Lecture 10 – YSP and Chronological Minimization

1. Yale Shooting Problem
2. Chronological Minimization
   (a) Lifschitz's Pointwise Circumscription Solution
   (b) Kautz's Chronological Minimization Solution
   (c) Shoham's Maximal Ignorance Solution
3. Conclusions

Yale Shooting Problem

– In 1986, McDermott and Hank's presented a paper questioning the feasibility of logic in general and nonmonotonic logic in particular as KR language for reasoning about action and change.

– As a vehicle for discussion they presented the following problem of temporal projection, where facts about the initial situation are given along with descriptions of causal rules. The goal is to predict the facts that hold after performing a sequence of actions.

The Original Axiomatization:

$$Holds(alive, S0) \tag{1}$$

$$\forall s.Holds(loaded, result(load, s)) \tag{2}$$

$$\forall s.Holds(loaded, s) \supset Ab(alive, shoot, s) \wedge Holds(dead, result(shoot, s)) \tag{3}$$

$$\forall f, e, s.Holds(f, s) \wedge \neg Ab(f, e, s) \supset Holds(f, result(e, s)) \tag{4}$$

An Equivalent Abnormality Theory

Initial Situation

$$Holds(alive, S0) \tag{5}$$

Laws of Motion

$$\forall s.Holds(loaded, result(load, s)) \tag{6}$$

$$\forall s.Holds(loaded, s) \supset Holds(dead, result(shoot, s)) \tag{7}$$

Frame Axioms

$$\forall f, e, s.Holds(f, s) \wedge \neg Ab(f, e, s) \supset Holds(f, result(e, s)) \tag{8}$$

$$\forall s.Holds(loaded, s) \supset Ab(alive, shoot, s) \tag{9}$$

State Constraints

$$S1 = result(load, S0) \tag{10}$$
$$S2 = result(wait, S1) \tag{11}$$
$$S3 = result(shoot, S2) \tag{12}$$

where

$$S3 = result(shoot, result(wait, result(load, S0))) \tag{13}$$

Ppoint Diagram 1:(YSP(I))

Circumscription Strategy

– Given the fact that the turkey is $alive$ in $S0$ and the actions $load$, $wait$, and $fire$ are performed in sequence, the temporal projection problem is to determine if the turkey is $alive$ or $dead$ in the resulting situation $S3$.

– The following circumscription strategy appears reasonable:

- **Minimize:** Ab.
- **Vary:** $Holds$
- **Fix:** Everything else.

– Unfortunately, this policy suffers from the "weak disjunction" problem and the preferred model set contains an unintended model. (See diagram).

– At best, we can derive:

$$Holds(alive, S3) \vee Holds(dead, S3).$$

– **Clue to a solution:** Direction in which Ab is minimized as preferred models are constructed.

Ppoint Diagram 2: (YSP(II))

Chronological Minimization

– Recall the manner in which the two preferred model classes were constructed.

1. Reasoning forward in time: Preference for $\neg Ab(loaded, wait, S1)$ and $Ab(alive, shoot, S2)$.

2. Reasoning backward in time: Preference for $\neg Ab(alive, shoot, S2)$ and $Ab(loaded, wait, S1)$.

IDEA: Prefer those models where abnormal change occurs as late as possible in the temporal ordering.

– **Problem with initial solution:** General nonmonotonic reasoning disregards temporal nature of the domain. Preference is set-theoretical.

–**Chronological Minimization:** Minimize abnormal change chronologically, relative to the particular temporal structure.

Lifschitz's Pointwise Minimization

– Recall the technique of pointwise minimization, where given a predicate P, one has the capability of minimizing at points (tuples) in the extension of P, while allowing other points in P or other predicates to vary.

– Recall again, the manner in which the two preferred model classes for the YSP were constructed.

1. Reasoning forward in time: Preference for $\neg Ab(loaded, wait, S1)$ and $Ab(alive, shoot, S2)$.

2. Reasoning backward in time: Preference for $\neg Ab(alive, shoot, S2)$ and $Ab(loaded, wait, S1)$.

IDEA: Place on ordering on situations (and indirectly tuples $\langle f, a, s\rangle$) and minimize at points in Ab chronologically.

CIRC POLICY:

– Given $S < S' < S''$, when minimizing Ab at a point $\langle --, --, S'\rangle$:

- **Minimize:** $\langle --, --, S'\rangle$ in Ab.
- **Vary:** All points $\langle --, --, S''\rangle$ in Ab greater than S'. Also $Holds$ in the standard manner.
- **Fix:** All points $\langle --, --, S\rangle$ in Ab less than S'.

