# FIRST-ORDER LOGIC

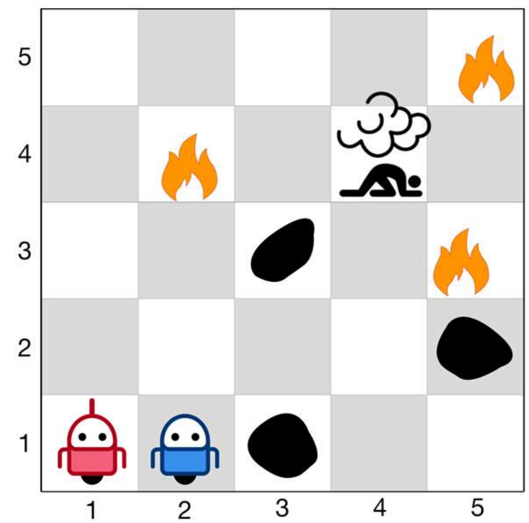
Fabrizio Santini | COMP 131A

VERSION 1.1

TODAY ON AI

- Disaster zone
- Propositional Logic is a weak language
- First-order Logic
- Questions?

- A building of size **WIDTH** × **LENGTH**
- RR1 starts in square (0, 0)
- BR2 starts in square (1, 0)
- There is a victim somewhere
- Robots of type SRR have sensors that can detect **heat** and **drop-offs**:
  - They reliably sense drop-offs and victims from exactly one (non-diagonal) block away
  - They are not good enough to say in what direction the victim or hole was



## Propositional Logic is a weak language

GOOD AND BAD

- Propositional Logic is **declarative**
- Propositional Logic is **compositional**:
  - The meaning of $S1 \land S2$ is derived from meaning of $S1$ **and** of $S2$
- Propositional Logic is **context-independent** (unlike natural language)
- Propositional Logic assumes that the world contains only **facts**

Unfortunately Propositional Logic has very **limited** expressive power (unlike natural language):

- Hard to identify **individuals**:
  **Every pit causes breeze in adjacent squares**
- Can’t directly talk about **properties** and relations between individuals:
  **RR1 is red**
- **Generalizations**, **patterns**, **regularities** can’t easily be represented:
  **All SRR robots have a left arm**

## First-order Logic

SECTION 01

**First-order Logic** is a powerful evolution of the Propositional Logic that is expressive enough to concisely represent generalizations, relationships, properties.

- Like natural language, FOL assumes that the world contains concepts like:
  - **Objects**: robot, people, houses, numbers, colors, baseball games, wars, …
  - **Relations**: red, blue, round, prime, brother of, bigger than, part of, comes between, …
  - **Functions**: leader of, father of, best friend, one more than, plus, …

- First-order logic (FOL) models the world in terms of:
  - **Objects**: things with individual identities
  - **Properties**: properties of objects that distinguish them from other objects
  - **Relations**: relationships that hold among sets of objects
  - **Functions**: a subset of relations where there is only one value for any given input

- Examples:
  - **Objects**: $Victim, Robot, Sq00, Sq01 \ldots$
  - **Properties**: $blue(Robot)$
  - **Relations**: $\mathrm{C}olor(x, Red)$
  - **Functions**: $Loc(Robot, Step0) = Sq00$

- **Term**: Constant symbols, variable symbols, or functions
- **Atom**: Relations
- Atoms are combined by **connectives** to produce other **atoms**:

| | | |
|---|---|---|
| $\wedge$ | AND | Conjunction |
| $\vee$ | OR | Disjunction |
| $\rightarrow$ | IMPLIES | Implication / conditional |
| $\leftrightarrow$ | IS EQUIVALENT | Biconditional |
| $\neg$ | NOT | Negation |

- **Sentence**: quantified sentences

| | | |
|---|---|---|
| $\exists$ | EXISTS | Existential quantifier |
| $\forall$ | FOR ALL | Universal quantifier |

