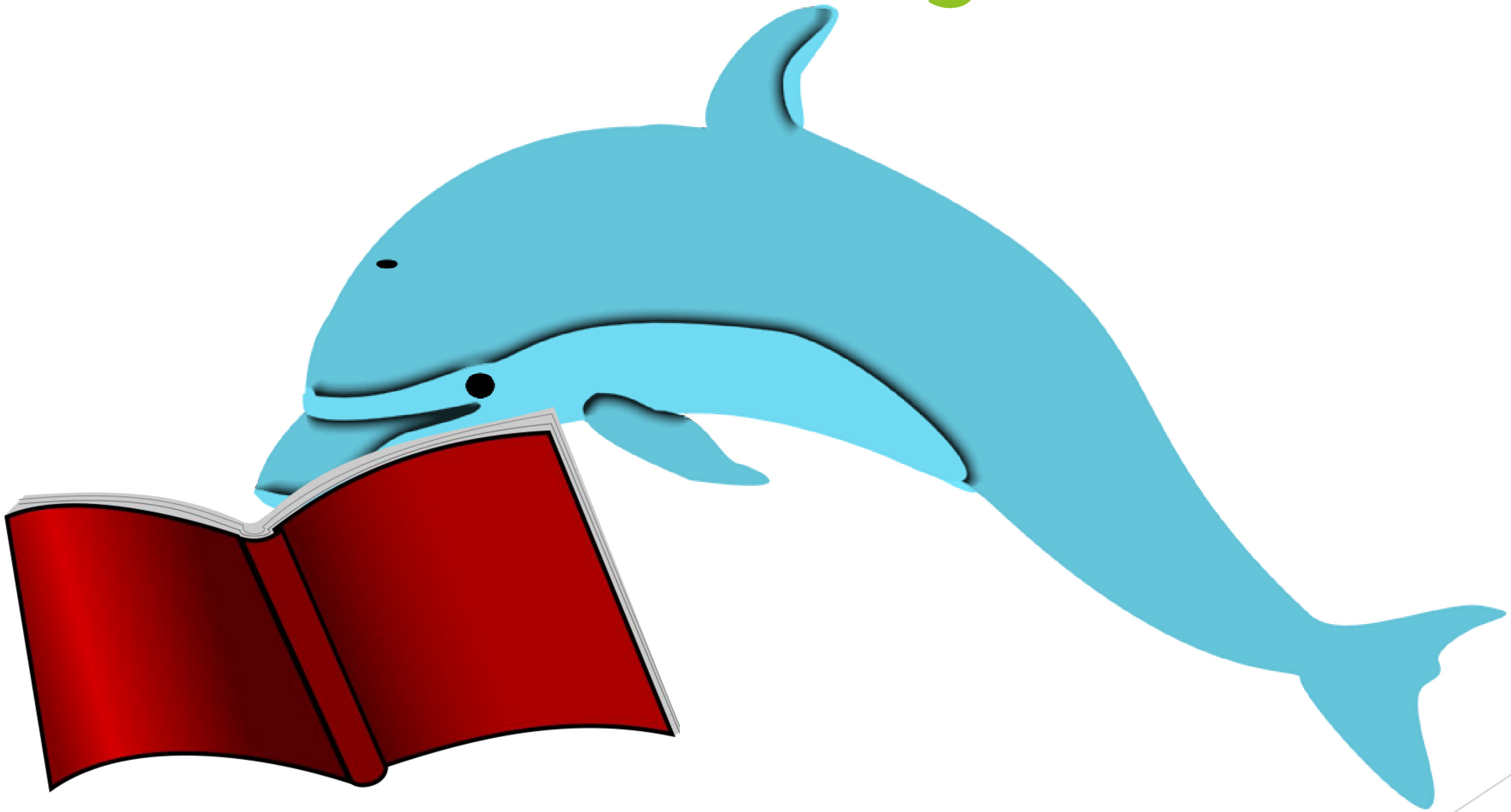
Resolution in First-Order Logic
Default Reasoning



Review

Substitution

Subst({x/Nemo}, Fish(x) ⇒ Swims(x)) gives us:

Fish(Nemo) ⇒ Swims(Nemo)

Universal Instantiation (entailment)

$$\frac{\forall x\ \mathrm{Fish}(x) \Rightarrow \mathrm{Swims}(x)}{\mathrm{Fish}(\mathrm{Nemo}) \Rightarrow \mathrm{Swims}(\mathrm{Nemo})}$$

Existential Instantiation / Skolemization

$$\frac{\exists x\ \mathrm{Bird}(x) \land \neg \mathrm{Flies}(x)}{\mathrm{Bird}(C1) \land \neg \mathrm{Flies}(C1)}$$

$$\frac{\forall\ x\ \exists y\ \mathrm{Faculty}(y) \land \mathrm{Advises}(y, x)}{\forall\ x\ \mathrm{Faculty}(F(x)) \land \mathrm{Advises}(F(x), x)}$$

Review

Unification

Unify(Knows(John, x), Knows(y, Anna)): {x/Anna, y/John}

Conversion to CNF

1. Eliminate biconditionals and implications
2. Move ¬ inwards
3. Standardize variables
4. Skolemize
5. Drop universal quantifiers
6. Apply distributivity law (∨ over ∧)

Resolution Rule - Propositional Logic

$$\frac{A \vee B, \neg A \vee C}{B \vee C}$$

We took two sentences of the form (A V B and ¬A V C)

We produced a new sentence containing the literals of the original sentences except for the two complementary literals (A and ¬A): B V C

Resolution Rule - First Order Logic

$$\frac{A \lor B, \; \neg A' \lor C}{(B \lor C)\theta}$$

We take two sentences of the form ($A \lor B$ and $\neg A' \lor C$).

We unify A and A' with a substitution θ.

We produce a new sentence containing the literals of the original sentences except for the two literals (A and ¬A'):

$(B \lor C)\theta$

Resolution Rule - First Order Logic

$$\frac{A \vee B,\ \neg A' \vee C}{(B \vee C)\theta}$$

Wanda is a fish or Wanda is a cat: $Fish(Wanda) \vee Cat(Wanda)$

All fish can swim: $\forall x\ Fish(x) \Rightarrow Swims(x)$

in CNF: $\neg\ Fish(x) \vee Swims(x)$

We unify $Fish(Wanda)$ and $Fish(x)$ with the substitution

$\theta = \{x/Wanda\}$

We can conclude that: $Cat(Wanda) \vee Swims(Wanda)$

Resolution

Goal is to determine if KB $\models \alpha$

- To show that KB $\models \alpha$, we show KB $\wedge \neg\alpha$ is unsatisfiable.
- Convert KB $\wedge \neg\alpha$ to CNF (Conjunctive Normal Form)
- Apply resolution rule repeatedly
- At the end: empty clause - unsatisfiable

Resolution Example

KB :

Whoever can read is literate.

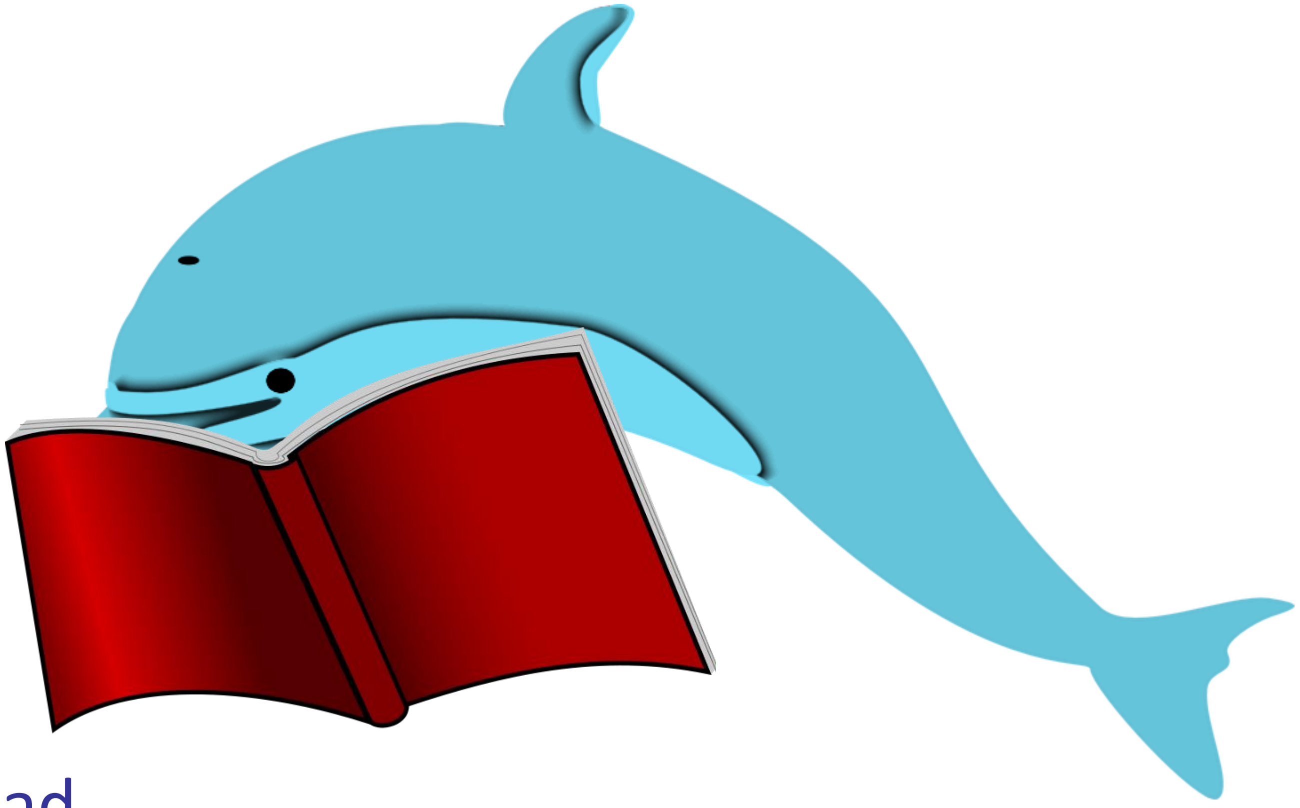
$\forall x\ Read(x) \Rightarrow Literate(x)$

Dolphins are not literate.

