Knowledge Representation and Reasoning

Introduction and Motivation

Maurice Pagnucco
School of Computer Sc. & Eng.
University of New South Wales
NSW 2052, AUSTRALIA

morri@cse.unsw.edu.au


Knowledge Representation and Reasoning

■ What is (the nature of) knowledge?
■ How can we represent what we know?
■ How can we use the representation to infer new knowledge?

Knowledge Representation and Reasoning

- Foundations
  1. Introduction to KRR
  2. Nonmonotonic Reasoning
  3. Reasoning about Action
  4. Belief Revision

Outline

- What is knowledge?
- Representation and Reasoning
- Why Knowledge?
- Why Representation?
- Why Knowledge Representation?
- Advantages of Knowledge Representation
- Forms of Knowledge Representation
What is Knowledge

- “John knows that…”
  - the ‘…” are replaced by a proposition
  - proposition can be true/false

- Other types of knowledge:
  - know how, know who, know what, know when, …
  - sensorimotor: riding a bike
  - affective: deep understanding

- Belief is similar but may not necessarily be true

- Note: we do not distinguish between knowledge and belief

- Main idea: take world to be one way and not another

Representation

- Symbols stand for things in the world
  - + → first aid
  - ! → restaurant
  - “John” → John
  - “John loves Mary” → the proposition that John loves Mary

- Knowledge representation
  - symbolic encoding of believed propositions
Reasoning

- Manipulation of symbols that encode propositions to produce representations of new propositions
- Analogy: arithmetic
  
  "1011" + "10" → "1101"
  ↓  ↓  ↓
  eleven  two  thirteen

  "John is Mary’s father" → “John is an adult male”
  ↓  ↓

Why Knowledge?

- For systems that are reasonably complex it is often useful to describe that system in terms of beliefs, goals, fears and intentions
  
  e.g., chess-playing program
  “because program believed that its queen was in danger but still wanted to control centre of chess board”

  sometimes more useful than describing actual technique:
  “because evaluation using minimax procedure returned value of 7 for this position”

- Intentional stance (Daniel Dennet)

- However is KR just a convenient way of describing complex systems?
  
  anthropomorphising can be inappropriate
  
  …and misleading
Why Representation?

- Intentional stance says nothing about what is and is not represented symbolically
- Knowledge Representation Hypothesis (Brian Smith)
  Any mechanically embodied intelligent process will be comprised of structural ingredients that a) we as external observers naturally take to represent a propositional account of the knowledge that the overall process exhibits, and b) independent of such external semantic attribution, play a formal but causal and essential role in engendering the behaviour that manifests that knowledge
- In other words, existence of structures that
  - can be interpreted propositionally
  - determine how the system behaves
- Knowledge-based system: a system designed in accordance with these principles

Example

- Contrast:
  ```prolog
  printColour(snow) :- !, write("It’s white.").
  printColour(grass) :- !, write("It’s green.").
  printColour(sky) :- !, write("It’s yellow.").
  printColour(X) :- write("Beats me.").
  ```
- with:
  ```prolog
  printColour(X) :- colour(X,Y), !,
                  write("It’s "), write(Y), write(".").
  printColour(X) :- write("Beats me.").
  colour(snow,white).
  colour(sky,yellow).
  colour(X,Y) :- madeof(X,Z), colour(Z,Y).
  madeof(grass,vegetation).
  colour(vegetation,green).
  ```
Example

- Both examples can be described intentionally
- Second example has a separate collection of symbolic structures (like the KR hypothesis)
- It is a knowledge-based system

KR in AI

- Much of AI is concerned with building systems that are knowledge-based
  - natural language understanding
  - planning and scheduling
  - diagnosis
  - “expert systems”
- Some to a certain extent
  - game-playing
  - vision
- Some to a lesser extent
  - speech recognition
  - motor control
Why KR?

- Why not “compile out” knowledge into specialised procedures?
  - distribute KB to procedures that need it
  - usually achieves better performance

- Don’t think. Just do it!
  - riding a bike
  - driving a car
  - playing soccer
  - playing chess?
  - doing math?
  - staying alive??

- Skills (Hubert Dreyfus)
  - novices think; experts react

Advantages of KR

- Knowledge-based system most suitable for open-ended tasks

- Good for
  - explanation and justification
  - “informability”: debugging the KB
  - “extensibility”: new relations
  - new applications

- Hallmark of a knowledge-based system: ability to be told facts about world and adjust behaviour accordingly

- “Cognitive penetrability” (Zenon Plylyshyn)
  - actions conditioned by what is currently believed (e.g., don’t leave the room when alarm sounds if you believe alarm is being tested)
Advantages of Reasoning

- Want knowledge to affect action
  - **not** do action $A$ if sentence $P$ is in KB
  - **but** do action $A$ if world believed in satisfies $P$

- Difference
  - $P$ may not be explicitly represented
  - need to apply what is known to particulars of given situation

- Patient $x$ is allergic to medication $m$
  - Anybody allergic to medication $m$ is also allergic to $m'$
  - Is it ok to prescribe $m'$ for $x$?

