Section 7

Belief Revision

Subsection 7.1

Introduction

Introduction

Nonmonotonic reasoning:
make plausible conjectures in the absence of complete information.

Intelligent systems:

- adapt the knowledge base when new knowledge arrives.
- resolving potential conflicts of old and new knowledge

Belief revision:

- is a process that can be used to modify a knowledge base when new information is acquired
- needs to cope with inconsistencies due to the addition of new knowledge (Recall: anything follows logically from an inconsistent knowledge base; rendering it useless for reasoning.)

Goal of this chapter:
lay down the foundations of principled mechanisms for modifying a knowledge base in a rational and coherent way.

Our abstract setting for belief revision

The setting considered in this chapter:

- agent (or information system)
- knowledge base (KB)
  - represents the collection of facts (/information) the agent believes in
  - represented by a theory T
- some language $\mathcal{L}$ that is
  - a set of closed formulae (sentences) over a fixed signature.
  - used for content of KB and the fact to be added
- one change φ
  one fact at a time is added or retracted
- modifications of the KB are addition or removal of facts.
  It is not the transformation of formulae (e.g. “all birds, but Tweety fly.”)

Considerations for designing change functions

- Belief Revision is a mechanism for modeling rational decisions concerning modifications of the KB.
- When incorporating new knowledge that is inconsistent with the knowledge base, the agent must decide which knowledge to give up.
- A rational agent would modify the KB as little as possible to incorporate new information. Usually there are several options.

Example for different choices for retraction of a fact

Example 7.1 (Tweety again)

Consider the beliefs: "all birds fly" "Tweety is a bird"
Consequence: "Tweety can fly"

New fact: "Tweety cannot fly"

How is the KB to be modified?
Retract "Tweety can fly", but also all beliefs that are in conflict with the new fact.

There are several choices:

1. retract "Tweety is a bird"
2. retract "all birds fly"
3. retract both

Which choice is picked depends on the preferences of the agent!
Agent might be less confident about one fact and prefer to retract that.

$\leadsto$ The choice for the "right" option depends on aspects outside of logic!

We employ a preference ordering of the facts in the KB to realize a choice function.

Principles for modifications

Aim: study rational modifications to KBs guided by
the "Principle of minimal change"

Notions of rationality and minimality are not so easy to capture formally!

Principle of rationality:
only facts are retracted that cause the inconsistencies.

Principle of minimal change:
as much as possible of the underlying information should be kept
(in accordance of the preference relation).

Preference relations capture

- the information content of the KB
- the agent's commitment to this information
- how the information should behave under change

Considering principle of minimal change
—when having different choices for retraction

Reconsider Example 7.1:

Are choices 1. and 2. more minimal than 3.?
—They give up a smaller number of facts!

But: Interdependencies among beliefs might necessitate to give up more than the minimal number of beliefs.

How to measure the magnitude of change?
Is it better to discard one strongly held belief or several weakly held beliefs?

The AGM framework

The basis for most belief revision approaches (or systems) is the well-known AGM frame work

The AGM framework:

- developed by Carlos Alchourrón, Peter Gärdenfors and David Makinson in the early eighties.
- is a formal framework for modeling ideal and rational changes of KBs under the Principle of minimal change
- provides coherent retraction and incorporation of information
- for our Tweety example it admits all three choices for retraction

Two approaches to belief revision

The approaches to belief revision we consider:

1. Axiomatic
   Describe change functions axiomatically by use of postulates.
   (Similarly to our way of showing properties of inference relations)
2. Constructive
   By use of the information encoded in a preference relation, called epistemic entrenchment relation, one can construct the actual change functions.

We discuss the classical change operations for KBs:

- expansion,
- contraction,
- withdrawal,
- revision.

Subsection 7.2

Axiomatizing change functions

Change function: expansion

Expansion[1] of a knowledge base:

- accepting new information
- not removing any of the previously accepted information
- can lead to inconsistencies in the KB

Definition 7.2 (Expansion)

Let T be a theory and $\varphi \in \mathcal{L}$. The expansion of a theory T w.r.t. a sentence φ is the logical closure of $T \cup \varphi$.

