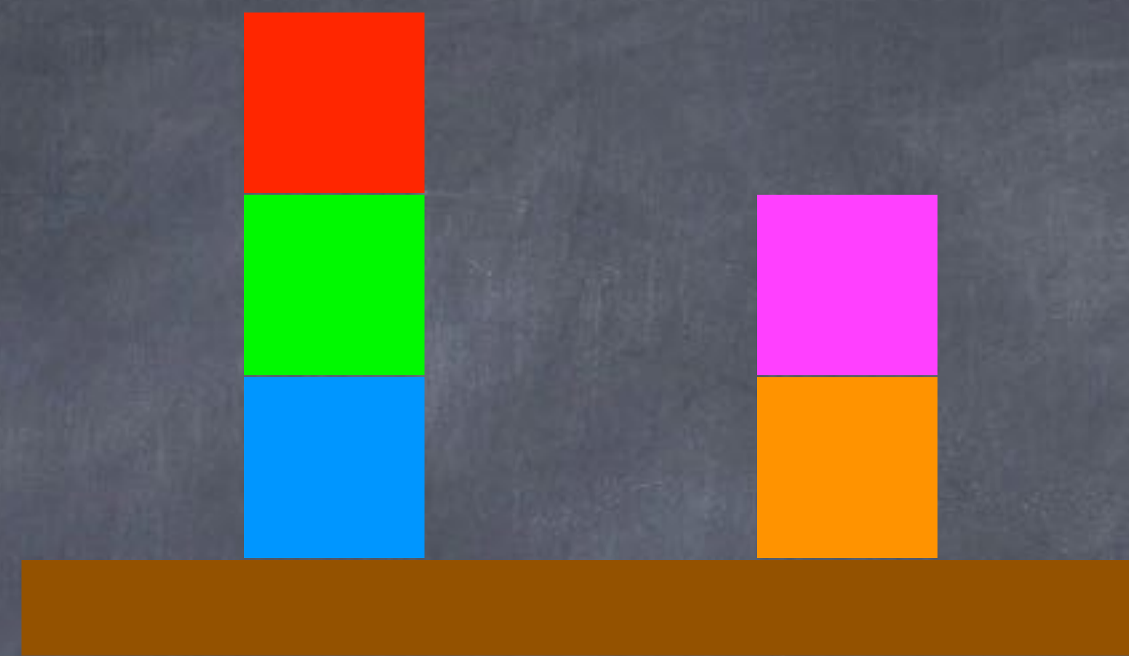
Constant Symbols: A, B, C, D, E

Predicate Symbols: On(·,·), Above(·,·),

OnTable(·), Clear(·)

Function Symbols: Hat(·)

$$\Omega_I = \{\ \color{red}\blacksquare\color{white},\ \color{green}\blacksquare\color{white},\ \color{blue}\blacksquare\color{white},\ \color{magenta}\blacksquare\color{white},\ \color{orange}\blacksquare\ \color{white}\}$$



$$I(\mathtt{A}) = \color{red}\blacksquare$$

$$I(\mathtt{B}) = \color{green}\blacksquare$$

$$I(\mathtt{C}) = \color{blue}\blacksquare$$

$$I(\mathtt{D}) = \color{magenta}\blacksquare$$

$$I(\mathtt{E}) = \color{orange}\blacksquare$$

$$I(\mathtt{On}) = \{\ \langle \color{red}\blacksquare\color{white}, \color{green}\blacksquare\color{white} \rangle,\ \langle \color{green}\blacksquare\color{white}, \color{blue}\blacksquare\color{white} \rangle,\ \langle \color{magenta}\blacksquare\color{white}, \color{orange}\blacksquare\color{white} \rangle\ \}$$

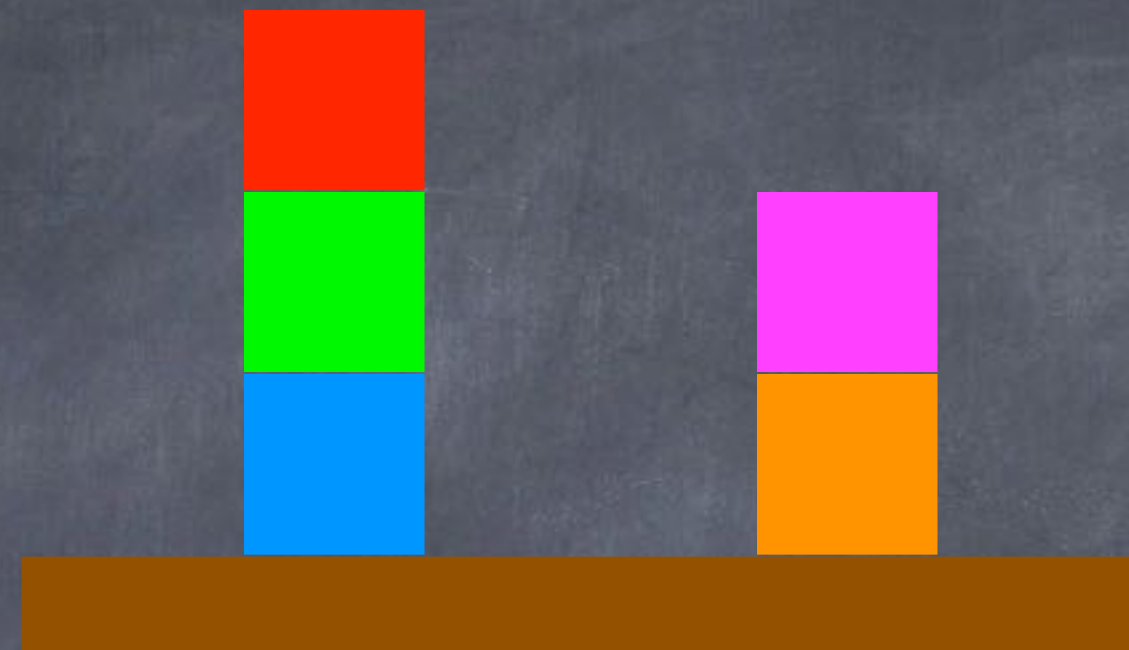
$$I(\mathtt{Above}) = \{\ \langle \color{red}\blacksquare\color{white}, \color{green}\blacksquare\color{white} \rangle,\ \langle \color{green}\blacksquare\color{white}, \color{blue}\blacksquare\color{white} \rangle,\ \langle \color{red}\blacksquare\color{white}, \color{blue}\blacksquare\color{white} \rangle,\ \langle \color{magenta}\blacksquare\color{white}, \color{orange}\blacksquare\color{white} \rangle\ \}$$

$$I(\mathtt{OnTable}) = \{\ \langle \color{blue}\blacksquare\color{white} \rangle,\ \langle \color{orange}\blacksquare\color{white} \rangle\ \}$$

$$I(\mathtt{Clear}) = \{\ \langle \color{red}\blacksquare\color{white} \rangle,\ \langle \color{magenta}\blacksquare\color{white} \rangle\ \}$$

$$I(\mathtt{Hat}) = \{\ \langle \color{green}\blacksquare\color{white} \rangle \rightarrow \color{red}\blacksquare\color{white},\ \langle \color{blue}\blacksquare\color{white} \rangle \rightarrow \color{green}\blacksquare\color{white},\ \langle \color{orange}\blacksquare\color{white} \rangle \rightarrow \color{magenta}\blacksquare\color{white}\ \}$$