- Usually need more than just DB-style retrieval of facts in KB

Forms of KRR

- Many forms of knowledge representation and reasoning have been proposed and studied
- A large number of these are logic-based or closely related to logic
- In the following slides we will briefly list some approaches
Inheritance Networks

Frames

(Trip18
  <IS-A Trip>
  <firstStep TravelStep18-1>)
(Trip18-1
  <IS-A TravelStep>
  <beginDate 12/21/98>
  <endDate 12/21/98>
  <means>
  <origin>
  <destination Toronto>
  <nextStep>
  <previousStep>
  <departureTime>
  <arrivalTime>)
Frames

(Trip <totalCost
  [IF-NEEDED
   {let x $\leftarrow$ SELF.firstStep;
    let result $\leftarrow$ 0;
    repeat
    {if x.nextStep then
     {result $\leftarrow$ result + x.cost + x.destinationLodgingStay.c
      x $\leftarrow$ x.nextStep}
    else return result+x.cost}]}>)

Uncertainty

- Probabilistic vs. possibilistic
- Bayesian networks/belief networks
- Dempster-Shafer theory of evidence
- Fuzzy logic
Nonmonotonic Reasoning

- Classical logics obey property of monotonicity
  \[ \text{If } \Delta \subseteq \Gamma, \text{ then } Cn(\Delta) \subseteq Cn(\Gamma) \]
- Nonmonotonic logics try to capture “commonsense” reasoning
- e.g., “birds usually fly”, “emus normally don’t fly”
- Next lecture will investigate some forms of nonmonotonic reasoning

Description Logic

- Object-centred representation and reasoning
- Concepts: types, categories
- Roles: properties, descriptions
  RED-BORDEAUX-WINE \iff
  (AND WINE
  (FILLS color Red)
  (FILLS region Bordeaux)
  (FILLS sugar-content Dry))
- Form basis of Semantic Web
Conclusion

- Tradeoff between expressiveness of knowledge representation and computational effort required to realise correct reasoning in that representation
- Less expressive KR often means more tractable reasoning
- Can also look to “approximative” inference
Maurice Pagnucco

KRR Program, National ICT Australia and
ARC Centre of Excellence for Autonomous Systems
School of Computer Science and Engineering
The University of New South Wales
Sydney, NSW, 2052

January 4, 2009

Nonmonotonic reasoning is an attempt to capture a form of
commonsense reasoning

In classical logic the more facts (premises) we have, the
more conclusions we can draw

This property is known as Monotonicity

\[ \text{If } \Delta \subseteq \Gamma, \text{ then } Cn(\Delta) \subseteq Cn(\Gamma) \]

(where \( Cn \) denotes classical consequence)

However, the previous example shows that we often do not
reason in this manner

Might a nonmonotonic logic—one that does not satisfy the
Monotonicity property—provide a more effective way of
reasoning?
Why Nonmonotonicity?

Problems with the classical approach to consequence
- It is usually not possible to write down all we would like to say about a domain
- Inferences in classical logic simply make implicit knowledge explicit; we would also like to reason with tentative statements
- Sometimes we would like to represent knowledge about something that is not entirely true or false; uncertain knowledge
- Nonmonotonic reasoning is concerned with getting around these shortcomings

Nonmonotonicity

Classical logic satisfies the following property
- Monotonicity: If \( \Delta \subseteq \Gamma \), then \( Cn(\Delta) \subseteq Cn(\Gamma) \) (equivalently, \( \Gamma \vdash \phi \) implies \( \Gamma \cup \Delta \vdash \phi \))
- However, we often draw conclusions based on ‘what is normally the case’ or ‘true by default’
- More information can lead us to retract previous conclusions
- We shall adopt the following notation
  - \( \vdash \) classical consequence relation
  - \( \not\vdash \) nonmonotonic consequence relation
Outline
Nonmonotonicity  Closed World Assumption  Predicate Completion  Circumscription  Default Logic  Nonmonotonic Consequence

Example

Suppose I tell you ‘Tweety is a bird’
You might conclude ‘Tweety flies’
I then tell you ‘Tweety is an emu’
You conclude ‘Tweety does not fly’

bird(Tweety) \models flies(Tweety)
bird(Tweety) \land emu(Tweety) \models \neg flies(Tweety)

The Closed World Assumption

- A **complete** theory is one in which for every ground atom in the language, the atom or its negation appears in the theory
- The **closed world assumption** (CWA) completes a base (non-closed) set of formulae by including the negation of a ground atom whenever the atom does not follow from the base
- In other words, if we have no evidence as to the truth of (ground atom) $P$, we assume that it is false
- Given a base set of formulae $\Delta$ we first calculate the assumption set
  \[ \neg P \in \Delta_{asm} \text{ if for ground atom } P, \Delta \not\models P \]
- $CWA(\Delta) = Cn(\Delta \cup \Delta_{asm})$

Example

$\Delta = \{ P(a), P(b), P(a) \rightarrow Q(a) \}$
$\Delta_{asm} = \{ \neg Q(b) \}$

**Theorem:** The CWA applied to a consistent set of formulae $\Delta$ is inconsistent iff there are positive ground literals $L_1, \ldots, L_n$ such that $\Delta \models L_1 \lor \ldots \lor L_n$ but $\Delta \not\models L_i$ for $i = 1, \ldots, n$.

- Note that in the example above we limited our attention to the object constants that appeared in $\Delta$ however the language could contain other constants. This is known as the Domain Closure Assumption (DCA)
- Another common assumption is the Unique-Names Assumption (UNA).

  *If two ground terms can’t be proved equal, assume that they are not.*

Predicate Completion

**Idea:** The only objects that satisfy a predicate are those that must

- For example, suppose we have $P(a)$. Can view this as
  \[ \forall x. \ x = a \rightarrow P(x) \]
  the *if*-half of a definition
- Can add the *only if* part:
  \[ \forall x. \ P(x) \rightarrow x = a \]
- Giving:
  \[ \forall x. \ P(x) \leftrightarrow x = a \]
**Predicate Completion**

- **Definition:** A clause is *solitary* in a predicate $P$ if whenever the clause contains a positive instance of $P$, it contains only one instance of $P$.
- For example, $Q(a) \lor P(a) \lor \neg P(b)$ is not solitary in $P$.
- Completion of a predicate is only defined for sets of clauses solitary in that predicate.

Each clause can be written:

- $\forall y.\ Q_1 \land \ldots \land Q_m \rightarrow P(t)$ ($P$ not contained in $Q_i$)
- $\forall y.\ \forall x.\ (x = t) \land Q_1 \land \ldots \land Q_m \rightarrow P(x)$
- $\forall x.\ (\forall y.\ (x = t) \land Q_1 \land \ldots \land Q_m \rightarrow P(x))$ (normal form of clause)

Doing this to every clause gives us a set of clauses of the form:

- $\forall x.\ E_1 \rightarrow P(x)$
- ...$\forall x.\ E_n \rightarrow P(x)$

Grouping these together we get:

- $\forall x.\ E_1 \lor \ldots \lor E_n \rightarrow P(x)$

Completion becomes: $\forall x.\ P(x) \leftrightarrow E_1 \lor \ldots \lor E_n$

and we can add this to the original set of formulae.