$\forall x\ Dolphin(x) \Rightarrow \neg Literate(x)$

Some dolphins are intelligent.

$\exists x\ Dolphin(x) \wedge Intelligent(x)$



To prove (α):

Some intelligent beings cannot read.

$\exists x\ Intelligent(x) \wedge \neg Read(x)$

Resolution Example

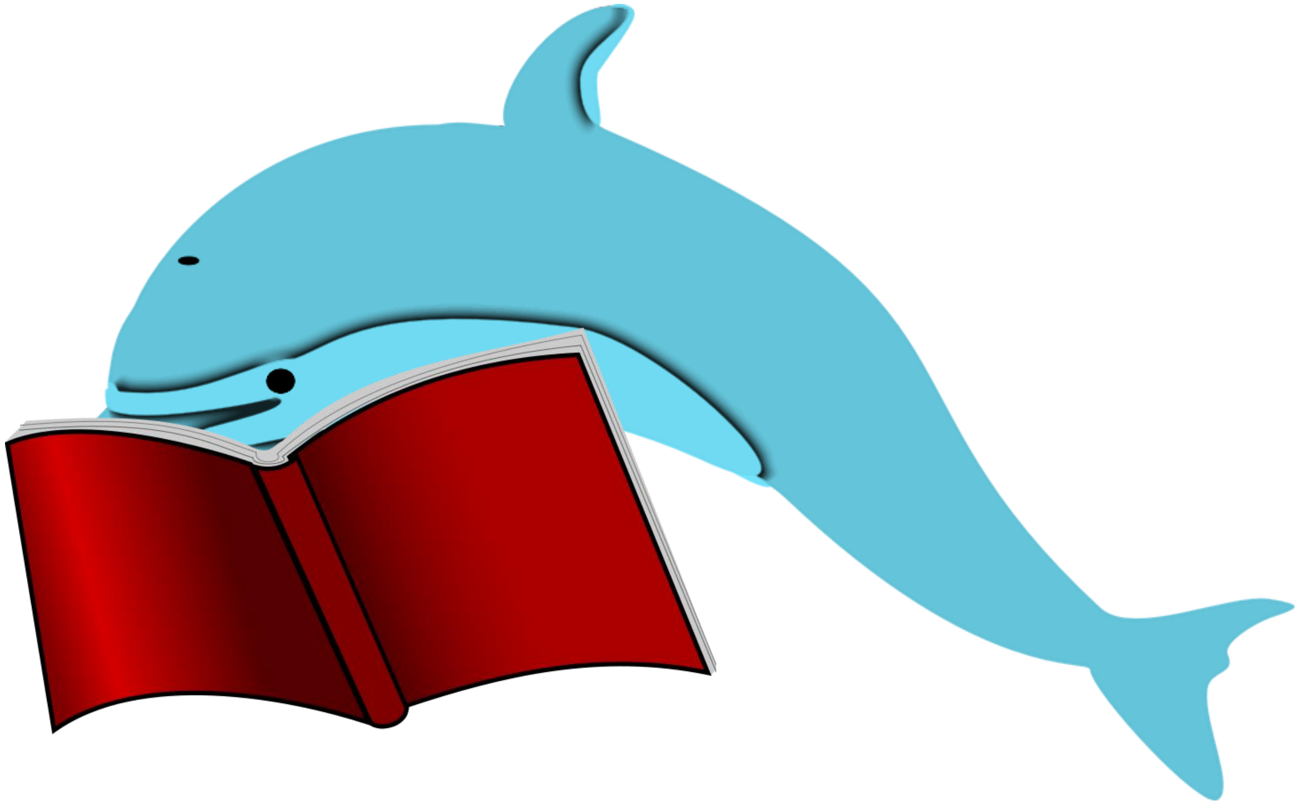
KB ∧ ¬ α :

$\forall x\ Read(x) \Rightarrow Literate(x)$

$\forall x\ Dolphin(x) \Rightarrow \neg Literate(x)$

$\exists x\ Dolphin(x) \wedge Intelligent(x)$

$\neg (\exists x\ Intelligent(x) \wedge \neg Read(x))$



Conversion to CNF

Step 1: Eliminate biconditionals and implications.

Replace $\alpha \Rightarrow \beta$ with $\neg\alpha \vee \beta$

$KB \wedge \neg \alpha$:

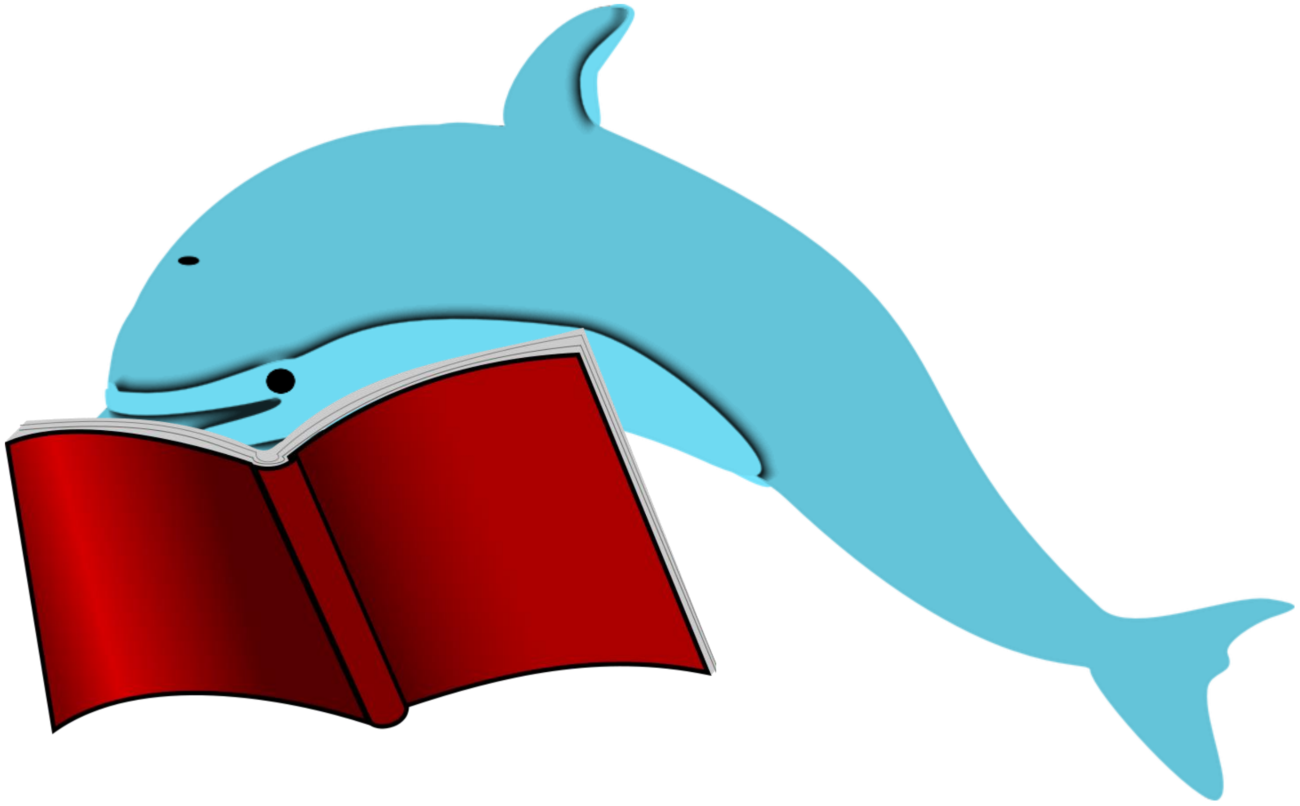
$\forall x\ Read(x) \Rightarrow Literate(x)$

$\forall x\ (\neg Read(x) \vee Literate(x))$

$\forall x\ Dolphin(x) \Rightarrow \neg Literate(x)$

$\forall x\ (\neg Dolphin(x) \vee \neg Literate(x))$

$\exists x\ Dolphin(x) \wedge Intelligent(x)$



$\neg (\exists x\ Intelligent(x) \wedge \neg Read(x))$

Conversion to CNF

$KB \wedge \neg \alpha$:

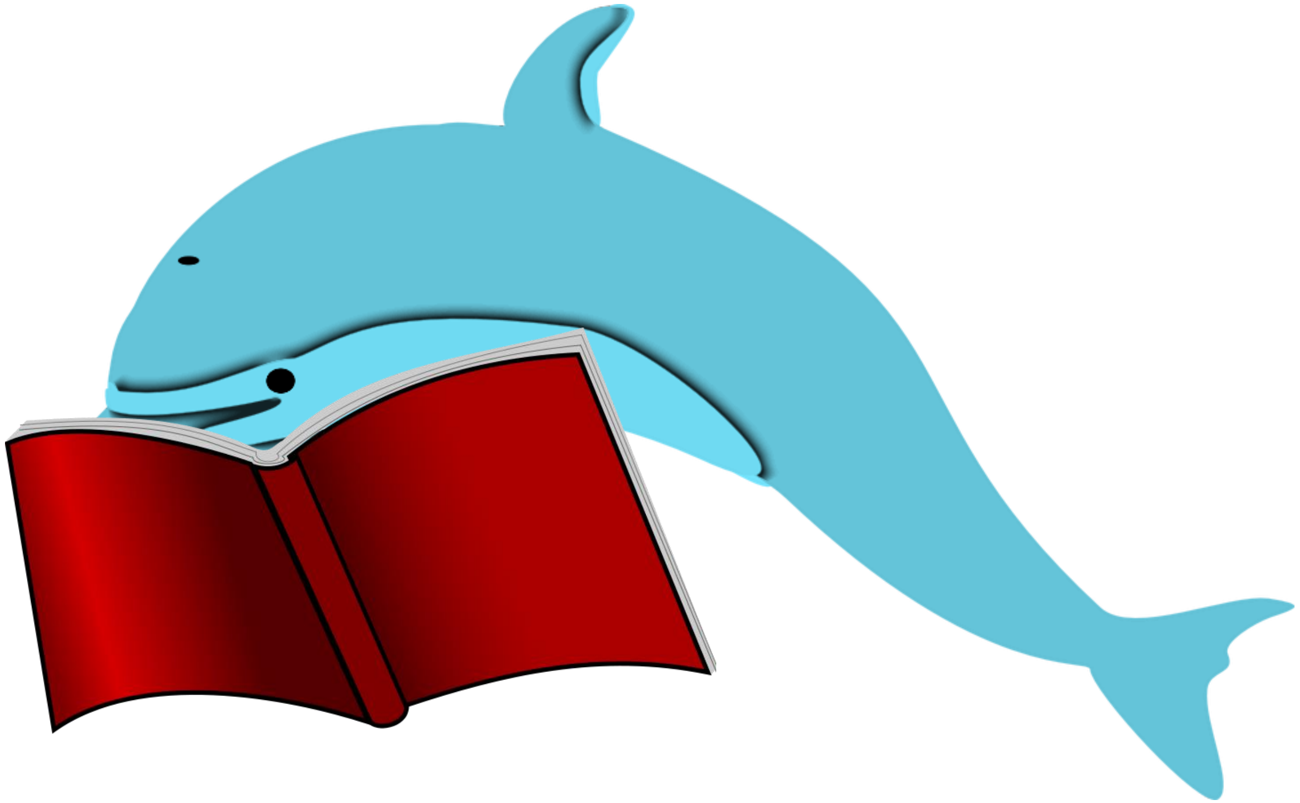
$\forall x\, (\neg\, Read(x) \vee Literate(x))$

$\forall x\, (\neg Dolphin(x) \vee \neg Literate(x))$

$\neg$

$\exists x\, Dolphin(x) \wedge Intelligent(x)$

$\neg\, (\exists x\, Intelligent(x) \wedge \neg Read(x))$



Conversion to CNF

Step 2: Move ¬ inwards.

$\forall x\ (\neg \text{Read}(x) \vee \text{Literate}(x))$

$\forall x\ (\neg \text{Dolphin}(x) \vee \neg \text{Literate}(x))$

$\exists x\ \text{Dolphin}(x) \wedge \text{Intelligent}(x)$

$\neg\ (\exists x\ \text{Intelligent}(x) \wedge \neg \text{Read}(x))$

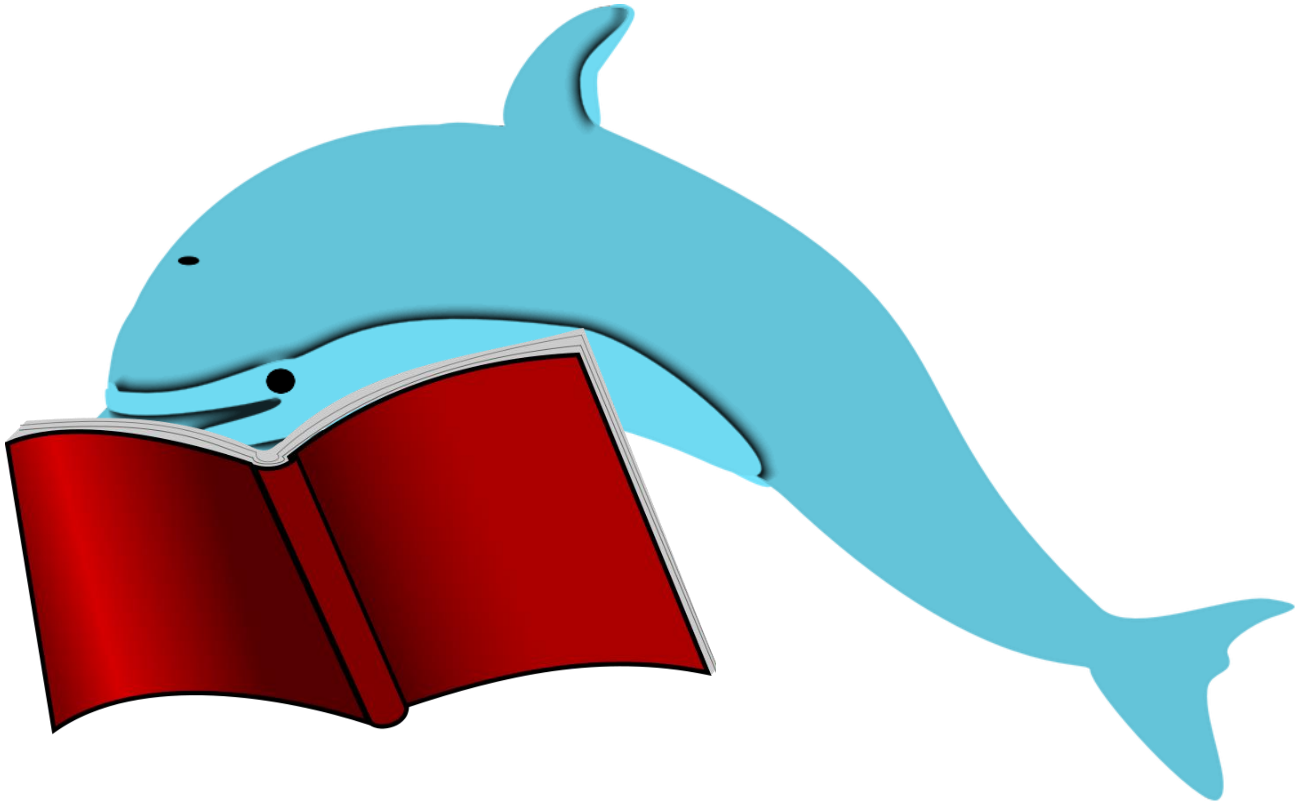
$\neg \exists\ x, p \equiv \forall\ x\ \neg p$