$$\Omega_I = \{ {\color{red}\blacksquare}, {\color{green}\blacksquare}, {\color{blue}\blacksquare}, {\color{magenta}\blacksquare}, {\color{orange}\blacksquare} \}$$



$$I(\mathtt{A}) = {\color{red}\blacksquare}$$

$$I(\mathtt{B}) = {\color{green}\blacksquare}$$

$$I(\mathtt{C}) = {\color{blue}\blacksquare}$$

$$I(\mathtt{D}) = {\color{magenta}\blacksquare}$$

$$I(\mathtt{E}) = {\color{orange}\blacksquare}$$

$$I(\mathtt{On}) = \{ \langle {\color{green}\blacksquare}, {\color{red}\blacksquare} \rangle, \langle {\color{blue}\blacksquare}, {\color{green}\blacksquare} \rangle, \langle {\color{orange}\blacksquare}, {\color{magenta}\blacksquare} \rangle \}$$

$$I(\mathtt{Above}) = \{ \langle {\color{green}\blacksquare}, {\color{red}\blacksquare} \rangle, \langle {\color{blue}\blacksquare}, {\color{green}\blacksquare} \rangle, \langle {\color{blue}\blacksquare}, {\color{red}\blacksquare} \rangle, \langle {\color{orange}\blacksquare}, {\color{magenta}\blacksquare} \rangle \}$$

$$I(\mathtt{OnTable}) = \{ \langle {\color{red}\blacksquare} \rangle, \langle {\color{magenta}\blacksquare} \rangle \}$$

$$I(\mathtt{Clear}) = \{ \langle {\color{blue}\blacksquare} \rangle, \langle {\color{orange}\blacksquare} \rangle \}$$

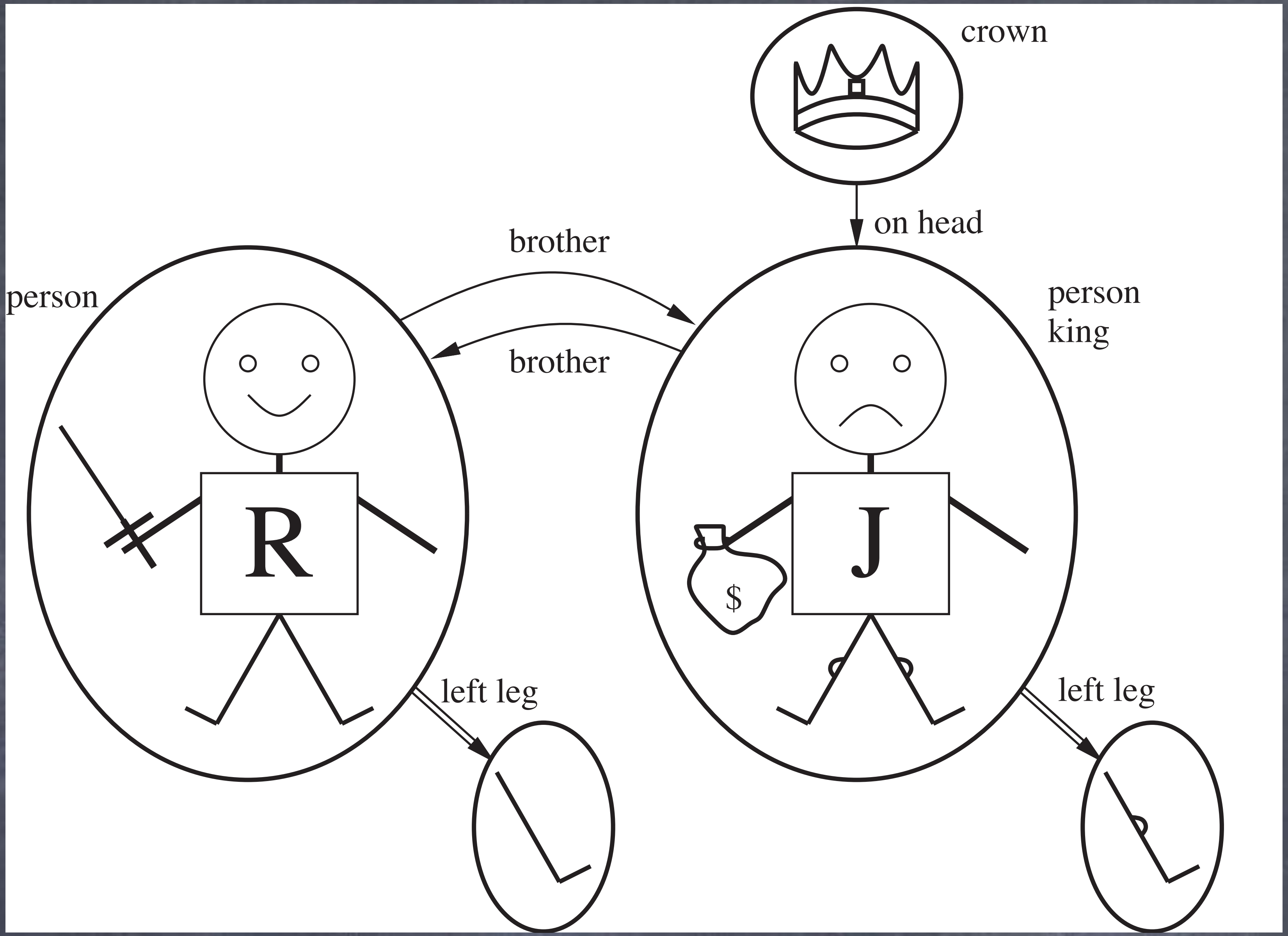
$$I(\mathtt{Hat}) = \{ \langle {\color{green}\blacksquare} \rangle \rightarrow {\color{red}\blacksquare}, \langle {\color{blue}\blacksquare} \rangle \rightarrow {\color{green}\blacksquare}, \langle {\color{orange}\blacksquare} \rangle \rightarrow {\color{magenta}\blacksquare} \}$$