**Example**

Suppose $\Delta = \{\forall x.\ Emu(x) \rightarrow Bird(x),\ Bird(Tweety),\ \neg Emu(Tweety)\}$

We can write this as:

- $\forall x.\ (Emu(x) \lor x = Tweety) \rightarrow Bird(x)$

Predicate completion of $P$ in $\Delta$ becomes $\Delta \cup \{\forall x.\ Bird(x) \rightarrow Emu(x) \lor x = Tweety\}$

**Circumscription**

- **Idea:** Make extension of predicate as small as possible.
- **Example:**

  - $\forall x.\ Bird(x) \land \neg Ab(x) \rightarrow Flies(x)$
  - $Bird(Tweety),\ Bird(Sam),\ Tweety \neq Sam,\ \neg Flies(Sam)$

  Want to be able to conclude $Flies(Tweety)$ but $\neg Flies(Sam)$

  Accept interpretations where $Ab$ predicate is as "small" as possible.

  That is, we minimise abnormality.
Circumscription

- Given interpretations $I_1 = \langle D, I_1 \rangle$, $I_2 = \langle D, I_2 \rangle$, $I_1 \preceq I_2$ iff for every predicate $P \in P$, $I_1[P] \subseteq I_2[P]$.
- $\Gamma \models \text{circ } \phi$ iff for every interpretation $I$ such that $I \models \Gamma$, either $I \models \phi$ or there is a $I' \prec I$ and $I' \models \Gamma$.
- $\phi$ is true in all minimal models
- Now consider
  \begin{align*}
  \forall x. & Bird(x) \land \neg Ab(x) \rightarrow Flies(x) \\
  \forall x. & Emu(x) \rightarrow Bird(x) \land \neg Flies(x) \\
  & Bird(Tweety)
  \end{align*}

Examples

- $W = \emptyset; D = \{ p \mid \frac{0}{p} \}$ — no extensions
- $W = \{ p \lor r \}; D = \{ p, q \mid \frac{p}{q}, \frac{r}{q} \}$ — one extension $\{ p \lor r \}$
- $W = \{ p \lor q \}; D = \{ p, q \mid \frac{p}{p}, \frac{q}{q} \}$ — two extensions $\{ \neg p, p \lor q \}, \{ \neg q, p \lor q \}$
- $W = \{ \text{emu}(Tweety), \forall x. \text{emu}(x) \rightarrow \text{bird}(x) \}; D = \{ \text{bird}(x) \mid \text{flies}(x) \}$ — one extension $\{ \text{emu}(Tweety), \forall x. \text{emu}(x) \rightarrow \text{bird}(x) \}$
- What if we add $\exists x. \text{emu}(x) \land \neg \text{flies}(x)$?
- Poole (1988) achieves a similar effect (but not quite as general) by changing the way the underlying logic is used rather than introducing a new element into the syntax

Default Theories—Properties

Observation: Every normal default theory (default rules are all normal) has an extension

Observation: If a normal default theory has several extensions, they are mutually inconsistent

Observation: A default theory has an inconsistent extension iff $D$ is inconsistent

Theorem: (Semi-monotonicity)

Given two normal default theories $\langle D, W \rangle$ and $\langle D', W \rangle$ such that $D \subseteq D'$ then, for any extension $\epsilon(D, W)$ there is an extension $\epsilon(D', W)$ where $\epsilon(D, W) \subseteq \epsilon(D', W)$ (The addition of normal default rules does not lead to the retraction of consequences.)
Nonmonotonic Consequence

- Abstract study and analysis of nonmonotonic consequence relation $\not\models$ in terms of general properties Kraus, Lehmann and Magidor (1991)

- Some common properties include:
  
  **Supraclassicality** If $\models \phi \implies \psi$, then $\not\models \phi \not\models \psi$
  
  **Left Logical Equivalence** If $\models \phi \iff \psi$ and $\not\models \phi \not\models \chi$, then $\not\models \psi \not\models \chi$

  **Right Weakening** If $\models \phi \implies \chi$ and $\not\models \psi \not\models \phi$, then $\not\models \phi \not\models \chi$

  **And** If $\not\models \phi \not\models \psi$ and $\not\models \phi \not\models \chi$, then $\not\models \phi \not\models \psi \wedge \chi$

- Plus many more!

KLM Systems

- Kraus, Lehman and Magidor (1991) study various classes of nonmonotonic consequence relations

  - Cumulative-Loop
  - Preferential
  - Monotonic
  - Cumulative
  - Cumulative-Monotonic

- This has been extended since. A good reference for this line of work is Schlechta (1997)

Summary

- Nonmonotonic reasoning attempts to capture a form of commonsense reasoning

- Nonmonotonic reasoning often deals with inferences based on defaults or 'what is usually the case'

- Belief change and nonmonotonic reasoning: two sides of the same coin?

- Can introduce abstract study of nonmonotonic consequence relations in same way as we study classical consequence relations

- Similar links exist with conditionals

- One area where nonmonotonic reasoning is important is reasoning about action (dynamic systems)
Reasoning About Actions

- One method to reason about action is to simply change the agent's knowledge base.
- Erase some sentence(s) that should no longer be true and add sentences that will now be true (i.e., after performing action).
- However, we can only answer questions about the current state.
- It will not be possible to reason about past or future states.
- On the other hand, if all we want to do is reason about which actions to perform, this may be a viable approach (may return to this later).
Modelling Domains and Actions

- Aspects we need to consider:
  - The state of the world
  - Actions that change state of the world and what changes they effect
  - Constraints on legal scenarios (won’t deal much with these in this lecture)
  - Can you think of anything else?