$\forall x\ \neg\ (\text{Intelligent}(x) \wedge \neg \text{Read}(x))$

$\neg(p \wedge q) \equiv \neg p \vee \neg q$ - De Morgan's law

$\forall x\ \neg\ \text{Intelligent}(x) \vee \neg\ \neg \text{Read}(x)$

$\forall x\ \neg\ \text{Intelligent}(x) \vee \text{Read}(x)$



Conversion to CNF

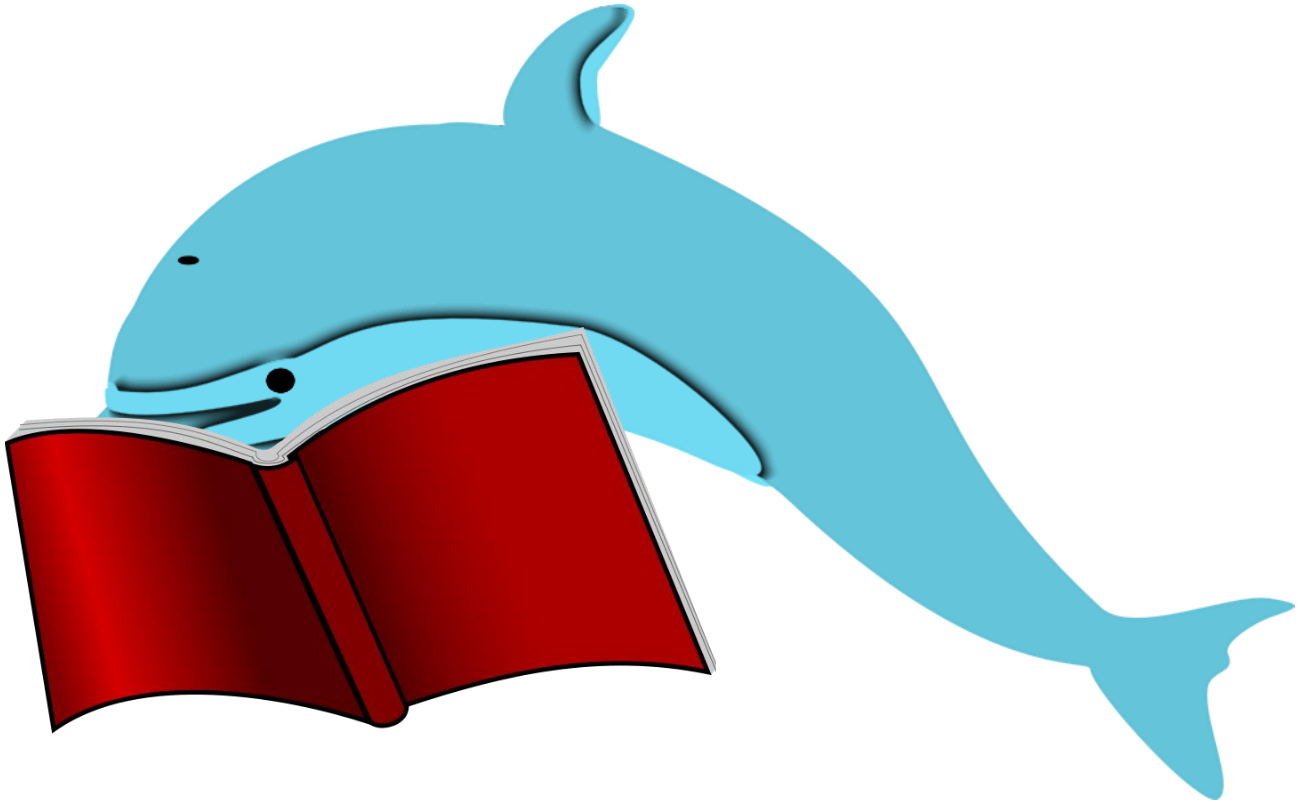
$KB \wedge \neg \alpha$:

$\forall x\ (\neg\ Read(x) \vee Literate(x))$

$\forall x\ (\neg Dolphin(x) \vee \neg Literate(x))$

$\exists x\ Dolphin(x) \wedge Intelligent(x)$

$\forall x\ \neg\ Intelligent(x) \vee Read(x)$



Conversion to CNF

Step 3: Standardize variables: each quantifier should use a different one

KB ∧ ¬ α:

∀x (¬ Read(x) V Literate(x))

∀x (¬Dolphin(x) V ¬Literate(x))

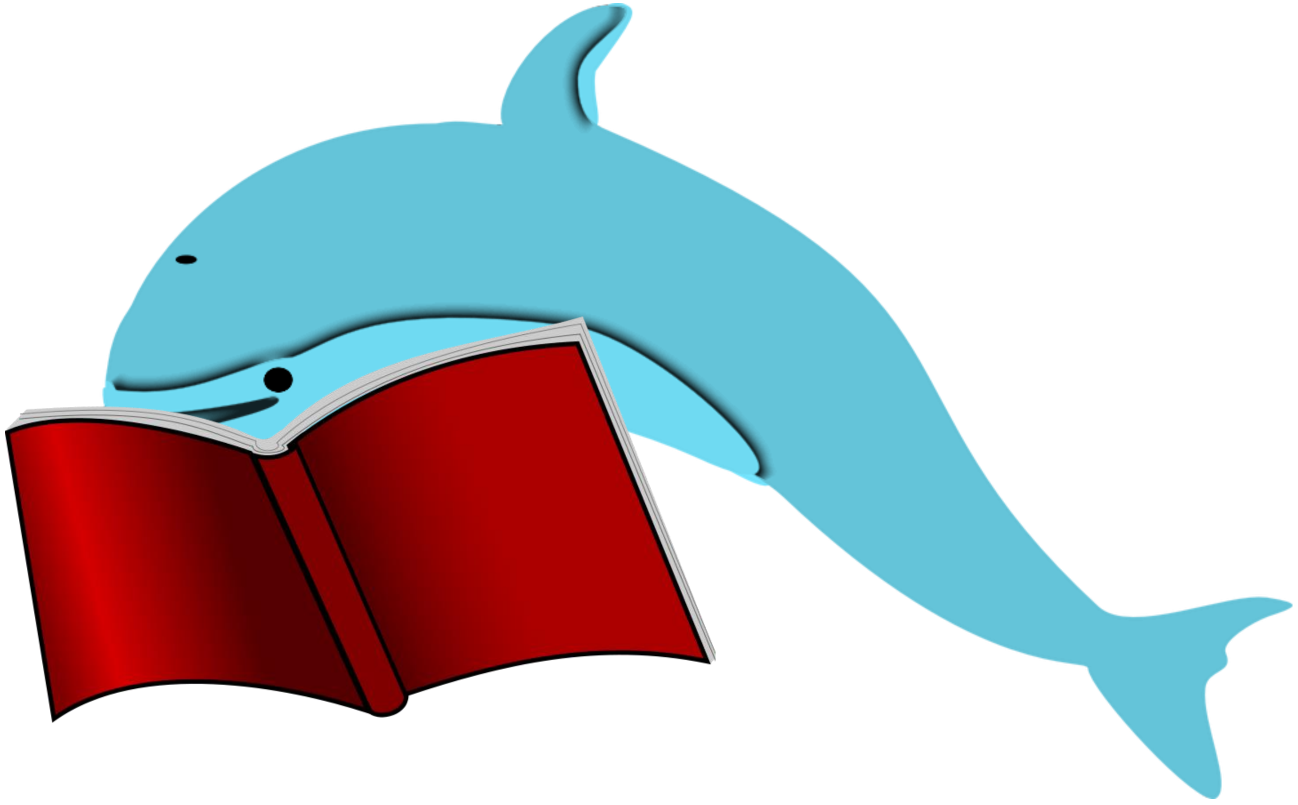
∀y (¬Dolphin(y) V ¬Literate(y))

∃x Dolphin(x) ∧ Intelligent(x)

∃z Dolphin(z) ∧ Intelligent(z)

∀x ¬ Intelligent(x) V Read(x)

∀t ¬ Intelligent(t) V Read(t)



Conversion to CNF

Step 4: Skolemize. The existential variable is replaced by a Skolem constant (since it is not inside a universal quantifier)

KB ∧ ¬ α:

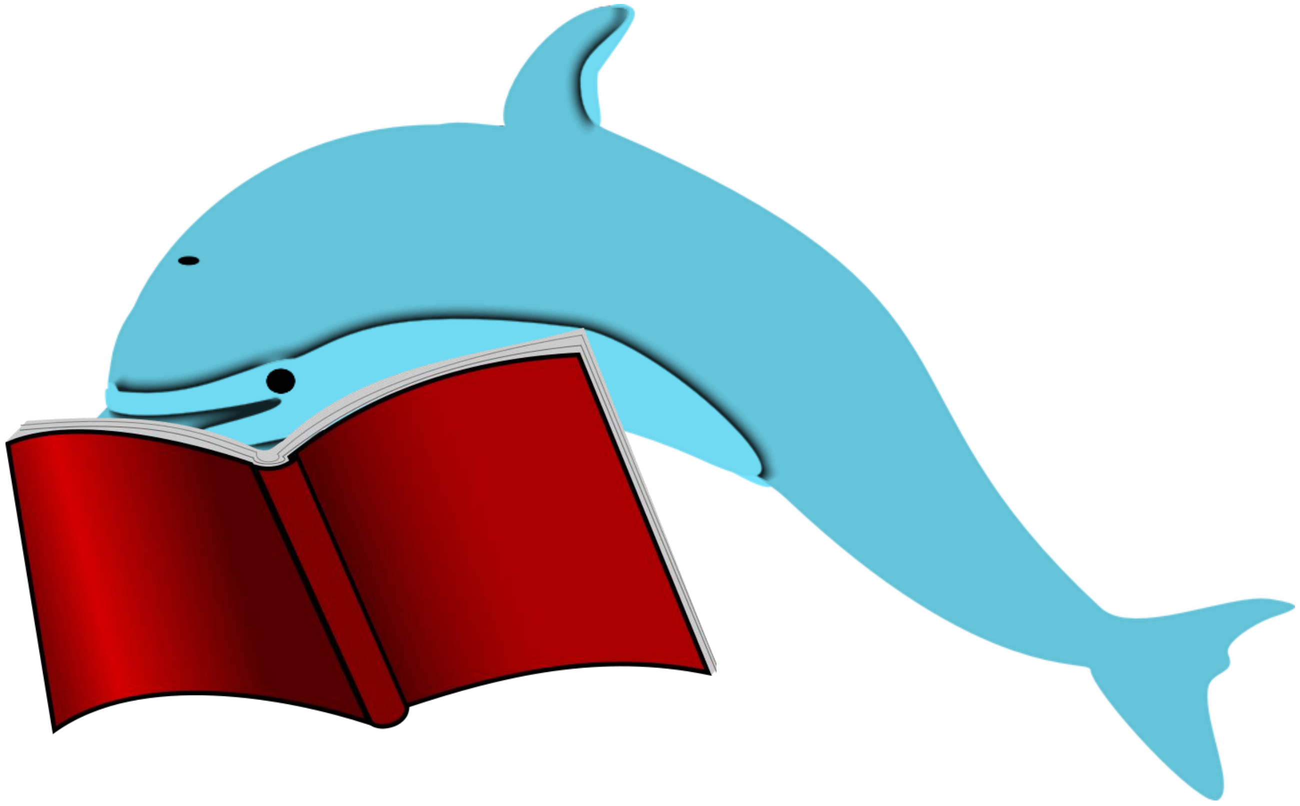
∀x (¬ Read(x) V Literate(x))