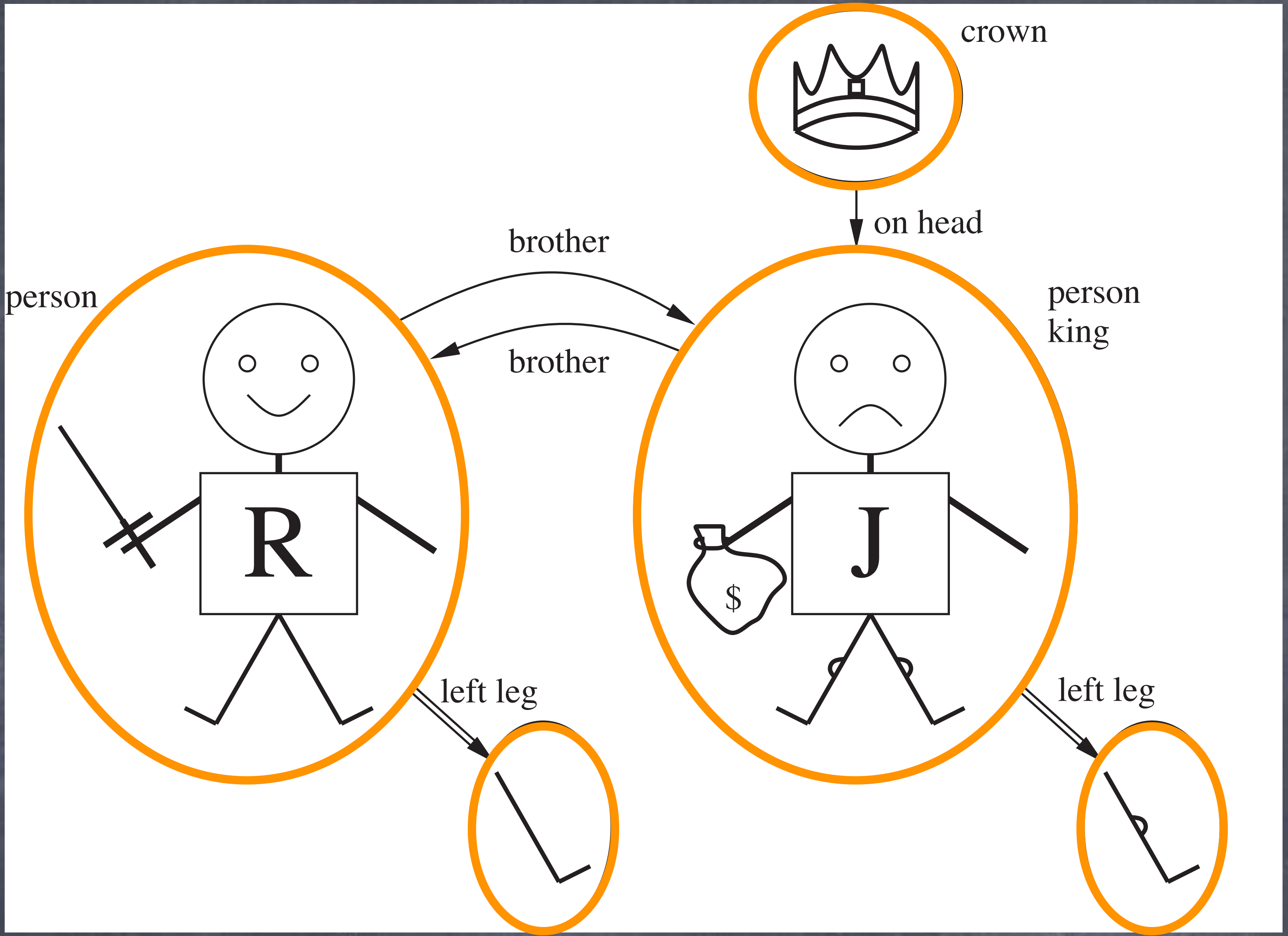
Richard

(1157-1199)

John

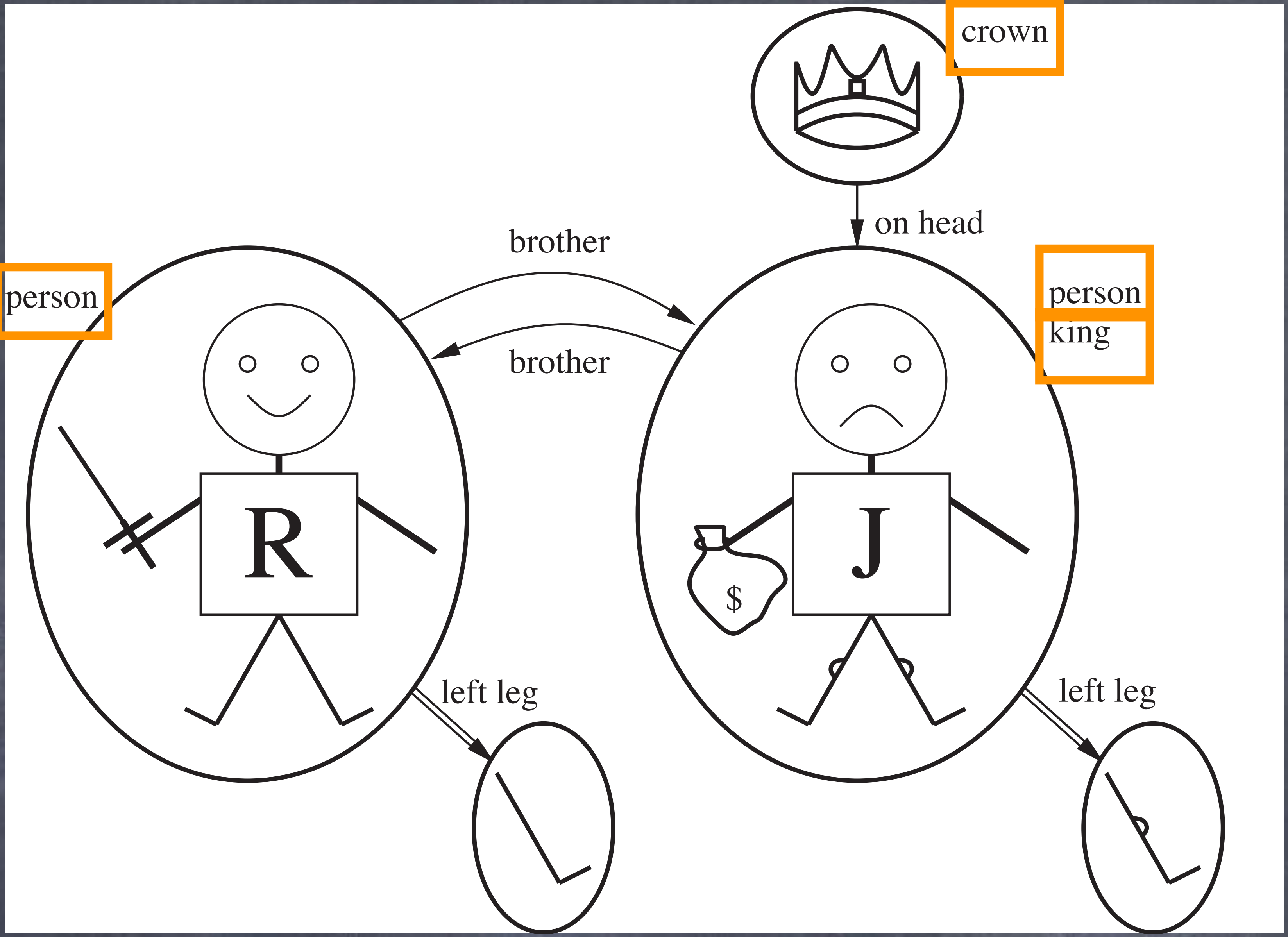
(1166-1216)


crown
on head
brother
brother
person
person
king
R
J
$
left leg
left leg



Objects (Ω_I):