Situation Calculus

- The situation calculus is a way of describing change in first-order logic
- In simple terms it may be viewed as a dialect of FOL
  - Terms
    - actions
    - situations
  - Fluents—predicates or functions whose values may vary

State of the World

- Method 1:
  - \( \text{on}(C, A, S_1) \)
  - \( \text{on}(A, \text{Table}, S_1) \)
  - \( \text{on}(B, \text{Table}, S_1) \)
  - \( \text{clear}(B, S_1) \)
  - \( \text{clear}(C, S_1) \)
  - **Note**: we reify states (i.e., make them entities in our formalisation)

- Another common way using the situation calculus is as follows
  - Method 2:
    - \( \text{holds}(\text{on}(C, A), S_1) \)
    - \( \text{holds}(\text{on}(A, \text{Table}), S_1) \)
    - \( \text{holds}(\text{on}(B, \text{Table}), S_1) \)
    - \( \text{holds}(\text{clear}(B), S_1) \)

Actions

- Actions are named
  - \( \text{put}(x, y) \) — put object \( x \) on top of object \( y \)
  - \( \text{move}(x, y, z) \) — move block \( x \) from \( y \) to \( z \)
  - \( \text{clear}(x) \) — clear \( x \)
### Situations

- **Situation** — a snapshot of the world at a particular point in time
- Alternate view — world histories
  - $S_0/init$ — initial situation (no actions have been performed)
  - $do(a, s)$ — situation resulting from performing action $a$ in situation $s$
- For example, $do(put(A, B), do(put(B, C), S_0))$
  
  Situation resulting from putting block $B$ on block $C$ in the initial situation and then placing block $A$ on block $B$

### Preconditions

- Predicates and functions whose values may vary from situation to situation
- For example,
  $$\neg Broken(x, s) \land Broken(x, do(drop(r, x), s))$$
- Special predicate $Poss(a, s)$ denotes that action $a$ may be performed in state $s$
  - For example, $Poss(pickup(r, x), s) \equiv \forall z \neg Holding(r, z, s) \land \neg Heavy(x) \land NextTo(r, x, s)$

### Effects

- Actions can have **positive effects**
  $$Fragile(x) \supset Broken(x, do(drop(r, x), s))$$
- and **negative effects**
  $$\neg Broken(x, do(repair(r, x), s))$$

### Domain Constraints

- Also known as **state constraints**
  - True at all (legal) states even though they involve state-dependent relations
  
  - $x$ is on the table iff it is not on top of another block
    $$on(x, Table, s) \equiv \neg \exists y (on(x, y, s) \land y \neq Table)$$
  - $x$ is clear iff there is no block on top of it
    $$clear(x, s) \equiv \neg \exists y \ on(y, x, s)$$
  - If $y$ is a block and there is another block on it, then $y$ is not clear
    $$on(x, y, s) \land (y = Table) \rightarrow \neg clear(y, s)$$
  - etc.
The Frame Problem

- Action descriptions are not complete:
  - They describe what changes BUT do not specify what stays the same!
- The (famous) Frame Problem:
  *The problem of characterising those aspects of the state description that are not changed by an action*

  - One solution — Frame Axioms
    - Moving an object does not change its colour
      \[ \text{Colour}(x, c, s) \supset \text{Colour}(x, c, \text{do(put}(x, y), s)) \]
    - Fragile things do not break
      \[ \neg \text{Broken}(x, s) \land (x \neq y \lor \neg \text{Fragile}(x)) \supset \neg \text{Broken}(x, \text{do(drop}(r, y), s)) \]
    - Since actions often leave most fluents unchanged, many frame axioms may be required

What counts as a solution to the frame problem?

- Once we have described the actions of a system, we would like a systematic method for automatically generating frame axioms
- Preferably, the representation should be concise
- Reasons:
  - Require frame axioms for reasoning
  - They are not entailed by other axioms
  - Reduce possibility of errors in determining frame axioms
  - Can easily update frame axioms if additional effects are specified

Qualification Problem

- What are the ramifications (direct and indirect effects) of performing an action
  \[ \neg \text{clear}(b, \text{do(move}(c, a, b), S_0)) \]
- Recent approaches have investigated the use of explicit notions of causality in an attempt to solve this problem efficiently
- What qualifications (preconditions) do we require in specifying actions and their effects
- Trying to specify exactly under which conditions an action has a particular effect is very difficult (in principle, the list of preconditions can be vast)

Projection

- Determining what is true in the situation resulting from the performing of a sequence of actions \( a_1, \ldots, a_n \)
- Suppose we gather all the axioms above in a sentence \( F \). To determine whether a formula \( \phi \) is true after performing the sequence of actions \( a_1, \ldots, a_n \), we need to determine
  \[ \Gamma \models \phi(\text{do}(a_n), \text{do}(a_{n-1}, \ldots, \text{do}(a_1, S_0)\ldots)) \]
 knowledge representation and reasoning


text content here

Legal

- However, we don’t know whether the sequence of actions \( a_1, \ldots, a_n \) can be performed
- A situation is legal iff:
  - \( \text{Legal}(S_0) \) — it is the initial situation
  - \( \text{Legal} \left( \text{do}(a, s) \right) \subseteq \text{Legal}(s) \land \text{Poss}(a, s) \) — it results from performing the action in a legal situation where its precondition is satisfied
- Adding these axioms to \( \Gamma \), we can determine whether a sequence of actions can be performed by showing that they lead to a legal situation
  \( \Gamma \models \text{Legal} \left( \text{do}(a_n, \, \text{do}(a_{n-1}, \ldots, \text{do}(a_1, \, S_0)) \right) \)