Two Additions to the Axiomatization

– To make this work, we must make explicit what we mean by "temporally earlier". We do this by adding the following axiom to our premise set,

$$\begin{aligned}\forall s, s'.s < s' \equiv \quad & (s = S0 \wedge s' = S1) \vee \\ & (s = S0 \wedge s' = S2) \vee \\ & (s = S0 \wedge s' = S3) \vee \\ & (s = S1 \wedge s' = S2) \vee \\ & (s = S1 \wedge s' = S3) \vee \\ & (s = S2 \wedge s' = S3) \end{aligned} \tag{14}$$

– In addition, we add an axiom defining an ordering on the fluent, action, situation tuples,

$$\forall f, e, s, f', e', s'.\langle f, e, s\rangle <_p \langle f', e', s'\rangle \equiv s < s'. \tag{15}$$

The Circumscription Axiom

Lifschitz then proposes the following circumscription axiom

$$CIRC_{PW}(\Gamma; Ab/<_P, Holds) \equiv$$

$$\forall f, e, s, Ab', Holds'.\neg[Ab(f,e,s) \tag{16}$$

$$\wedge \forall f', e', s'.\langle f', e', s'\rangle <_p \langle f, e, s\rangle$$

$$\supset (Ab'(f', e', s') \equiv Ab(f', e', s')) \tag{17}$$

$$\wedge \Gamma(\lambda f', e', s'.(Ab'(f', e', s') \wedge \langle f, e, s\rangle \neq \langle f', e', s'\rangle), Holds')] \tag{18}$$

The circumscription axiom states that for any substitutions Ab' and $Holds'$ of Ab and $Holds$ and point $\langle f, e, s\rangle$, it can not be the case that

16. Ab is true at $\langle f, e, s\rangle$

    and that

17. Ab and Ab' agree on all points temporally earlier than s

    and that

18. Ab' satisfies the theory Γ at all points other than $\langle f, e, s\rangle$.

Applying the Circumscription Axiom

Step1: Choose substitutions for Ab' and $Holds'$.

We use some wishful thinking and choose substitutions that correspond to the preferred model.

$$\begin{aligned} Ab_c \equiv \lambda f, e, s.(\langle f, e, s\rangle \in \{&\langle alive, shoot, result(shoot, S1)\rangle \\ &\langle alive, shoot, S2\rangle \\ &\langle alive, shoot, S3\rangle \\ &\langle alive, shoot, result(shoot, S3)\rangle\} \end{aligned} \tag{19}$$

$$\begin{aligned} Holds_c \equiv \lambda f, s.(\langle f, s\rangle \in \{&\langle alive, S0\rangle \\ &\langle alive, S1\rangle, \langle loaded, S1\rangle \\ &\langle alive, S2\rangle, \langle loaded, S2\rangle \\ &\langle dead, S3\rangle, \langle loaded, S3\rangle \\ &\langle dead, result(shoot, S1)\rangle \\ &\langle dead, result(shoot, S3)\rangle\} \end{aligned} \tag{20}$$

Given the substitutions, we would like to be able to conclude that

$$Ab(loaded, wait, S1) \tag{21}$$

is not true.

– Let's instantiate the circumscription axiom with this tuple and our substitutions,

$$\neg[Ab(loaded, wait, S1) \tag{22}$$
$$\wedge \forall f', e', s'.\langle f', e', s'\rangle <_p \langle loaded, wait, S1\rangle$$
$$\supset (Ab_c(f', e', s') \equiv Ab(f', e', s')) \tag{23}$$
$$\wedge \Gamma(\lambda f', e', s'.(Ab_c(f', e', s') \wedge \langle loaded, wait, S1\rangle \neq \langle f', e', s'\rangle),$$
$$Holds')] \tag{24}$$

Step2: Proof conjuncts are true.

–Let's call the three conjuncts C1', C2' and C3'. If we can show that C2' and C3' are true with the help of Γ then by the circumscription axiom it must follow that $\neg Ab(loaded, wait, S1)$.

– In fact, this can be done. We save the details for the lecture with practical exercises.

Counterexamples

– Postdiction

– Stanford Murder Mystery

– Parked Car Example

Ppoint Diagram:(SMM1(1-3))

Comments

– The circumscription axiom produces exactly those theorems true in the chronologically minimal models.

– Is the criteria general enough to deal with all projection problems?

– **Major problem:** using the circumscription axiom to generate useful inferences. Our example shows the strategy and difficulties involved.