∀y (¬Dolphin(y) V ¬Literate(y))

∃z Dolphin(z) ∧ Intelligent(z)

Dolphin(A) ∧ Intelligent(A)

∀t ¬ Intelligent(t) V Read(t)



Conversion to CNF

Step 5: Drop universal quantifiers

KB ∧ ¬ α :

∀x (¬ Read(x) V Literate(x))

¬ Read(x) V Literate(x)

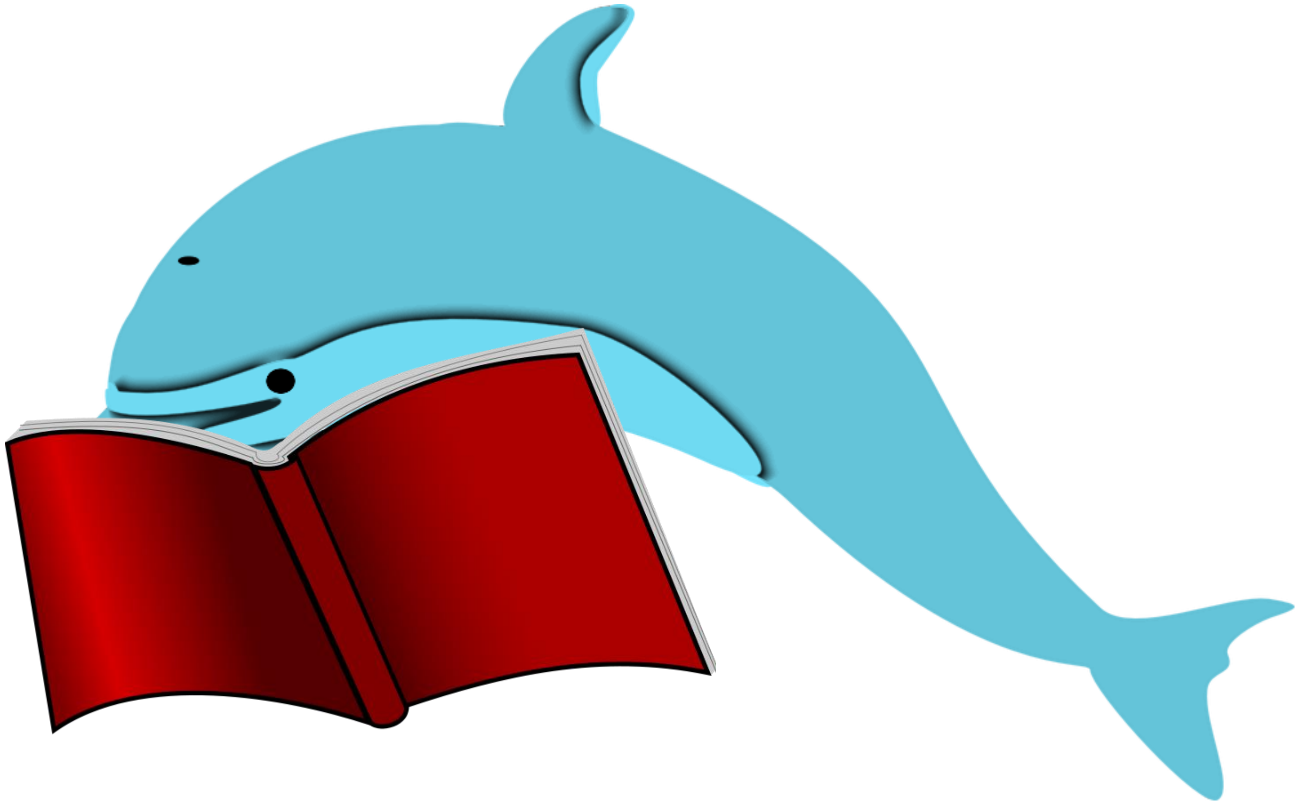
∀y (¬Dolphin(y) V ¬Literate(y))

¬Dolphin(y) V ¬Literate(y)

Dolphin(A) ∧ Intelligent(A)

∀t ¬ Intelligent(t) V Read(t)

¬ Intelligent(t) V Read(t)



Resolution

KB ∧ ¬ α :

(¬ Read(x) V Literate(x)) ∧ (¬Dolphin(y) V ¬Literate(y)) ∧ (Dolphin(A) ∧ Intelligent(A)) ∧ (¬ Intelligent(t) V Read(t))

(¬ R(x) V L(x)) ∧ (¬D(y) V ¬L(y)) ∧ D(A) ∧ I(A) ∧ (¬ I(t) V R(t))

¬ R(x) V L(x))	¬D(y) V ¬L(y)	D(A)	I(A)	¬ I(t) V R(t)

Resolution

$(\neg R(x) \lor L(x)) \land (\neg D(y) \lor \neg L(y)) \land D(A) \land I(A) \land (\neg I(t) \lor R(t))$

$\neg R(x) \lor L(x))$	$\neg D(y) \lor \neg L(y)$	$D(A)$	$I(A)$	$\neg I(t) \lor R(t)$

Resolution

$(\neg R(x) \lor L(x)) \land (\neg D(y) \lor \neg L(y)) \land D(A) \land I(A) \land (\neg I(t) \lor R(t))$

$\neg R(x) \lor L(x))$	$\neg D(y) \lor \neg L(y)$	$D(A)$	$I(A)$	$\neg I(t) \lor R(t)$

Resolution

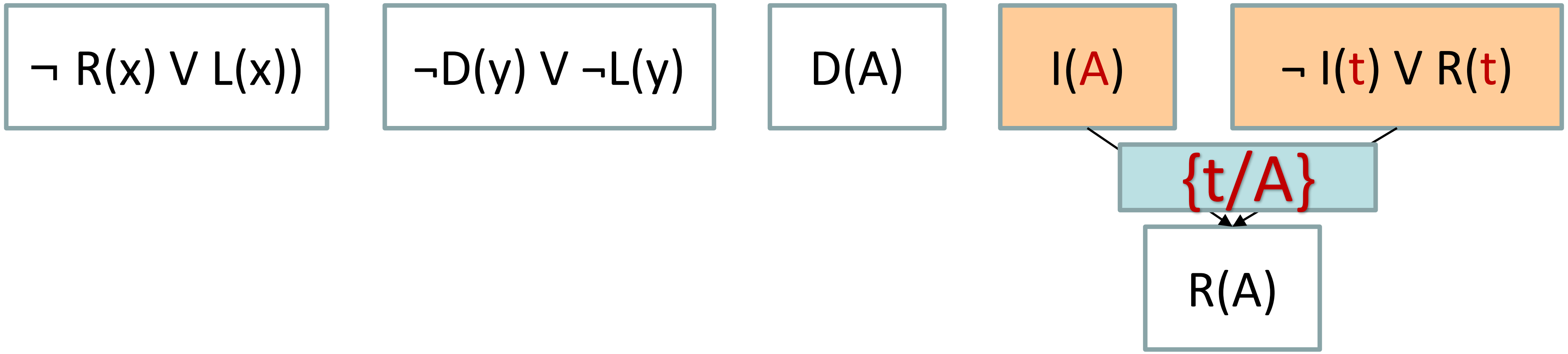
$(\neg R(x) \vee L(x)) \wedge (\neg D(y) \vee \neg L(y)) \wedge D(A) \wedge I(A) \wedge (\neg I(t) \vee R(t))$

$\neg R(x) \vee L(x))$	$\neg D(y) \vee \neg L(y)$	$D(A)$	$I(A)$	$\neg I(t) \vee R(t)$

$\{t/A\}$

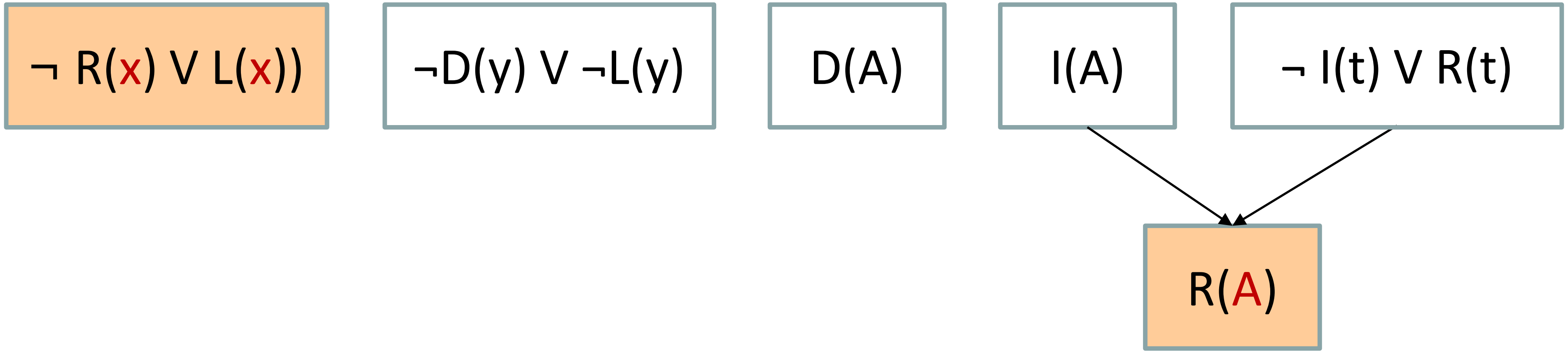
Resolution

$(\neg R(x) \lor L(x)) \land (\neg D(y) \lor \neg L(y)) \land D(A) \land I(A) \land (\neg I(t) \lor R(t))$