Explanation Closure

- Assumption: The previous two formulae characterise the only way in which a fluent may change
- Explanation Closure Axioms
  - \( \neg F(x, s) \land F(x, \, \text{do}(a, s)) \supset P_F(x, \, a, \, s) \)
  - \( F(x, s) \land \neg F(x, \, \text{do}(a, s)) \supset N_F(x, \, a, \, s) \)
- Disguised frame axioms:
  - \( \neg F(x, s) \land \neg P_F(x, \, a, \, s) \supset \neg F(x, \, \text{do}(a, s)) \)
  - \( F(x, s) \land \neg N_F(x, \, a, \, s) \supset \neg F(x, \, \text{do}(a, s)) \)

Normal form for effect axioms

- Given positive effect axioms for fluent \( \text{Broken} \):  
  - \( \text{Fragile}(x) \supset \text{Broken}(x, \, \text{do}(\text{drop}(r, \, x), \, s)) \)
  - \( \text{NextTo}(b, \, x, \, s) \supset \text{Broken}(x, \, \text{do}(\text{explode}(b), \, s)) \)
- Rewrite them:
  - \( \exists r \{ a = \text{drop}(r, \, x) \land \text{Fragile}(x) \} \lor \exists b \{ a = \text{explode}(b) \land \text{NextTo}(b, \, x, \, s) \} \supset \text{Broken}(x, \, \text{do}(a, \, s)) \)
- Negative effect axiom:
  - \( \neg \text{Broken}(x, \, \text{do}(\text{repair}(r, \, x), \, s)) \)
- Rewrite as:
  - \( \exists a = \text{repair}(r, \, x) \supset \neg \text{Broken}(x, \, \text{do}(a, \, s)) \)
- These formulae or of the form:
  - \( P_F(x_1, \ldots, x_n, \, a, \, s) \supset F(x_1, \ldots, x_n, \, a, \, s) \)
  - \( N_F(x_1, \ldots, x_n, \, a, \, s) \supset \neg F(x_1, \ldots, x_n, \, a, \, s) \)

Successor State Axioms

- Additional axioms:
  - Integrity of effect axioms
    - \( \neg \exists x, \, a, \, s \, P_F(x, \, a, \, s) \land N_F(x, \, a, \, s) \)
  - Unique names for actions
    - \( A(x_1, \ldots, x_n) = A(y_1, \ldots, y_n) \supset (x_1 = y_1) \land \ldots \land (x_n = y_n) \)
    - \( A(x_1, \ldots, x_n) \neq B(y_1, \ldots, y_m) \) for distinct \( A \) and \( B \)
  - Together, axioms on last three slides equivalent to successor state axiom for \( F \):
    - \( F(x, \, \text{do}(a, \, s)) \equiv P_F(x, \, a, \, s) \lor (F(x, \, s) \land \neg N_F(x, \, a, \, s)) \)
  - \( \text{Broken}(x, \, \text{do}(a, \, s)) \equiv \exists r \{ a = \text{drop}(r, \, x) \land \text{Fragile}(x) \} \lor \exists b \{ a = \text{explode}(b) \land \text{NextTo}(b, \, x, \, s) \} \lor \text{Broken}(x, \, \text{do}(a, \, s)) \)

Summary

Maurice Pagnucco
Knowledge Representation and Reasoning

Maurice Pagnucco
Knowledge Representation and Reasoning

Maurice Pagnucco
Knowledge Representation and Reasoning

Maurice Pagnucco
Knowledge Representation and Reasoning
What we cannot do

- Explicit time
- Exogenous actions
- Concurrent actions
- Continuous actions
- Complex actions
- ...

Summary

- Reasoning about actions is a very interesting area of artificial intelligence and often makes use of nonmonotonic reasoning techniques
- We have seen that a number of challenging problems arise that we must deal with in order to reason effectively
- One of the problems, however, is the possible proliferation of axioms
- The search continues for a concise solution to the frame problem (and associated problems)
- Other formalisms include the event calculus, A languages, features and fluents, fluent calculus
- Current research: causal approaches, cognitive robotics, planning (an area in its own right)
Belief Change — An example

(Gärdenfors & Rott 1995)

- **Beliefs**
  - The bird caught in the trap is a swan
  - The bird caught in the trap comes from Sweden
  - Sweden is part of Europe
  - All European swans are white

- **Consequences**
  - The bird caught in the trap is white

- **New information**
  - The bird caught in the trap is black

Which sentence(s) would you give up?

Logical considerations alone are not sufficient to answer this question.

- **One possibility**
  - The bird caught in the trap is a swan
  - The bird caught in the trap comes from Sweden
  - Sweden is part of Europe
  - All European swans, except for some of the Swedish, are white
  - The bird caught in the trap is black
Belief Change—The General Idea

Dynamics of epistemic states

However, we are not interested in all such ways of changing beliefs but only those that follow certain principles (i.e., rational belief change)

Rationality Criteria

Principle of Categorial Matching: The representation of a belief state after change should be of the same format as that prior to change

Consistency: Beliefs in belief state should be consistent

Deductive Closure: If the beliefs in a belief state logically entail a sentence $\phi$, then $\phi$ should be included in the state

Principle of Informational Economy: The amount of information lost during change should be kept to a minimum

Preference: Beliefs considered more important or entrenched should be retained in favour of less important ones.

AGM Framework

(Alchourrón, Gärdenfors and Makinson 1985) While there are many frameworks for belief change we concentrate on the AGM here as it is very common in the literature

- Epistemic states: belief sets (closed under $Cn$)
- Epistemic input: formula
- Operators: belief expansion, belief contraction, belief revision
  - Guided by principles of rationality (e.g., minimal change)
  - Characterised by postulates for rational belief change (delineating a class of functions with desirable properties)
  - Constructions: Selection functions, systems of spheres, epistemic entrenchment
Epistemic Attitudes

- An agent can have three attitudes towards a formula $\phi$:
  - $\phi \in K$: agent believes $\phi$
  - $\neg \phi \in K$: agent disbelieves $\phi$
  - $\phi, \neg \phi \not\in K$: agent is indifferent towards $\phi$

- In other models of belief change there may be a much larger number of attitudes (for example, suppose we attach probabilities to beliefs).