Richard, John, left leg 1, left leg 2, the crown

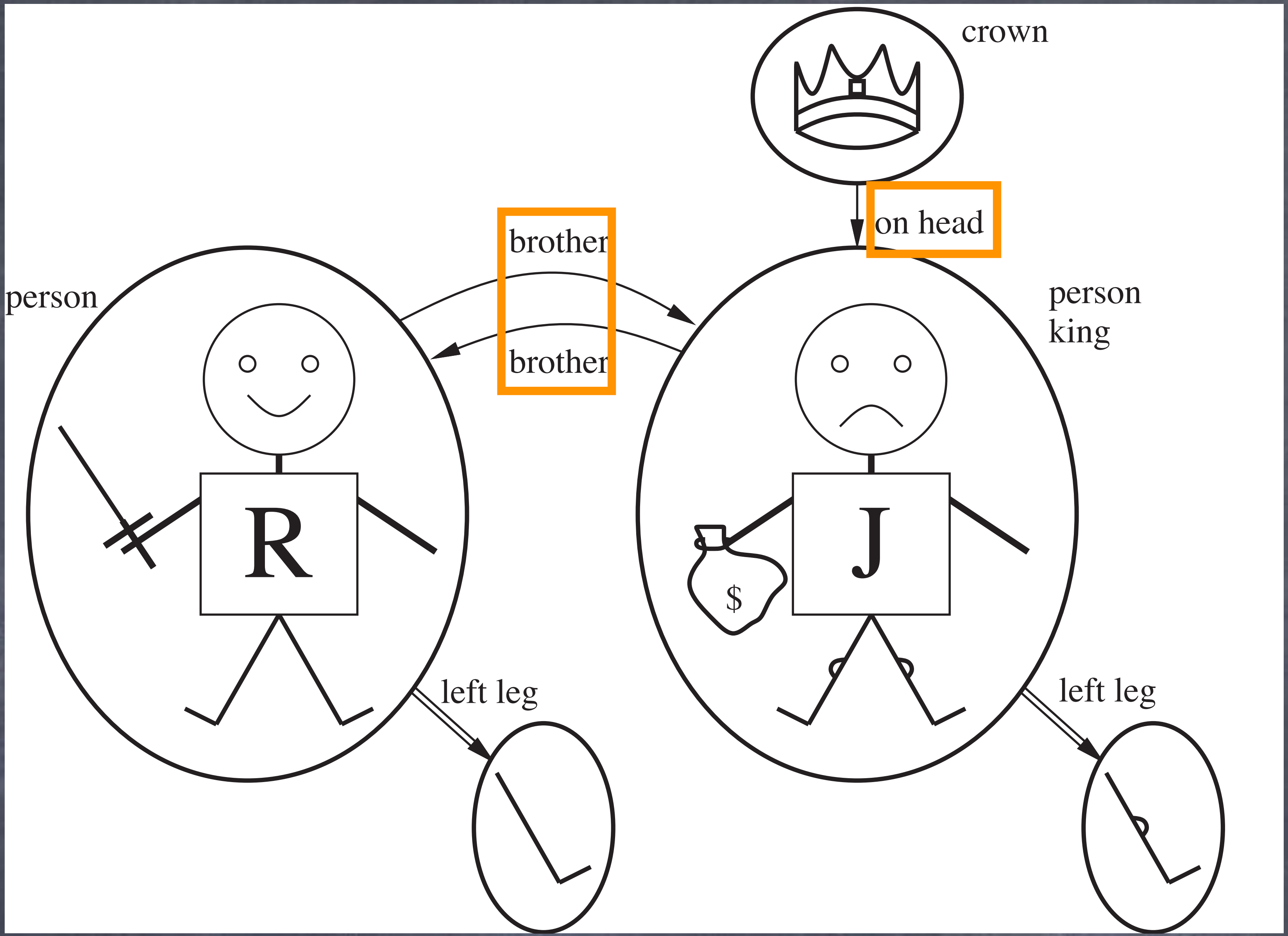


Unary Relations (Properties):

being a person

being a crown

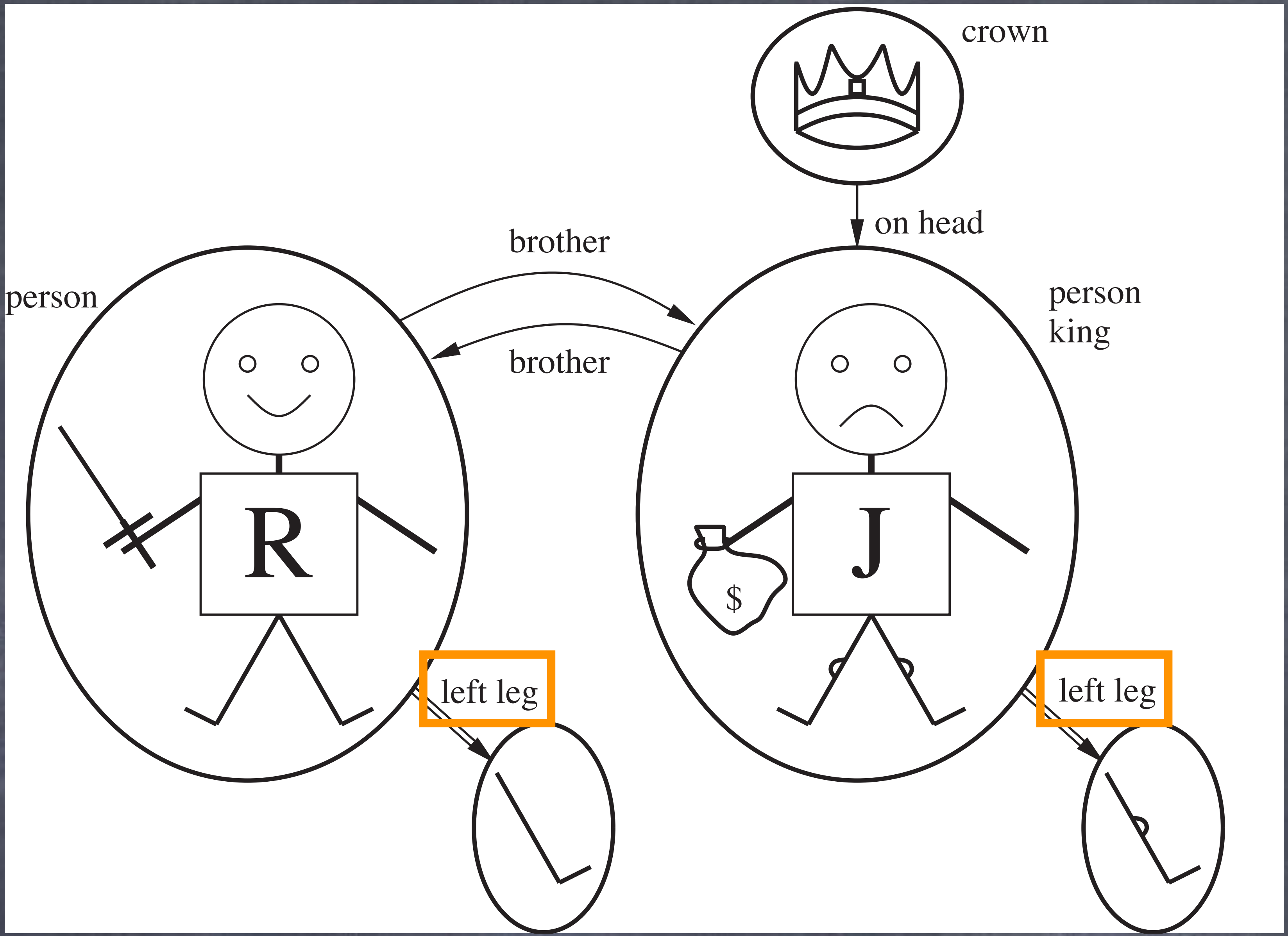
being a king



Binary Relations:

two things being brothers

one thing being on the head of another



Functions:

the left leg of something

$\Omega_I = \{$ *Richard, John, the crown, left leg 1, left leg 2* $\}$

$I(\texttt{Richard}) = \textit{Richard}$

$I(\texttt{John}) = \textit{John}$

$I(\texttt{Person}) = \{ \langle \textit{Richard} \rangle, \langle \textit{John} \rangle \}$

$I(\texttt{King}) = \{ \langle \textit{John} \rangle \}$

$I(\texttt{Crown}) = \{ \langle \textit{the crown} \rangle \}$

$I(\texttt{Brother}) = \{ \langle \textit{Richard,John} \rangle, \langle \textit{John,Richard} \rangle \}$