The set of all theories of a language $\mathcal{L}$ is $\mathcal{K}_{\mathcal{L}}$.

An expansion function $^+$ is a function $^+ : \mathcal{K}_{\mathcal{L}} \times \mathcal{L} \mapsto \mathcal{K}_{\mathcal{L}}$ that maps (T, φ) to T^+_φ, where $T^+_\varphi = Th(T \cup \varphi)$.

[1] Not to be confused with the expansions of autoepistemic theories.

Expansion operation

Note, that the expansion operation . . .

- is a monotonic operation,
  i.e. $T \subseteq T^{+}_{\varphi}$ and if $\neg\varphi \in T$, then T^{+}_{φ} is inconsistent.
- of T by φ, if $\neg\varphi \notin T$, has T^{+}_{φ} as the smallest change to incorporate φ in T—realizing the Principle of minimal change.
- is a unique operation! (No choice here.)

Change function: contraction

A contraction of a KB w.r.t. a formula φ causes the removal of a set of formulae from the KB, so that φ is no longer implied (unless φ is a tautology).

The challenge is to determine which sentences should be given up.

Definition 7.3 (Contraction)

Let T be a theory and $\varphi \in \mathcal{L}$.
A contraction function $^-$ is a function $^- : \mathcal{K}_\mathcal{L} \times \mathcal{L} \mapsto \mathcal{K}_\mathcal{L}$ mapping (T, φ) to T^-_φ, where T^-_φ satisfies the following postulates (for any $\varphi, \psi \in \mathcal{L}$ and any $T \in \mathcal{K}_\mathcal{L}$):

($^-$1) $T^-_\varphi \in \mathcal{K}_\mathcal{L}$

($^-$2) $T^-_\varphi \subseteq T$

($^-$3) If $\varphi \notin T$ then $T \subseteq T^-_\varphi$

($^-$4) If $\not\vdash \varphi$ then $\varphi \notin T^-_\varphi$

($^-$5) $T \subseteq (T^-_\varphi)^+_\varphi$ (recovery)

($^-$6) If $\vdash (\varphi \leftrightarrow \psi)$ then $T^-_\varphi = T^-_\psi$

($^-$7) $T^-_\varphi \cap T^-_\psi \subseteq T^-_{(\varphi \wedge \psi)}$

($^-$8) If $\varphi \notin T^-_{(\varphi \wedge \psi)}$ then $T^-_{(\varphi \wedge \psi)} \subseteq T^-_\varphi$

On the contraction postulates

The AGM postulates for contraction

- incorporate the "Principle of minimal change"
- act as integrity constraints for change functions
- do not determine a unique change function, but rather identify the set of possible KBs resulting from retracting information
- identify a class of functions for a KB
- are motivated by the criterion of informational economy

Intuition for the individual contraction postulates I

First postulate (¯1):
The result of contraction is a theory, new KB is logically closed.

Second postulate (¯2)
Contracting a theory involves only removal of old information, never the addition of new information. No spurious information is added.

Third postulate (¯3)
When information φ is is not accepted, then taking it away has no effect.

(¯2) and (¯3) say that "if φ is not in T, then taking it away has no effect".

Forth postulate (¯4)
If φ is not a tautology, then it must be removed in the contraction T^-_φ.

Intuition for the individual contraction postulates II

Fifth postulate (¯5), a.k.a. recovery
If φ is retracted and then replace it by using expansion, then T is restored.

Together with the last 4 postulates says that if $\varphi \in T$, then $T = (T)^+_\varphi$ i.e., no more information is lost than can be reintroduced by an expansion w.r.t. the information retracted.

Sixth postulate (¯6)
The same result is obtained if contraction is performed w.r.t. equivalent sentences; contraction is oblivious to syntax.

Seventh postulate (¯7)
The theory obtained from retracting w.r.t. $\varphi \wedge \psi$ should never be smaller than taking the intersection of T^-_φ and T^-_ψ.