Belief Expansion

- Want to add a belief(s) without giving anything up

  - (K+1) For any sentence $\alpha$ and any belief set $K$, $K + \alpha$ is a belief set (closure)
  - (K+2) $\alpha \in K + \alpha$ (success)
  - (K+3) $K \subseteq K + \alpha$ (inclusion)
  - (K+4) If $\alpha \in K$, then $K + \alpha = K$ (vacuity)
  - (K+5) If $K \subseteq H$, then $K + \alpha \subseteq H + \alpha$ (monotonicity)
  - (K+6) For all belief sets $K$ and sentences $\alpha$, $K + \alpha$ is the smallest belief set satisfying (K+1) – (K+5) (minimality)

**Theorem:** The expansion function $+$ satisfies (K+1) – (K+6) iff $K + \alpha = Cn(K \cup \{\alpha\})$.

- There is only one AGM expansion function—classical consequence $Cn$.
- This is not the case for AGM contraction and revision.

Belief Contraction

**Contraction Example**

Suppose $K$ contains the closure of the following formulae:

- $\text{rain} \rightarrow \text{wet\_grass}$
- $\text{rain}$
- $\text{shop\_open}$
- $\text{shop\_open} \rightarrow \text{light\_on}$

Consider possibilities for $K - \text{wet\_grass}$:

- $\text{rain} \rightarrow \text{wet\_grass}$
- $\text{rain}$
- $\text{shop\_open}$
- $\text{shop\_open} \rightarrow \text{light\_on}$
- $\text{shop\_open} \rightarrow \text{light\_on}$
- $\text{shop\_open} \rightarrow \text{light\_on}$
Belief Contraction

Want to give up a belief or suspend judgement; do not want to add any beliefs

(K−1) For any sentence \( \phi \) and any belief set \( K \),
\( K \vdash \phi \) is a belief set
\( K \vdash \phi \subseteq K \)
\( K \vdash \phi = K \)
\( K \vdash \phi \subseteq K \)
\( \forall \phi \text{ then } \phi \not\in K \vdash \phi \)
\( \phi \in K \), then \( K \vdash \psi \subseteq K \)
\( K \vdash \phi \land K \vdash \psi \subseteq K \vdash (\phi \land \psi) \)
\( K \vdash (\phi \lor \psi) \subseteq K \vdash \phi \lor K \vdash \psi \)

In particular, note the difference between

- \( K \vdash (\phi \land \psi) \): need only give up either \( \phi \) or \( \psi \)
- \( K \vdash (\phi \lor \psi) \): must give up both \( \phi \) and \( \psi \)

The following properties follow from the AGM postulates for belief contraction.

1. If \( \phi \in K \), then \( (K \vdash \phi) + \phi \subseteq K \)
2. \( K \vdash \phi = K \cap (K \vdash \phi) + \neg \phi \)
3. \( K \vdash \phi \cap \text{Con}n(\{\phi\}) \subseteq K \vdash (\phi \land \psi) \)
4. Either \( K \vdash (\phi \land \psi) \subseteq K \vdash \phi \) or \( K \vdash (\phi \land \psi) \subseteq K \vdash \psi \)
5. Either \( K \vdash (\phi \land \psi) = K \vdash \phi \) or \( K \vdash (\phi \land \psi) = K \vdash \psi \)
6. If \( \phi \rightarrow \phi \in K \vdash \phi \) and \( \phi \rightarrow \psi \in K \vdash \psi \), then \( K \vdash \phi = K \vdash \psi \)

Belief Contraction

Digression—Commensurability Thesis (Levi 1991)

“Given an initial state of full belief \( K_1 \) and another state of full belief \( K_2 \), there is always a sequence of expansions and contractions, beginning with \( K_1 \), remaining within the state of potential states of full belief and terminating with \( K_2 \).”

Levi Identity: \( K \ast \phi = (K \vdash \neg \phi) + \phi \)
- Given a contraction function we can construct a (associated) revision function: contract anything that would cause the addition of \( \phi \) to lead to inconsistency and then expand by \( \phi \)

Harper Identity: \( K \vdash \phi = K \cap K \ast \neg \phi \)
- Given a revision function we can construct a (associated) contraction function: revise by \( \neg \phi \) (which would remove \( \phi \) if it were currently believed so as to have a consistent revision) and keep those beliefs in \( K \) that were maintained by this revision

Constructing Contraction Functions

- AGM contraction function is simply a mapping from a belief set and a formula to a new belief set that satisfies certain restrictions
- How would we go about “constructing” such a function especially if we wanted to implement one in a computer program?
- Storing all the possible mappings is out of the question!
- There are a number of constructions for contraction functions that we shall investigate
- The first idea is to consider removing just enough formulae from \( K \) so that it no longer implies \( \phi \)

Maurice Pagnucco
UNSW
Knowledge Representation and Reasoning
Maximal Non-implying Subsets

Definition: $K'$ is a maximal subset of $K$ that fails to imply $\phi$ (a $\phi$-remainder) iff
(i) $K' \subseteq K$
(ii) $\phi \notin K'$
(iii) for any $\psi \in K$ such that $\psi \notin K'$, $\psi \rightarrow \phi \in K'$

We denote by $K \perp \phi$ the set of all such maximal non-implying subsets.

Maxichoice Contraction

Idea: select the “best” element from $K \perp \alpha$ (minimal change).

Definition: Let $\gamma$ be an opinionated selection function. A maxichoice contraction function over $K$ may be defined as follows
(Def Max) $K \div \phi = \begin{cases} \gamma(K \perp \phi) & \text{whenever } K \perp \phi \neq \emptyset \\ K & \text{otherwise} \end{cases}$

Theorem: If $\div$ is a maxichoice contraction function over $K$, then it satisfies $(K \div 1) - (K \div 6)$.