$I(\texttt{OnHead}) = \{ \langle \textit{the crown, John} \rangle \}$

$$I(\texttt{leftLegOf}) = \{ \langle \textit{Richard} \rangle \rightarrow \textit{left leg 1}, \\ \langle \textit{John} \rangle \rightarrow \textit{left leg 2} \}$$

$$\Omega_I = \{ \textit{Richard, John, the crown, left leg 1, left leg 2} \}$$

$$I(\texttt{Richard}) = \textit{John}$$

$$I(\texttt{John}) = \textit{the crown}$$

$$I(\texttt{Person}) = \{ \langle \textit{Richard} \rangle, \langle \textit{John} \rangle \}$$

$$I(\texttt{King}) = \{ \langle \textit{John} \rangle \}$$

$$I(\texttt{Crown}) = \{ \langle \textit{the crown} \rangle \}$$

$$I(\texttt{Brother}) = \{ \langle \textit{Richard,John} \rangle, \langle \textit{John,Richard} \rangle \}$$

$$I(\texttt{OnHead}) = \{ \langle \textit{the crown, John} \rangle \}$$

$$I(\texttt{leftLegOf}) = \{ \langle \textit{Richard} \rangle \to \textit{left leg 1}, \langle \textit{John} \rangle \to \textit{left leg 2} \}$$

$$\Omega_I = \{\ \textit{Richard, John, the crown, left leg 1, left leg 2}\ \}$$

$I(\texttt{Richard}) = \textit{left leg 1}$

$I(\texttt{John}) = \textit{left leg 1}$

$I(\texttt{Person}) = \{\ \langle \textit{Richard} \rangle, \langle \textit{John} \rangle\ \}$

$I(\texttt{King}) = \{\ \langle \textit{John} \rangle\ \}$

$I(\texttt{Crown}) = \{\ \langle \textit{the crown} \rangle\ \}$

$I(\texttt{Brother}) = \{\ \langle \textit{Richard}, \textit{John} \rangle, \langle \textit{John}, \textit{Richard} \rangle\ \}$

$I(\texttt{OnHead}) = \{\ \langle \textit{the crown, John} \rangle\ \}$

$$I(\texttt{leftLegOf}) = \{\ \langle \textit{Richard} \rangle \rightarrow \textit{left leg 1}, \\ \langle \textit{John} \rangle \rightarrow \textit{left leg 2}\ \}$$

$$\Omega_I = \{ \textit{Richard, John, the crown, left leg 1, left leg 2} \}$$

$$I(\texttt{Richard}) = \textit{Richard}$$

$$I(\texttt{John}) = \textit{John}$$

$$I(\texttt{Person}) = \{ \langle \textit{Richard} \rangle, \langle \textit{John} \rangle \}$$

$$I(\texttt{King}) = \{ \langle \textit{John} \rangle \}$$

$$I(\texttt{Crown}) = \{ \langle \textit{the crown} \rangle \}$$

$$I(\texttt{Brother}) = \{ \langle \textit{Richard,John} \rangle, \langle \textit{John,Richard} \rangle \}$$

$$I(\texttt{OnHead}) = \{ \langle \textit{the crown, John} \rangle \}$$

$$I(\texttt{leftLegOf}) = \{ \langle \textit{Richard} \rangle \to \textit{left leg 1}, \langle \textit{John} \rangle \to \textit{left leg 2} \}$$

First-Order Model (Possible World)

- Ontology (Domain of Discourse, Conceptualization)
  - Objects, relations, and functions
- Interpretation function I
  - Constant symbols $\rightarrow$ Objects
  - Predicate symbols $\rightarrow$ Relations (sets of tuples)
  - Function symbols $\rightarrow$ Functions (mappings)

Satisfaction

- A model (possible world) satisfies a sentence if it makes the sentence true
  - "A model of the sentence"

Terms

- Constant term c
    - $I(c) \in \Omega_I$
- Function term $f(t_1, ..., t_n)$
    - $I(f)$ = some function F
    - $I(t_i)$ = some object d_i
    - $I(f(t_1, ..., t_n)) = F(d_1, ..., d_n)$

Terms

- Constant term c
  - $I(c) \in \Omega_I$
- Function term $f(t_1, \ldots, t_n)$
  - $I(f)$ = some function F
  - $I(t_i)$ = some object $d_i \in \Omega_I$
  - $I(f(t_1, \ldots, t_n)) = F(d_1, \ldots, d_n)$

"The interpretation fixes the referent (or denotation) of every term."

Atomic Sentences

- Atomic sentence $P(t_1, ..., t_n)$
  - $I(P)$ = some relation Φ
  - $I(t_i)$ = some object d_i
  - $P(t_1, ..., t_n)$ is true if $\langle d_1, ..., d_n \rangle \in \Phi$

Atomic Sentences

- Atomic sentence $P(t_1, \ldots, t_n)$
  - $I(P)$ = some relation Φ
  - $I(t_i)$ = some object d_i
  - $P(t_1, \ldots, t_n)$ is true if $\langle d_1, \ldots, d_n \rangle \in \Phi$

"An atomic sentence is true in a given model if the relation referred to by the predicate symbol holds among the objects referred to by the arguments."

Complex Sentences

α	β	$\neg\alpha$	$\alpha \wedge \beta$	$\alpha \vee \beta$	$\alpha \Rightarrow \beta$	$\alpha \Leftrightarrow \beta$
false	false		false	false	true	true
false	true	true	false	true	true	false
true	false		false	true	false	false
true	true	false	true	true	true	true

Semantics of First-Order Logic

- Set of objects, with relations & functions
- Interpretation function
  - Constant symbols → objects
  - Predicate symbols → relations (tuples)
  - Function symbols → functions (mappings)
- An interpretation satisfies a sentence if it makes the sentence true

- Rooms adjacent to pits are breezy
- Socrates is a person
  All people are mortal
- Anybody's grandmother is either their mother's or their father's mother

- Rooms adjacent to pits are breezy
- Socrates is a person
  All people are mortal
- Anybody's grandmother is either their mother's or their father's mother

$$Person(Socrates)$$

- Rooms adjacent to pits are breezy
- Socrates is a person
  All people are mortal
- Anybody's grandmother is either their mother's or their father's mother

$$Person(Socrates)$$

True in I if $\langle I(Socrates)\rangle \in I(Person)$

- Rooms adjacent to pits are breezy
- Socrates is a person
  All people are mortal
- Anybody's grandmother is either their mother's or their father's mother