Resolution

$(\neg R(x) \vee L(x)) \wedge (\neg D(y) \vee \neg L(y)) \wedge D(A) \wedge I(A) \wedge (\neg I(t) \vee R(t))$



Resolution

$(\neg R(x) \lor L(x)) \land (\neg D(y) \lor \neg L(y)) \land D(A) \land I(A) \land (\neg I(t) \lor R(t))$

$\neg D(y) \lor \neg L(y)$

$D(A)$

$I(A)$

$\neg I(t) \lor R(t)$

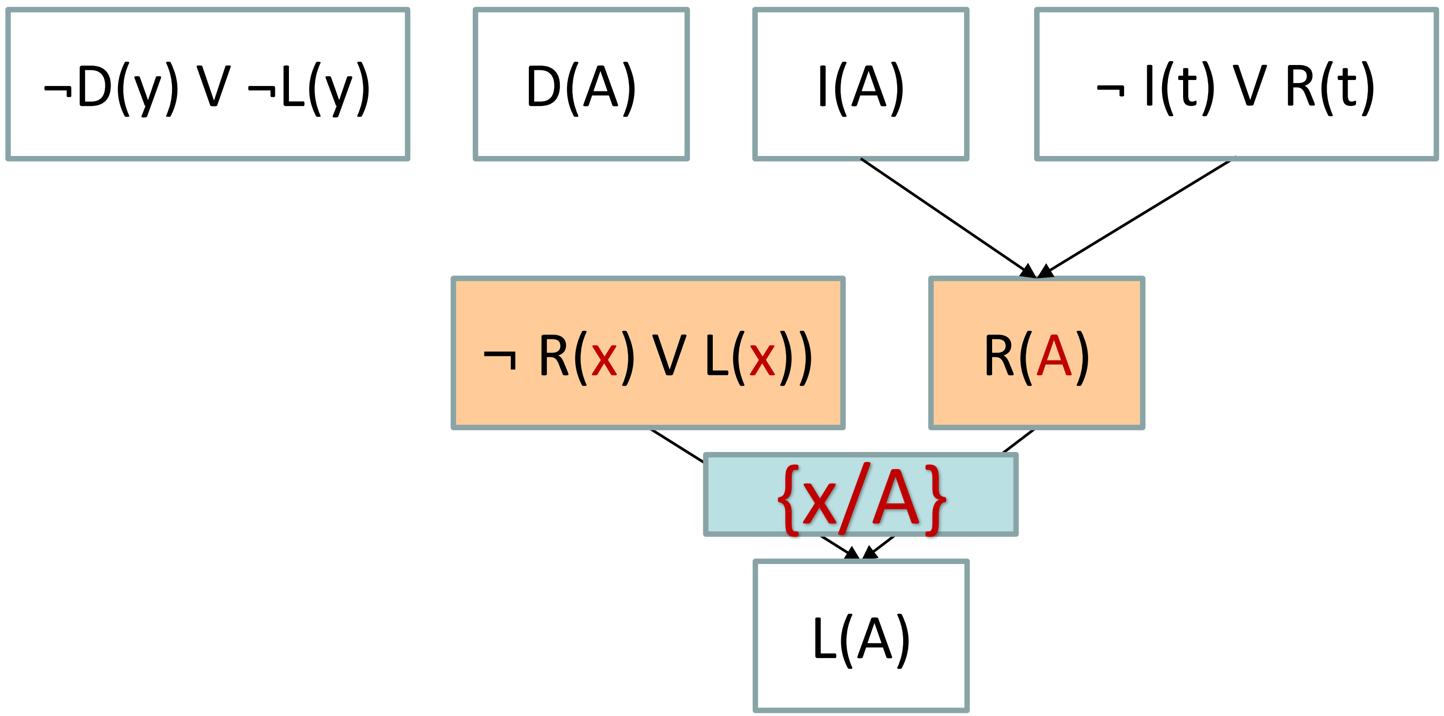
$\neg R(x) \lor L(x)$

$R(A)$

$\{x/A\}$

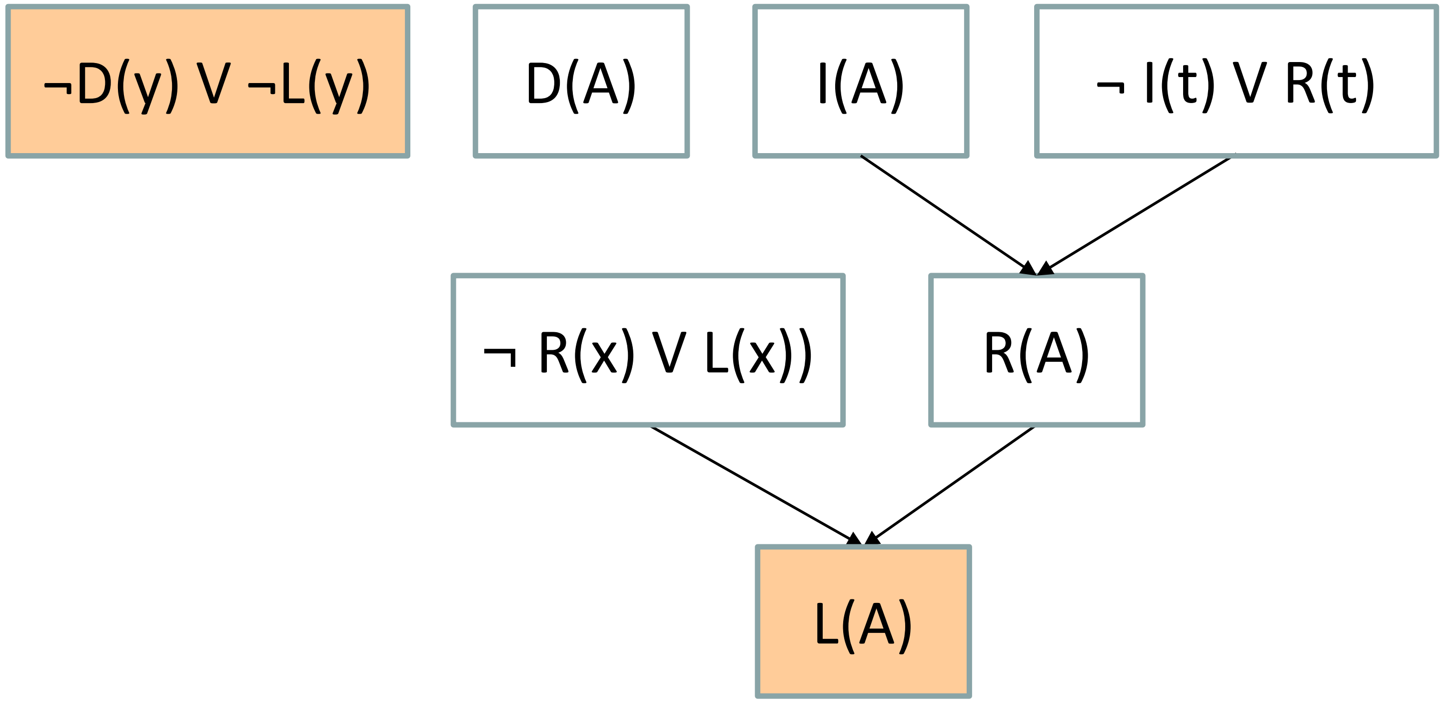
Resolution

$(\neg R(x) \lor L(x)) \land (\neg D(y) \lor \neg L(y)) \land D(A) \land I(A) \land (\neg I(t) \lor R(t))$