Theorem: If a revision function $\ast$ is obtained from a maxichoice contraction function $\div$ via the Levi Identity, then for any $\phi$ such that $\neg \phi \in K$, $K \ast \phi$ is complete.
- Maxichoice doesn’t remove enough

Full Meet Contraction

Idea: All or nothing!

Definition: A full meet contraction over $K$ may be defined as follows
(Def Full) $K \div \phi = \begin{cases} \bigcap(K \perp \phi) & \text{whenever } K \perp \phi \neq \emptyset \\ K & \text{otherwise} \end{cases}$

Theorem: Any full meet contraction function satisfies $(K \div 1) - (K \div 6)$.

Theorem: If a revision function $\ast$ is obtained from a full meet contraction function $\div$ via the Levi Identity, then for any $\phi$ such that $\neg \phi \in K$, $K \ast \phi = \text{Cn}(\phi)$.
- Full meet removes too much
Maximal Non-implying Subsets

Selection Functions — more details

We can define a selection function as follows and then apply restrictions to see what properties result:
Marking-off identity ≤
\[ \gamma(K \perp \phi) = \{ K' \in K \perp \phi : K'' \leq K' \text{ for all } K'' \in K \perp \phi \} \]

Definition: \( \gamma \) is a transitively relational iff it can be defined via a marking-off identity ≤ which is transitive.

Theorem:
For every belief set \( K \), \( \gamma \) is a transitively relational partial meet contraction function iff \( \gamma \) satisfies (\( K \gamma 1 \)) – (\( K \gamma 4 \)).

Withdrawals (Makinson 1986)

Idea: Let's consider functions that don't satisfy (Recovery)
Definition: A function \( \gamma \) is a withdrawal function iff it satisfies postulates (\( K \gamma 1 \)) – (\( K \gamma 4 \)) for contraction over \( K \).

Definition: Two withdrawal functions \( \gamma, \delta \) are revision equivalent iff they generate the same revision function via the Levi Identity.

Theorem: Let \( K \) be any belief set. Then for each withdrawal operation \( \gamma \) on \( K \), there is a unique contraction function \( \delta \) on \( K \) that is revision equivalent to \( \gamma \) and this \( \delta \) is the greatest element of \( \{ \gamma \} \).

Belief Revision

Want to incorporate a belief in a consistent fashion:

(K*1) For any sentence \( \phi \) and any belief set \( K \), \( K + \phi \) is a belief set (closure)
(K*2) \( \phi \in K \) (success)
(K*3) \( K + \phi \subseteq K + \phi \) (inclusion)
(K*4) If \( \neg \phi \not\in K \), then \( K + \phi \subseteq K + \phi \) (preservation)
(K*5) \( K + \phi = K \) if and only if \( \vdash \neg \phi \) (vacuity)
(K*6) If \( \neg \phi \leftrightarrow \psi \), then \( K + \phi = K + \psi \) (extensionality)
(K*7) \( K + \phi \land \psi \subseteq (K + \phi) + \psi \) (super expansion)
(K*8) If \( \neg \psi \not\in K + \phi \), then \( (K + \phi) + \psi \subseteq K + (\phi \land \psi) \) (sub expansion)
**Outline**

**Belief Change**

AGM

Constructions

Summary

---

**Maximal Non-implying Subsets**

**Additional Properties**

1. If $\phi \in K$, then $K * \phi = K$
2. $K * \phi = (K \cap K * \phi) + \phi$
3. $K * \phi = K * \psi$ if and only if $\psi \in K * \phi$ and $\phi \in K * \psi$
4. $K * \phi \cap K * \psi \subseteq K * (\phi \lor \psi)$
5. If $\neg \psi \notin K * (\phi \lor \psi)$, then $K * (\phi \lor \psi) \subseteq K * \psi$
6. $K * (\phi \lor \psi) = K * \phi$ or $K * (\phi \lor \psi) = K * \psi$ or $K * (\phi \lor \psi) = K * \phi \cap K * \psi$

---

**Epistemic Entrenchment**

**Ordering over formulae in $L$**

- Certain beliefs about the world are more important than others when planning future actions, etc.
- $\phi \leq \psi$: $\psi$ is at least as epistemically entrenched as $\phi$
- In contraction, sentences in $K$ with lower entrenchment given up
- Tautologies maximally entrenched, non-beliefs minimally entrenched
  - But what does an epistemic entrenchment relation look like?

---

**Additional Properties**

1. $\phi \leq \psi$ or $\psi \leq \phi$ (connectedness)
2. If $\phi \land \chi \leq \phi$, then $\phi \leq \phi$ or $\chi \leq \phi$
3. $\phi \leq \psi \iff \phi \land \psi < \psi$
4. If $\chi \leq \phi$ and $\chi \leq \psi$, then $\chi \leq \phi \land \psi$
5. If $\phi \leq \psi$, then $\phi \leq \phi \land \psi$
6. $\phi \land \psi = \text{min}(\phi, \psi)$
7. $\phi \lor \psi \geq \text{max}(\phi, \psi)$

(NB: $\phi < \psi \equiv \psi \not\leq \phi$)
### Belief Change via Entrenchment

<table>
<thead>
<tr>
<th>( \phi )</th>
<th>( Y )</th>
<th>( X \wedge Y )</th>
<th>( X \lor Y )</th>
<th>( T )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \phi ) not in ( K )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Gärdenfors and Makinson 1988)

\[ (C \leq) \phi \leq \psi \iff \phi \not\in K \land (\phi \land \psi) \lor \vdash \phi \land \psi \]

- Prefer \( \psi \) to \( \phi \) if we would give up \( \phi \) when given a choice between giving up \( \phi \) or \( \psi \) or if it's not possible to give up either formula.

\[ (C^-) \psi \in K^\phi \iff \psi \in K \text{ and either } \phi < \phi \lor \psi \lor \vdash \phi \]

- Retain belief if it was originally believed and there is “independent evidence” for maintaining it or if it is not possible to remove \( \phi \).

\[ (C^*) \psi \in K^\phi \iff \neg \phi \leq \neg \phi \lor \psi \lor \vdash \neg \phi \]

**Theorem:** If an ordering \( \leq \) satisfies \((EE1) – (EE5)\), then the contraction function which is uniquely determined by \((C^-)\) satisfies \((K^-1) – (K^-8)\) as well as condition \((C \leq)\).