All people are mortal

Every object that is a person is also mortal

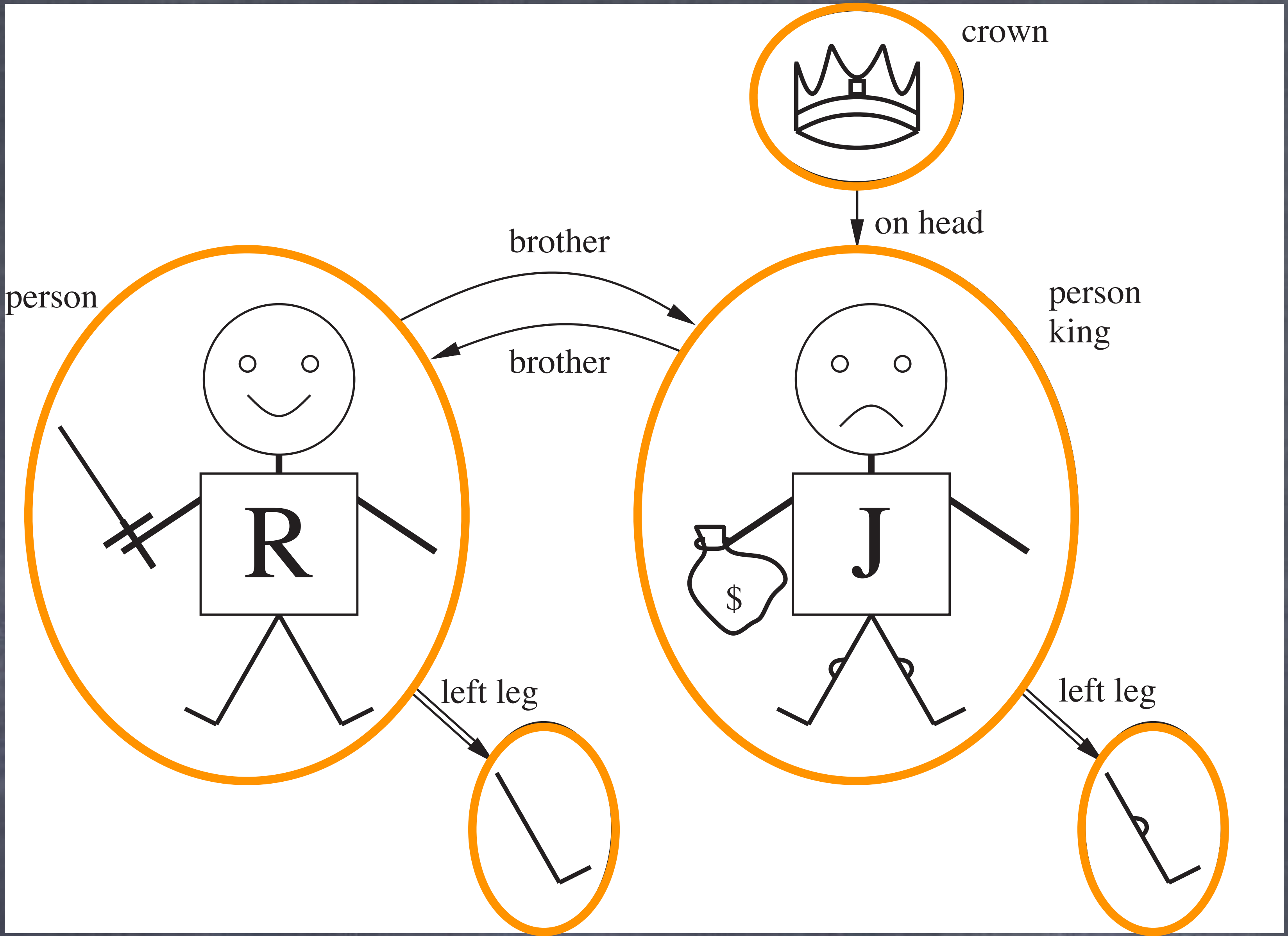
For every object x, if x is a person, then x is mortal

For every object x: $Person(x) \Rightarrow Mortal(x)$

Universal Quantification

- Syntax: $\forall x\, \varphi$
- Semantics: φ is true for every object x
  - Extended interpretation maps every variable to an object in the domain
  - $\forall x\, \varphi$ is true if φ is true in every extended interpretation

$$\forall x \; King(x) \Rightarrow Person(x)$$



Objects (Ω_I):

Richard, John, left leg 1, left leg 2, the crown

$$\forall x\ King(x) \Rightarrow Person(x)$$

$x \rightarrow$ Richard

$x \rightarrow$ John

$x \rightarrow$ Richard's left leg

$x \rightarrow$ John's left leg

$x \rightarrow$ the crown

$$\forall x\ King(x) \Rightarrow Person(x)$$

Richard is a king ⇒ Richard is a person

John is a king ⇒ John is a person

Richard's left leg is a king ⇒ Richard's left leg is a person

John's left leg is a king ⇒ John's left leg is a person

the crown is a king ⇒ the crown is a person

$$\forall x \; King(x) \Rightarrow Person(x)$$

Richard is a king ⇒ Richard is a person

true true

John is a king ⇒ John is a person True

Richard's left leg is a king ⇒ Richard's left leg is a person

John's left leg is a king ⇒ John's left leg is a person

the crown is a king ⇒ the crown is a person

$$\forall x\ King(x) \Rightarrow Person(x)$$

false

Richard is a king $\Rightarrow$ Richard is a person True

true true

John is a king $\Rightarrow$ John is a person True

false

Richard's left leg is a king $\Rightarrow$ Richard's left leg is a person True

false

John's left leg is a king $\Rightarrow$ John's left leg is a person True

false

the crown is a king $\Rightarrow$ the crown is a person True

$$\forall x\ King(x) \Rightarrow Person(x)$$ True!

false
Richard is a king ⇒ Richard is a person True

true true
John is a king ⇒ John is a person True

false
Richard's left leg is a king ⇒ Richard's left leg is a person True

false
John's left leg is a king ⇒ John's left leg is a person True

false
the crown is a king ⇒ the crown is a person True

Universal Quantification

- Syntax: $\forall x\, \varphi$
- Semantics: φ is true for every object x
  - Extended interpretation maps every variable to an object in the domain
  - $\forall x\, \varphi$ is true if φ is true in every extended interpretation

All people are mortal.

$$\forall x \; Person(x) \Rightarrow Mortal(x)$$

Rooms adjacent to pits are breezy.

$$\forall x \forall y \; Room(x) \land Pit(y) \land Adjacent(x, y) \Rightarrow Breezy(x)$$

Anybody's grandmother is either their mother's or their father's mother

$$\forall x \forall y \; Grandmother(x, y) \Rightarrow$$
$$x = mother(mother(y)) \lor x = mother(father(y)))$$

$$\forall x \; King(x) \land Person(x)$$

$$\forall x \; King(x) \wedge Person(x)$$ False!

Richard is a king ∧ Richard is a person — False

John is a king ∧ John is a person — True

Richard's left leg is a king ∧ Richard's left leg is a person — False

John's left leg is a king ∧ John's left leg is a person — False

the crown is a king ∧ the crown is a person — False

Rule: $\forall x \; King(x) \Rightarrow Person(x)$

Probably false statement:

$$\forall x \; King(x) \wedge Person(x)$$

Existential Quantification

- Syntax: $\exists x \, \varphi$
- Semantics: φ is true for some object x
  - Extended interpretation maps every variable to an object in the domain
  - $\exists x \, \varphi$ is true if φ is true in some extended interpretation

John has a crown on his head.

$$\exists x\ Crown(x) \wedge OnHead(x, John)$$

John has a crown on his head.

$$\exists x\ Crown(x) \land OnHead(x, John)$$

$x \rightarrow$ Richard

$x \rightarrow$ John

$x \rightarrow$ Richard's left leg

$x \rightarrow$ John's left leg

$x \rightarrow$ the crown

John has a crown on his head.

$\exists x \; Crown(x) \land OnHead(x, John)$ True!