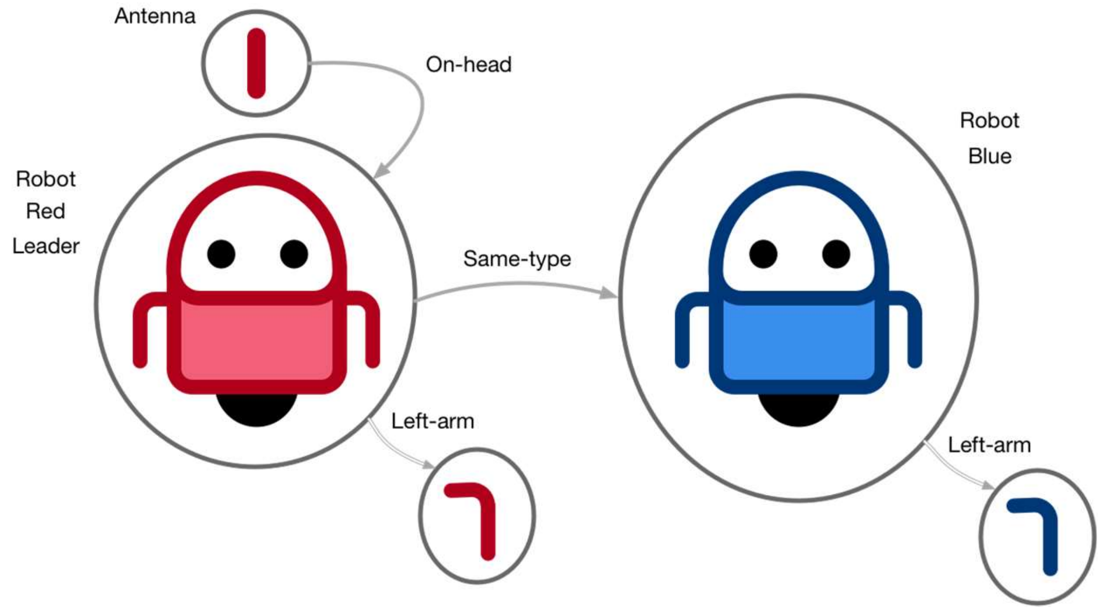
**FIRST-ORDER LOGIC GRAMMAR**

| | |
|---|---|
| **Term** → | Constant \| Variable \| Function |
| **Atom** → | Predicate(Term1, Term2, …) \|<br>$\neg s$ \| Atom Connective Atom |
| **Sentence** → | ( Sentence ) \|<br>Quantifier Sentence |
| **Connective** → | $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$ |
| **Quantifier** → | $\exists$, $\forall$ |

- Sentences are **true** with respect to a **model** and an **interpretation**
- Model contains objects (**domain elements**) and relations among them
- Interpretation specifies referents for:

| | | |
|---|---|---|
| **constant symbols** | → | objects |
| **predicate symbols** | → | relations |
| **function symbols** | → | functional relations |

- A sentence $predicate(term_1, \ldots, term_n)$ is **true** $iff$ the **objects** referred to by $term_1, \ldots, term_n$ are in the **relation** referred to by **predicate**


Antenna
On-head
Robot
Red
Leader
Same-type
Robot
Blue
Left-arm
Left-arm

**Existential** quantification:

- $\exists x:\ P(x)$ means that P holds for some value of x in the domain associated with that variable:

  ∃x: **Square**(x) ∧ **Unsafe**(x) means **There is an unsafe square**

- It permits one to make a statement about some object without naming it

- Roughly speaking the existential quantifier is equivalent to the **disjunction** of **instantiations** of $P(x)$:

  **Square**(x) ∧ **Unsafe**(x) ∨
  **Square**(x) ∧ **Unsafe**(x) ∨
  **Square**(x) ∧ **Unsafe**(x) ∨
  ⋮

- Existential quantifiers are normally used with **and** (∧) to specify a list of properties about that particular object:

  ∃x: **Square**(x) ∧ **Free**(x) means **There is a square which is also free**

- A common mistake is to use it with an implication (→):

  ∃x: **Square**(x) → **Free**(x) means **There is a square [therefore] it is free**

**Universal** quantification:

- $\forall x\colon P(x)$ means that $P$ holds for all values of $x$ in the domain associated with that variable:

∀x: **Adjacent**(x, **Sq01**) ∧ **Unsafe**(x, **Sq01**) → **Dropoff-detected**(x)

- Roughly speaking the existential quantifier is equivalent to the **disjunction** of instantiations of $P(x)$:

Adjacent(**Sq00**, **Sq01**) ∧ Unsafe(**Sq00**, **Sq01**) → Dropoff–detected(**Sq00**) ∧
Adjacent(**Sq02**, **Sq01**) ∧ Unsafe(**Sq02**, **Sq01**) → Dropoff–detected(**Sq02**) ∧
Adjacent(**Sq13**, **Sq01**) ∧ Unsafe(**Sq13**, **Sq01**) → Dropoff–detected(**Sq13**) ∧
⋮

- Universal quantifiers are often used with "→" to form **rules**
- Universal quantification is rarely used to make blanket statements about every individual in the world:

  ∀x: **Unsafe**(x) means **Every object [in the world] is unsafe**

  ∀x: **Square**(x) ∧ **Unsafe**(x) means **All the square [in the world] are unsafe**

- Switching the order of universal quantifiers does not change the meaning:

$$\forall x\ \forall y:\ P(x,y) \leftrightarrow \forall y \forall x\ P(x,y)$$

- Switching the order of existential quantifiers not does change the meaning:

$$\exists x\ \exists y:\ P(x,y) \leftrightarrow \exists y \exists x\ P(x,y)$$

- Switching the order of existential and universal quantifiers does change the meaning:

∀x ∃y: **Loves**(x, y) or ∀x (∃y: **Loves**(x, y)) means **...?**

∃y ∀x: **Loves**(x, y) or ∃y (∀x: **Loves**(x, y)) means **...?**

”Everybody loves someone”
“There is someone who is loved by everybody”

- Every gardener likes the sun.
  $\forall x$: **gardener**(x) → **likes**(x, **Sun**)

- You can fool some of the people all of the time.
  $\exists x\ \forall t$: (**person**(x) ∧ **time**(t)) → **can-fool**(x, t)

- You can fool all of the people some of the time.
  $\forall x\ \exists t$: (**person**(x) ∧ **time**(t)) → **can-fool**(x, t)

- All purple mushrooms are poisonous.
  $\forall x$: (**mushroom**(x) ∧ **purple**(x)) → **poisonous**(x)

- No purple mushroom is poisonous.
  $\nexists x$: **purple**$(x) \land$ **mushroom**$(x) \land$ **poisonous**$(x)$
  $\forall x$: **mushroom**$(x) \land$ **purple**$(x)) \rightarrow \neg$**poisonous**$(x)$

- There are exactly two purple mushrooms.
  $\exists x\ \exists y$: **mushroom**$(x) \land$ **purple**$(x) \land$ **mushroom**$(y) \land$ **purple**$(y) \land \neg(x = y) \land$
  $\forall z$: (**mushroom**$(z) \land$ **purple**$(z)) \rightarrow ((x = z) \lor (y = z))$

- Deb is not tall.
  $\neg$**tall**(**Deb**)

- We can relate sentences involving ∀ and ∃ using De Morgan's laws:

| SENTENCE | EQUIVALENT |
|---|---|
| $\forall x \ \neg P(x)$ | $\nexists x \ P(x)$ |
| $\neg \forall x \ P(x)$ | $\exists x \ \neg P(x)$ |
| $\forall x \ P(x)$ | $\nexists x \ \neg P(x)$ |
| $\exists x \ P(x)$ | $\neg \forall x \ \neg P(x)$ |

FOL introduces the concept of equality:

- At timestep 0, the robot was at Square (0, 0).
  $\mathbf{Location}(\mathbf{Robot}, \mathbf{T0}) = \mathbf{Sq00}$

- The robot was located at different places in timestep 0 and timestep 1.
  $x = \mathbf{Location}(\mathbf{Robot}, \mathbf{T0}) \land y = \mathbf{Location}(\mathbf{Robot}, \mathbf{T1}) \land \neg (x = y)$
  $x = \mathbf{Location}(\mathbf{Robot}, \mathbf{T0}) \land y = \mathbf{Location}(\mathbf{Robot}, \mathbf{T1}) \land (x \neq y)$

- The robot is different from the victim.
  $\neg(\mathbf{Robot} = \mathbf{Victim})$
  $\mathbf{Robot} \neq \mathbf{Victim}$

QUESTIONS ?