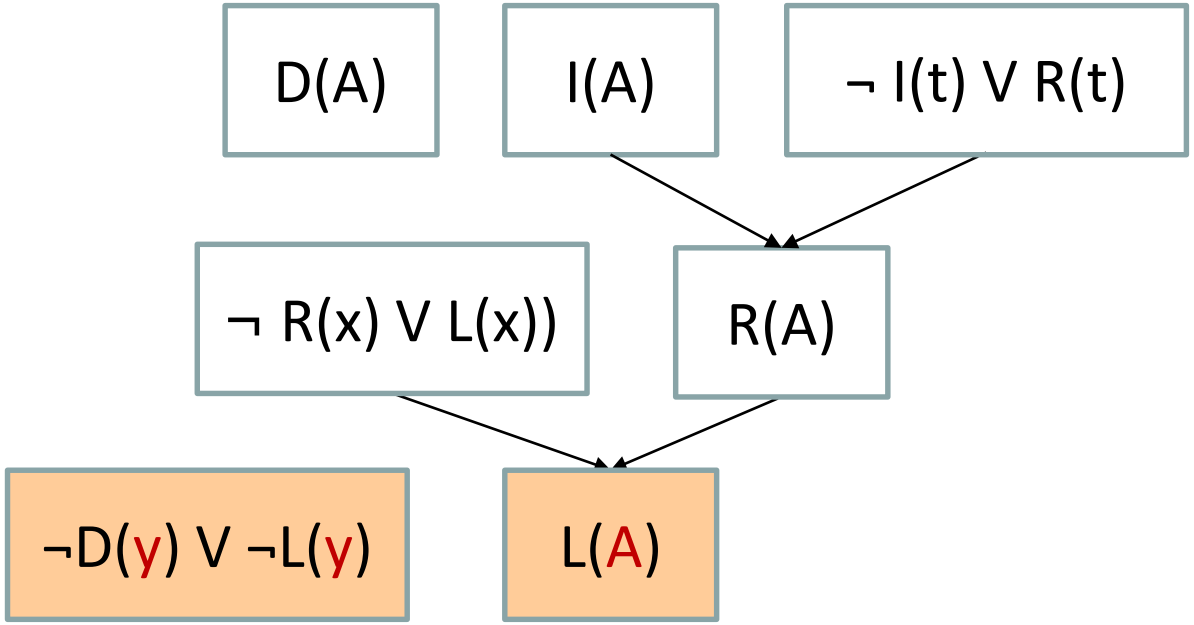
Resolution

$(\neg R(x) \vee L(x)) \wedge (\neg D(y) \vee \neg L(y)) \wedge D(A) \wedge I(A) \wedge (\neg I(t) \vee R(t))$



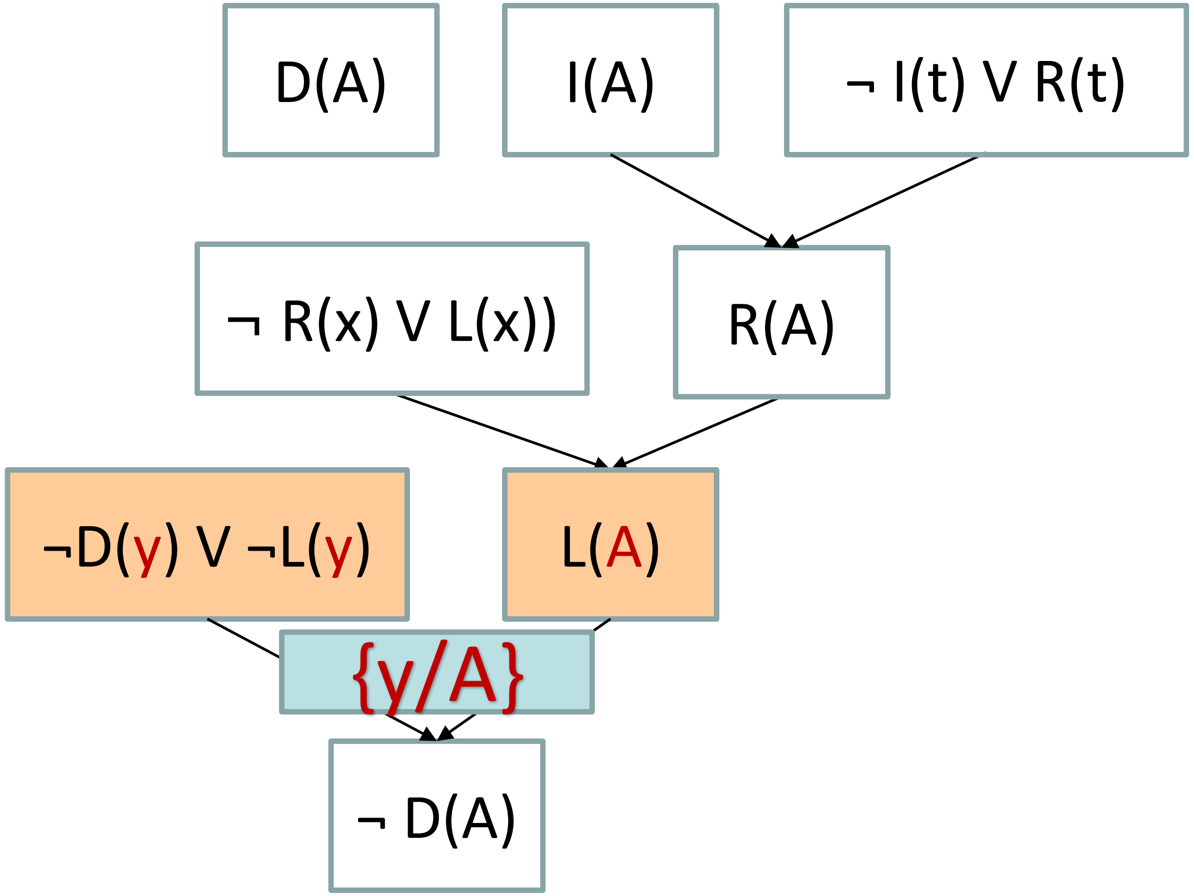
Resolution

$(\neg R(x) \vee L(x)) \wedge (\neg D(y) \vee \neg L(y)) \wedge D(A) \wedge I(A) \wedge (\neg I(t) \vee R(t))$



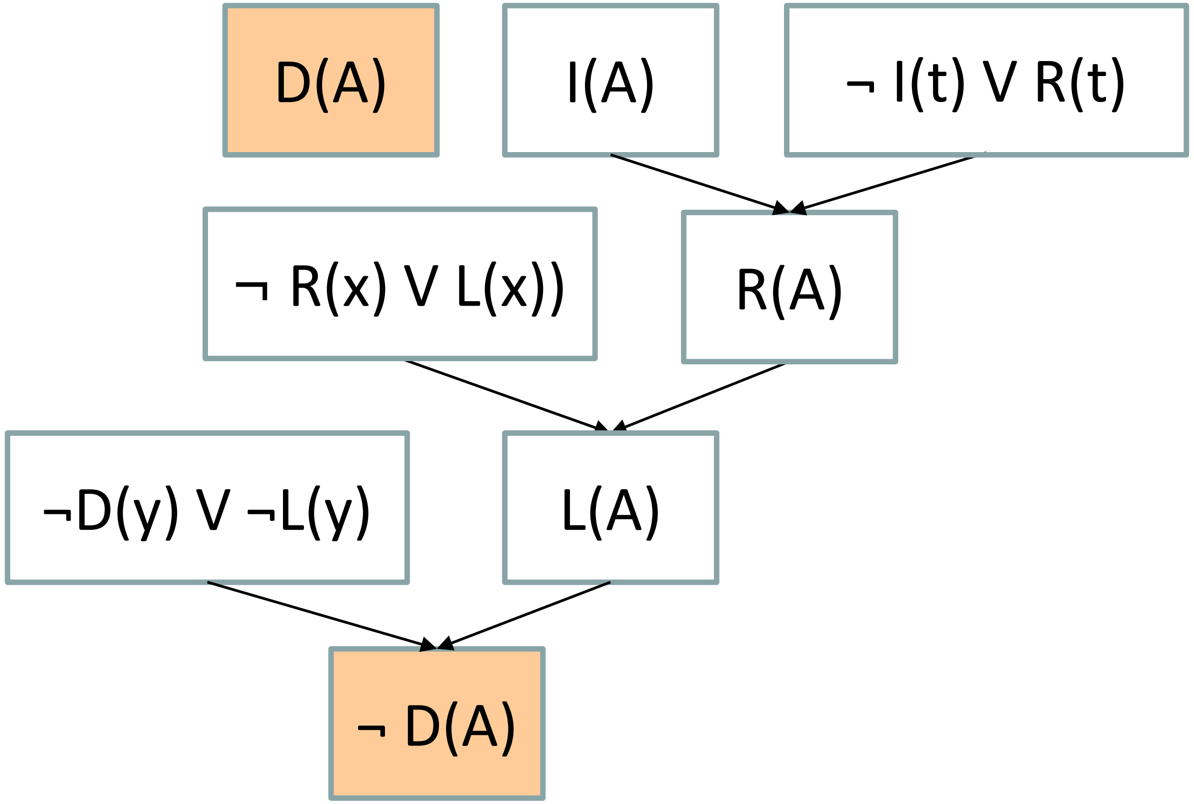
Resolution

$(\neg R(x) \lor L(x)) \land (\neg D(y) \lor \neg L(y)) \land D(A) \land I(A) \land (\neg I(t) \lor R(t))$