$x \rightarrow$ Richard

$x \rightarrow$ John

$x \rightarrow$ Richard's left leg

$x \rightarrow$ John's left leg

$x \rightarrow$ the crown True

Existential Quantification

- Syntax: $\exists x \, \varphi$
- Semantics: φ is true for <u>some</u> object x
  - <u>Extended interpretation</u> maps every variable to an object in the domain
  - $\exists x \, \varphi$ is true if φ is true in <u>some</u> extended interpretation

Nested Quantifiers

Brothers are siblings

$$\forall x \forall y \; Brother(x,y) \Rightarrow Sibling(x,y)$$

Being a sibling is a symmetric relationship

$$\forall x \forall y \; Sibling(x,y) \Rightarrow Sibling(y,x)$$

Nested Quantifiers

Everyone (every person) loves someone

$\forall x\ Person(x) \Rightarrow \exists y\ Person(y) \wedge Loves(x,y)$

Nested Quantifiers

Everyone (every person) loves someone

$$\forall x \; Person(x) \Rightarrow \exists y \; Person(y) \wedge Loves(x,y)$$

Someone is loved by everyone

$$\exists x \; Person(x) \wedge \forall y \; Person(y) \Rightarrow Loves(y,x)$$

Nested Quantifiers

Everyone (every person) loves someone

$$\forall x\ Person(x) \Rightarrow \exists y\ Person(y) \wedge Loves(x,y)$$

Someone is loved by everyone

$$\exists x\ Person(x) \wedge \forall y\ Person(y) \Rightarrow Loves(y,x)$$

Someone loves everyone

$$\exists x\ Person(x) \wedge \forall y\ Person(y) \Rightarrow Loves(x,y)$$

$$\exists x\ \forall y\ Person(x) \wedge Person(y) \Rightarrow Loves(x,y)$$

First-Order Predicate Logic

- Syntax:
  - Constant, predicate, and function symbols
  - Terms, atomic sentences, connectives
  - Quantifiers and variables
- Semantics:
  - Domain of objects, relations, functions
  - First-order interpretation
  - Extended interpretation
  - Satisfaction (sentence true in a possible world)

Entailment

- α entails β ($\alpha \vDash \beta$) when:
  - β is true in **every** world considered possible by α
  - Every model of α is also a model of β
  - Models(α) $\subseteq$ Models(β)

All Possible Models

- # of objects in the world from 1 to ∞
- Some constants refer to the same object
- Some objects are not referred to by any constant (“unnamed”)
- Relations and functions defined over sets of subsets of objects
- Variables range over all possible objects in extended interpretations