**Theorem:** If a contraction function \( \vdash \) satisfies \((K^-1) – (K^-8)\), then the ordering that is uniquely determined by \((C \leq)\) satisfies \((EE1) – (EE8)\) as well as condition \((C^-)\).

---

### Grove’s Spheres

- Ordering over “possible worlds” (maximally consistent sets of formulae)
- Motivated by Lewis’ sphere semantics for counterfactuals
- System of spheres: sets of possible worlds nested one within the other
- The set of all worlds \( \mathcal{M}_L \) is the outermost (largest) sphere
- \([K]\) is the set of worlds consistent with \( K \); these worlds form the innermost (smallest) sphere
- System of spheres centred on \([K]\)

---

### Systems of Spheres

**Definition:** Let \( S \) be any collection of subsets of \( \mathcal{M}_L \). We call \( S \) a system of spheres, centred on \( X \subseteq \mathcal{M}_L \), if it satisfies the following conditions:

\[ (S1) \] \( S \) is totally ordered by \( \subseteq \); that is, if \( U, \ V \in S \), then \( U \subseteq V \lor V \subseteq U \)

\[ (S2) \] \( X \) is the \( \subseteq \)-minimum of \( S \)

\[ (S3) \] \( \mathcal{M}_L \) is the \( \subseteq \)-maximum of \( S \)

\[ (S4) \] If \( \phi \in L \) and \( \not\vdash \phi \), then there is a smallest sphere in \( S \) intersecting \([\phi]\) (i.e., there is a sphere \( U \in S \) such that \( U \cap [\phi] \neq \emptyset \), and \( V \cap [\phi] \neq \emptyset \) implies \( U \subseteq V \) for all \( V \in S \))

**Lewis’ Limit Assumption**
Belief Change via SOSs

\[ \text{c}_S(\phi) \] — the smallest sphere intersecting \([\phi]\)
\[ f_S(\phi) = \text{c}_S(\phi) \cap [\phi] \] — innermost \(\phi\)-worlds

**Theorem:** Let \(S\) be any system of spheres in \(\mathcal{M}_L\) centred on \([K]\) for some theory \(K \in \mathcal{K}\). If one defines, for any \(\phi \in L\), \(K * \phi\) to be \(\text{th}(f_S(\phi))\), then the axioms (K*1) – (K*8) are satisfied.

**Theorem:** Let \(* : \mathcal{K} \times L \rightarrow \mathcal{K}\) be any function satisfying axioms (K*1) – (K*8). Then for any (fixed) theory \(K\) there is a system of spheres on \(\mathcal{M}_L\), \(S\) say, centred on \([K]\) and satisfying \(K * \phi = \text{th}(f_S(\phi))\) for all \(\phi \in L\).

---

Properties of \(\text{th}\) (Grove 1988)

\[ \text{th} : 2^{\mathcal{M}_L} \rightarrow \mathcal{K} \]

(i) \(\text{th}([K]) = K\) for all belief sets (i.e., theories) \(K\) if the underlying logic is compact

(ii) \(\text{th}(X) \neq K_L\) if and only if \(X\) is nonempty

(iii) For any sentence \(\phi \in L\) and \(X \subseteq \mathcal{M}_L\),
\[ \text{th}(X \cap [\phi]) = \text{Cn}(\text{th}(X) \cup \{\phi\}) \]

(iv) For \(X, X' \subseteq \mathcal{M}_L\), if \(X \subseteq X'\), then \(\text{th}(X') \subseteq \text{th}(X)\)

(v) For \(K, K' \in \mathcal{K}\), if \(K \subseteq K'\), then \([K'] \subseteq [K]\)
(Gärdenfors 1988)
Can translate back and forth from a Systems of Spheres $\mathcal{S}$ and an epistemic entrenchment relation $\leq$ using the following condition:

$$\phi \leq \psi \text{ iff } c(\neg \phi) \subseteq c(\neg \psi)$$
Belief Change and NMR

- Can define $\phi \leadsto \psi$ iff $\psi \in K \ast \phi$ (cf. Ramsey test)
- Translation of AGM revision postulates:
  (K*1) = If $\phi \leadsto \psi_i$ for all $\psi_i \in K$ and $K \vdash \chi$, then $\phi \leadsto \chi$ (Closure)
  (K*2) = $\phi \leadsto \phi$ (Reflexivity)
  (K*3) = If $\phi \leadsto \psi$, then $T \vdash \phi \rightarrow \psi$ (Weak Conditionalisation)
  (K*4) = If $T \vdash \psi \rightarrow \phi$ and $T \vdash \phi \rightarrow \psi$, then $\phi \leadsto \psi$ (Weak Rational Monotony)
  (K*5) = If $\phi \leadsto \bot$, then $\phi \vdash \bot$ (Consistency Preservation)
  (K*6) = Left Logical Equivalence
  (K*7) = If $\phi \land \psi \leadsto \chi$, then $\phi \vdash \psi \rightarrow \chi$ (Conditionalisation)
  (K*8) = If $\psi \rightarrow \phi$ and $\phi \leadsto \chi$, then $\phi \land \psi \leadsto \chi$ (Rational Monotony)
- Note that $\leadsto$ is with respect to a particular $K$
- All these postulates are known in the nonmonotonic

Areas of Study in Belief Change

- Belief bases and computational approaches
- Coherence (how do beliefs ‘cohere’)
- Contraction proposals
- Iterated revision
- Relationships between belief change, nonmonotonic reasoning, conditionals, rational choice, ...
- Non-prioritised belief change
- Abductive belief change
- Reasoning about action and belief change
- Dealing with uncertainty (Spohn, Bayesian networks, ...)