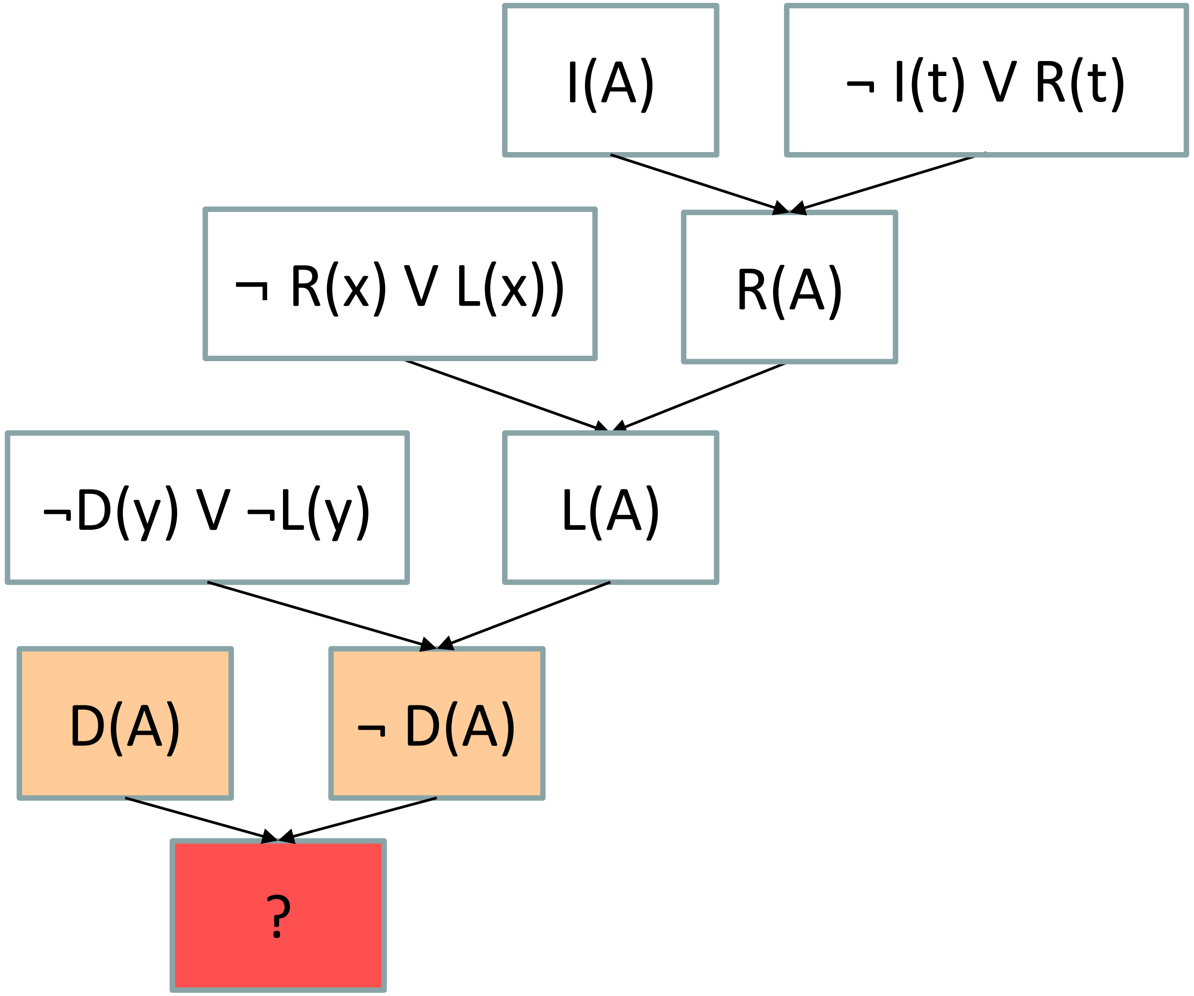
Resolution

$(\neg R(x) \vee L(x)) \wedge (\neg D(y) \vee \neg L(y)) \wedge D(A) \wedge I(A) \wedge (\neg I(t) \vee R(t))$



Resolution

$(\neg R(x) \lor L(x)) \land (\neg D(y) \lor \neg L(y)) \land D(A) \land I(A) \land (\neg I(t) \lor R(t))$



Monotonicity

First-order logic is monotonic.

Our knowledge base can only grow as new information is added.

Formally: if KB $\models \alpha$ then KB $\land\ \beta \models \alpha$

Non Monotonic Reasoning

In practice, our reasoning processes are often non-monotonic.

We hold some default beliefs about the world.

These beliefs are 'almost always' true, with a few exceptions.

We jump to conclusions.

As new evidence arrives, we revise our beliefs.

Default Reasoning

We see a car, we believe it has 4 wheels, even though we can't see them all.

The conclusion is reached by default.

If we see a wheel rolling down the street, then we revise our belief. We retract the earlier default conclusion. We now believe the car has 3 wheels.

Our reasoning here is non-monotonic. It shrank when new evidence came in.

Default and Exceptions

Most birds fly except for penguins, ostriches, the Maltese falcon etc.

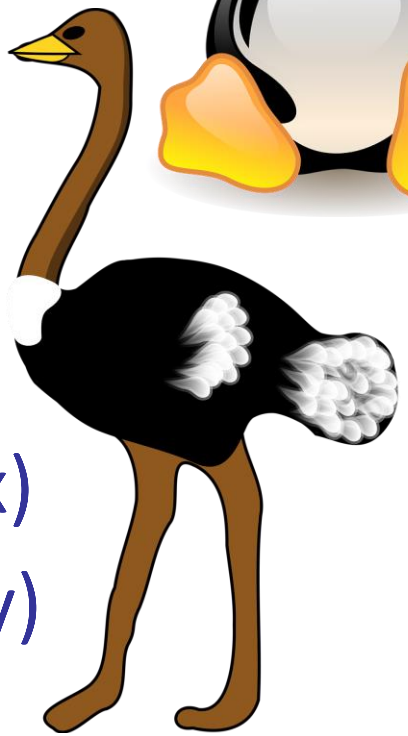
Given a particular bird, we conclude that it flies unless we happen to know that it satisfies one of these exceptions.

How do we represent that?

In first-order logic:

$$\forall x\ Bird(x) \wedge \neg Penguin(x) \wedge \neg Ostrich(x) \wedge \neg \ldots \Rightarrow Flies(x)$$

What can we conclude if all we know is: Bird(Tweety)



Default Logic

Default logic offers one way to formalize this default reasoning (AIMA 12.6).

We can write a default rule to generate contingent conclusions.

Bird(x) : Flies(x) / Flies(x)

If Bird(x) is true and if Flies(x) is consistent with our knowledge base, then Flies(x) may be concluded by default.

Default Logic

Bird(x) : Flies(x) / Flies(x)

If Bird(x) is true and if Flies(x) is consistent with our knowledge base, then Flies(x) may be concluded by default.

Let's say our KB contains the following facts:

Bird(Peter) ∧ Bird(Olga) ∧ Bird(Tweety)

Penguin(x) ⇒ ¬Flies(x)

Penguin(Peter)

Ostrich(Olga)

What can we conclude by default?

A. Only Tweety can fly
B. Tweety, Olga and Peter can fly
C. Only Tweety and Olga can fly
D. No one can fly
E. Only Peter and Olga can fly

Default Logic

In general, a default rule is written as:

$P : J_1 , \ldots, J_n / C$

P is the prerequisite

C is the conclusion

$J_1 , \ldots, J_n$ are the justifications; if any of them are proven false the conclusion cannot be drawn.

Default Logic & the Nixon Diamond

Richard Nixon was both a Quaker and a Republican.

$Quaker(Nixon) \wedge Republican(Nixon)$

Quakers are by default pacifists.

$Quaker(x) : Pacifist(x) / Pacifist(x)$

Republicans are by default not pacifists.

$Republican(x) : \neg Pacifist(x) / \neg Pacifist(x)$

What can we conclude?

Is Richard Nixon a pacifist?

A. $Pacifist(Nixon)$
B. $\neg Pacifist(Nixon)$

Default Logic & the Nixon Diamond

Richard Nixon was both a Quaker and a Republican.

$Quaker(Nixon) \wedge Republican(Nixon)$

Quakers are by default pacifists.

$Quaker(x) : Pacifist(x) \,/\, Pacifist(x)$

Republicans are by default not pacifists.

$Republican(x) : \neg Pacifist(x) \,/\, \neg Pacifist(x)$

The conclusion we draw will depend on the order in which we apply the default rules.

Planning

Planning

Today:

- What is classical planning?
- The frame problem

Next week:

- Planning Domain Definition Language (PDDL)

Planning

- Planning: coming up with a plan of actions to achieve some goals
- Planning is a central part of AI
- We have already seen a search based planning agent:
  - work in deterministic, static and fully observable environments
  - atomic representations of states
  - need domain specific knowledge (heuristic)
- Classical Planning
  - work in deterministic, static and fully observable environments
  - factored representation of states

Planning

To formulate a planning problem, we need to define:

- Initial state
- Actions available
- Results of the actions
- Goal state

In search problems, the result of an action was a single, atomic state.

In classical planning, a single state may be factored into many variables.

The Frame Problem

Do we need to specify the result of an action on ALL the state variables?

The Frame Problem

Let’s take as an example this classroom. To represent the current state, we can say that:

- The lights are on
- The door is open
- The window is closed

Now let me close the door. What are the results of my action?

✓ The door is closed

? The lights are on

? The window is closed

The Frame Problem

The frame problem is the challenge of representing the effects of an action without having to represent explicitly a large number of intuitively obvious non-effects.

How do we capture what is relevant and what is not relevant?

How do we represent what changes without representing what stays the same.