Constant symbols: $\{ R, J \}$

Relation symbol: $P(\cdot,\cdot)$

Constant symbols: $\{ \; R, \; J \; \}$

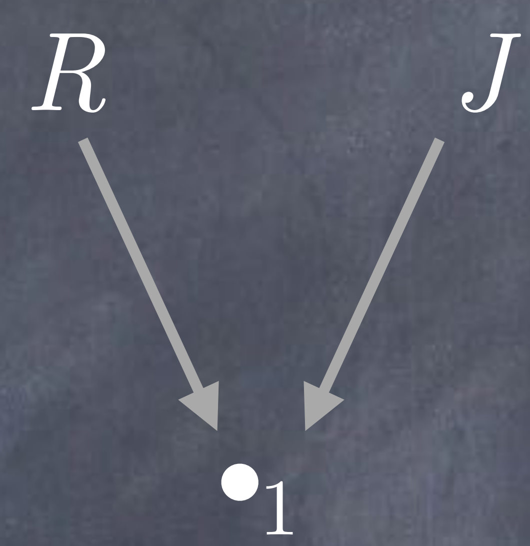
Relation symbol: $P(\cdot,\cdot)$

$$\Omega_I = \{ \; \bullet_1 \; \}$$

$$I(R) = \bullet_1$$

$$I(J) = \bullet_1$$

$$I(P) = \{ \; \}$$



Constant symbols: $\{ R, J \}$

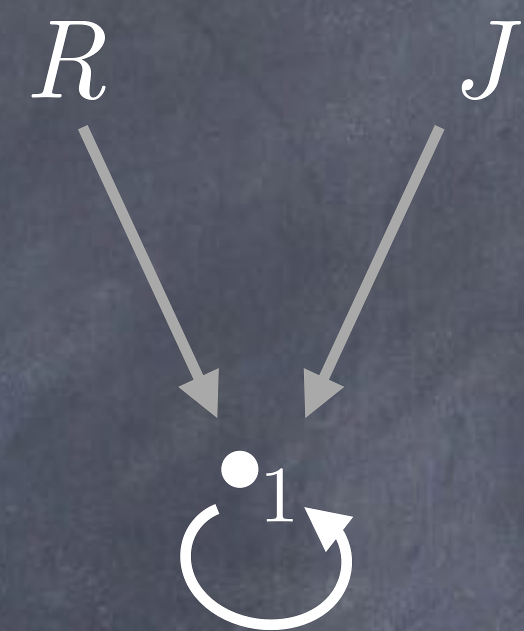
Relation symbol: $P(\cdot,\cdot)$

$$\Omega_I = \{ \bullet_1 \}$$

$$I(R) = \bullet_1$$

$$I(J) = \bullet_1$$

$$I(P) = \{ \langle \bullet_1, \bullet_1 \rangle \}$$



Constant symbols: $\{ R, J \}$

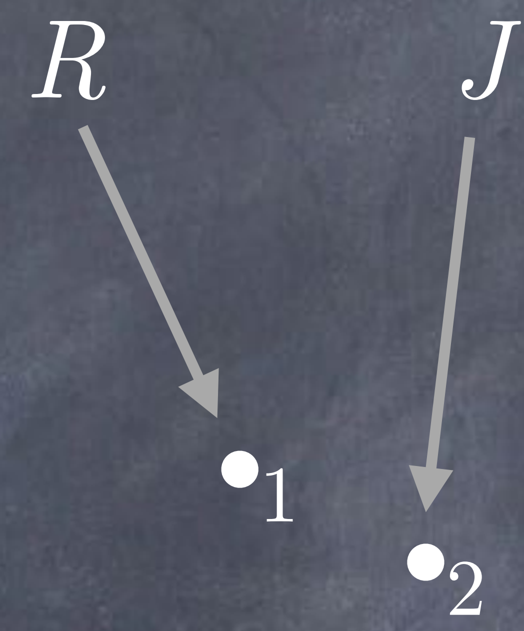
Relation symbol: $P(\cdot, \cdot)$

$$\Omega_I = \{ \bullet_1, \bullet_2 \}$$

$$I(R) = \bullet_1$$

$$I(J) = \bullet_2$$

$$I(P) = \ldots$$



Constant symbols: $\{ R, J \}$

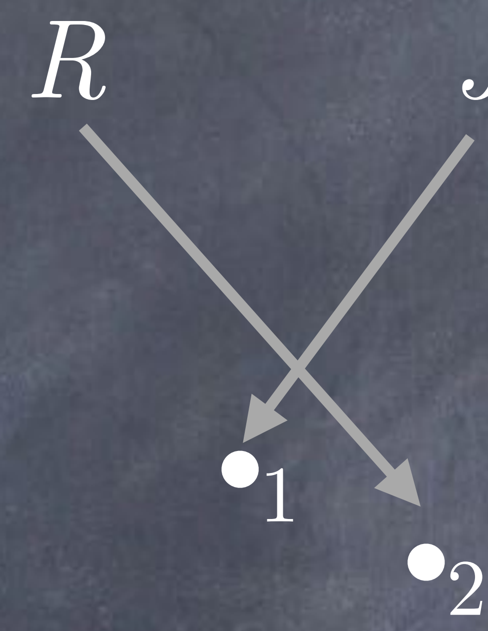
Relation symbol: $P(\cdot,\cdot)$

$$\Omega_I = \{ \bullet_1, \bullet_2 \}$$

$$I(R) = \bullet_2$$

$$I(J) = \bullet_1$$

$$I(P) = \ldots$$



Constant symbols: $\{\ R,\ J\ \}$

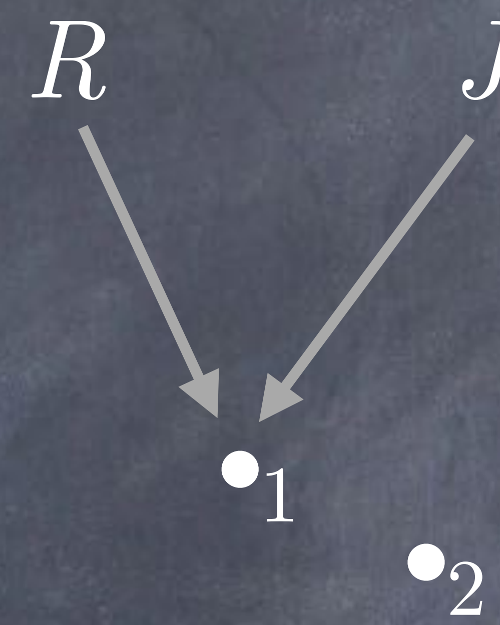
Relation symbol: $P(\cdot,\cdot)$

$$\Omega_I = \{\ \bullet_1,\ \bullet_2\ \}$$

$$I(R) = \bullet_1$$

$$I(J) = \bullet_1$$

$$I(P) = \ldots$$



Constant symbols: $\{\ R,\ J\ \}$

Relation symbol: $P(\cdot,\cdot)$

$$\Omega_I = \{\ \bullet_1,\ \bullet_2\ \}$$

$$I(R) = \bullet_2$$

$$I(J) = \bullet_2$$

$$I(P) = \ldots$$

Constant symbols: $\{ R, J \}$

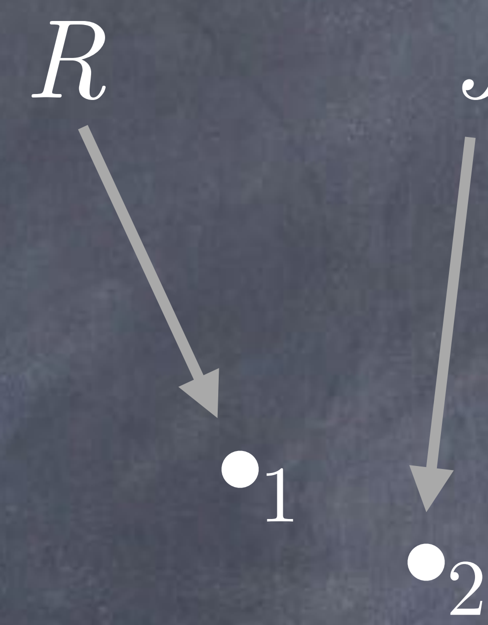
Relation symbol: $P(\cdot,\cdot)$

$$\Omega_I = \{ \bullet_1, \bullet_2 \}$$

$$I(R) = \bullet_1$$

$$I(J) = \bullet_2$$

$$I(P) = \ldots$$



$\langle \bullet_1, \bullet_1 \rangle$, $\langle \bullet_1, \bullet_2 \rangle$, $\langle \bullet_2, \bullet_1 \rangle$, $\langle \bullet_2, \bullet_2 \rangle$: $2^2=4$ binary tuples

$2^{2^2}=16$ interpretations of P

64 possible interpretations

Constant symbols: $\{\ R,\ J\ \}$

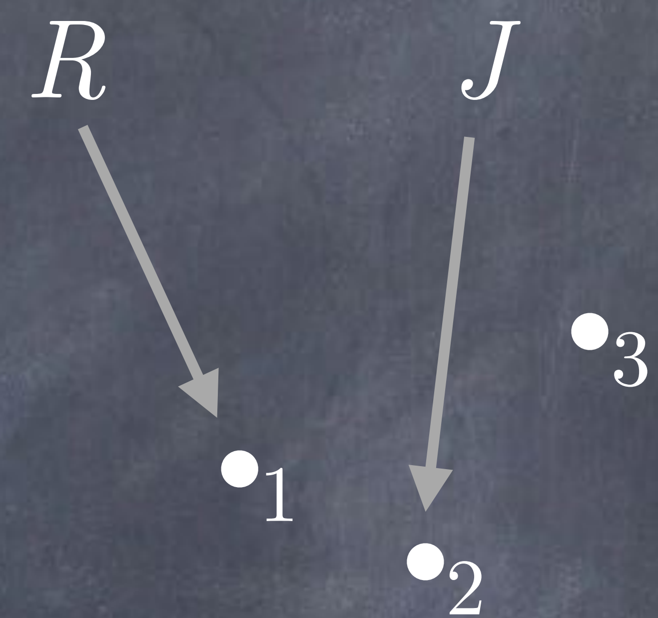
Relation symbol: $P(\cdot,\cdot)$

$$\Omega_I = \{\ \bullet_1,\ \bullet_2,\ \bullet_3\ \}$$

$$I(R) = \bullet_1$$

$$I(J) = \bullet_2$$

$$I(P) = \ldots$$



2^3=8 interpretations of R and J

2^{2^3}=2^8=256 interpretations of P

2048 possible interpretations

Constant symbols: $\{ R, J \}$

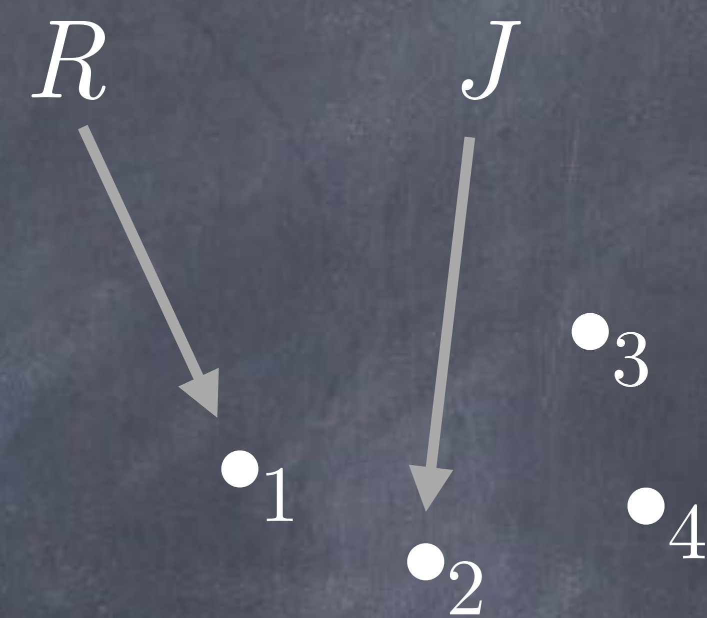
Relation symbol: $P(\cdot, \cdot)$

$$\Omega_I = \{ \bullet_1, \bullet_2, \bullet_3, \bullet_4 \}$$

$$I(R) = \bullet_1$$

$$I(J) = \bullet_2$$

$$I(P) = \ldots$$



1,048,576 possible interpretations

Computing Entailment

- Number of models HUGE (often unbounded)
- Can't do model checking

Computing Entailment

- Number of models HUGE (often unbounded)
- Can't do model checking
- Look for inference rules, do theorem proving

For next time:

AIMA Ch 9