Eighth postulate (¯8)
If φ is not contained in the contraction w.r.t. $(\varphi \wedge \psi)$ then the contraction w.r.t. this conjunction cannot be larger than the theory obtained from contraction of φ alone.

The recovery postulate is controversial!

Two strange effects of recovery:

1. A consequence of theory closure and recovery is that for all $\psi \in \mathcal{L}$ holds: if a sentence φ is contained in theory T, then $(\psi \longrightarrow \varphi) \in T^{-}_{\psi\vee\varphi}$ and the consequences $\varphi \in (T^{-}_{\psi\vee\varphi})^{+}_{\psi}$ and $\neg\psi \in (T^{-}_{\psi\vee\varphi})^{+}_{\neg\psi}$ are not always desirable.
2. Assume φ and ψ are in T, then so is $(\varphi \vee \psi)$. If $(\varphi \vee \psi)$ is learned to be wrong and retracted by the agent, but later on re-learned, then it should be possible not to add φ and ψ again.

Change function: withdrawal

The withdrawal operation is similar to contraction, but need not satisfy recovery. A withdrawal function satisfies the postulates $(^{-}1)$ to $(^{-}4)$ and $(^{-}6)$ to $(^{-}8)$, but not necessarily $(^{-}5)$.

Change function: revision

Revision attempts to change a KB as little as possible to accommodate new information. If the KB turns inconsistent, then information has to be retracted to regain consistency. Thus revision functions are nonmonotonic.

Definition 7.4 (Revision)

Let T be a theory and $\varphi \in \mathcal{L}$.
A revision function $*$ is a function $* : \mathcal{K}_{\mathcal{L}} \times \mathcal{L} \mapsto \mathcal{K}_{\mathcal{L}}$ mapping (T, φ) to T^*_φ, where T^*_φ satisfies the following postulates (for any $\varphi, \psi \in \mathcal{L}$ and any $T \in \mathcal{K}_{\mathcal{L}}$):

(*1) $T^*_\varphi \in \mathcal{K}_{\mathcal{L}}$

(*2) $\varphi \in T^*_\varphi$

(*3) $T^*_\varphi \subseteq T^+_\varphi$

(*4) If $\neg\varphi \notin T$ then $T^+_\varphi \subseteq T^*_\varphi$

(*5) If $T^*_\varphi = \bot$ then $\vdash \neg\varphi$

(*6) If $\vdash (\varphi \leftrightarrow \psi)$ then $T^*_\varphi = T^*_\psi$

(*7) $T^*_{(\varphi\wedge\psi)} \subseteq (T^*_\varphi)^+_\psi$

(*8) If $\neg\psi \notin T^*_\varphi$ then $(T^*_\varphi)^+_\psi \subseteq T^*_{\varphi\wedge\psi}$

Intuition for the revision postulates I

First postulate (*1):
Revising a theory results in a theory, new KB is logically closed.

Second postulate (*2)
Information to be added by a revision is always successfully added to the KB.

Third postulate (*3)
Revising a theory can never incorporate more information than an expansion operation.

Forth postulate (*4)
If $\neg\varphi$ is not contained in T then the expansion of T w.r.t. φ is contained in the revision w.r.t. φ.

Intuition for the revision postulates II

Fifth postulate (*5)
The only way to obtain an inconsistent theory is to revise w.r.t. an inconsistent sentence

Sixth postulate (*6)
Revision functions are syntax independent.

Seventh postulate (*7)
The theory that results from revising w.r.t. $\varphi \wedge \psi$ should never contain more than the revision w.r.t. φ followed by the expansion w.r.t. φ.

Eighth postulate (*8)
If $\neg\varphi$ is contained in the revision w.r.t. φ then the revision w.r.t. φ followed by the expansion w.r.t ψ should not contain more information than the theory that results from revising w.r.t. $\varphi \wedge \psi$

If the first 6 postulates hold. Then (*7) and (*8) imply that
$T^*_{\varphi\wedge\psi}$ is equivalent to T^*_φ, T^*_ψ or $T^*_\varphi \cap T^*_\psi$