1. We know what intended conclusions we want from our theory.

2. We use this knowledge to construct an instance of the circumscription axiom.

3. We use this instance together with the other axioms in the theory to see if we can derive the intended conclusions we have in mind.

In effect there is no inferential mechanism for deriving intended conclusions that we do not know about, which is really the point of using logic as an inference mechanism in the first place.

Kautz's Logic of Persistence

– Uses a reified temporal logic.

– Emphasizes model theory rather than circumscription axiom.

Basic Idea:

– Introduce a predicate constant $Clip$. Asserting $Clip(t, f)$ represents a termination or change in value from true to false of a fact f at t.

– At each time step, the value of a fact either persists or is clipped. Represented in the form of a persistence axiom:

$$\forall t, f. Holds(t, f) \supset Holds(t+1, f) \oplus Clip(t+1, f) \quad (25)$$

which states that if a fact f holds at time t, $Holds(t, f)$, then either it persists, $Holds(t+1, f)$, or its persistence is clipped, $Clip(t+1, f)$. The $\oplus$ operator is defined as an exclusive-or operator.

– Simply circumscribing $Clip$ while varying $Holds$ would still result in two classes of preferred models analogous to those discussed above.

– In order to avoid the unintended class, the criteria for preferring one model over the other would be to minimize $Clip$ chronologically, preferring models where the $Clip$ instances are as late in time as possible.

Original Axiomatization

Initial Situation

$$Holds(1, loaded) \tag{26}$$

$$Holds(1, alive) \tag{27}$$

Laws of Motion

$$\begin{aligned} &\forall t.Holds(t, fire) \wedge Holds(t, loaded) \supset \\ &\quad \neg Holds(t+1, loaded) \wedge \neg(t+1, alive) \end{aligned} \tag{28}$$

$$Holds(3, fire) \tag{29}$$

Frame Axioms

$$\forall t, f.Holds(t, f) \supset Holds(t+1, f) \oplus Clip(t+1, f) \tag{30}$$

State Constraints

$$\alpha \oplus \beta \equiv_{def} (\alpha \wedge \neg\beta) \vee (\neg\alpha \wedge \beta) \tag{31}$$

$$loaded \neq alive \wedge loaded \neq fire \wedge alive \neq fire \tag{32}$$

Semantic Characterization

Preliminaries:

– Let $M1$ and $M2$ be models and P a predicate of arity n.

– We write $M1[P]$ to denote the extension of P in $M1$ and $\langle x_1, \ldots, x_n \rangle$ to denote a tuple in $M1[P]$.

– Besides $Clip$ and $Holds$, predicate constant $<$. $t' < t$ denotes that t' is temporally before t.

– **Minimization policy:** Minimize $Clip$ chronologically, vary $Holds$, and fix everything else.

Definition 1 (Preferred Model) *$M1$ is preferred to $M2$, written $M1 \sqsubseteq M2$ iff*

1. *$M1$ and $M2$ have the same domain.*

2. *Every constant, function and predicate symbol other than $Clip$ and $Holds$ receives the same interpretation in $M1$ and $M2$.*

3. *The following (meta-theoretic) statement is true.*

$$\begin{aligned} \langle f, t \rangle \in M1[Clip] \supset (\langle f, t \rangle \in M2[Clip] \vee \\ (\exists t', f'. \langle f, t \rangle \in M2[Clip] \wedge \langle f', t' \rangle \notin M1[Clip] \wedge \\ \langle t', t \rangle \in M1[<])) \end{aligned} \quad (33)$$

A model $M1$ is said to be strictly preferred *to $M2$, written $M1 \sqsubset M2$, if $M1 \sqsubseteq M2$ and not $M2 \sqsubseteq M1$.*

Syntactic Characterization

– Like Lifschitz, Kautz also provides a syntactic characterization of his definition of preferential entailment by using a circumscription axiom.

$$\begin{aligned} &\forall Clip', Holds'.[\Gamma(Clip', Holds') \wedge \\ &\forall t, f.Clip'(t, f) \supset [Clip(t, f) \vee \\ &\exists t', f'.t' < t \wedge Clip(t', f') \wedge \neg Clip'(t', f')]] \\ &\supset \forall t, f.Clip(t, f) \equiv Clip'(t, f) \end{aligned} \tag{34}$$

– It is useful to compare Kautz's circumscription axiom with Lifschitz's.

– Example considered in Kautz's paper and might be done during the practical lecture.

Problems

- Postdiction
- Parked Car Example.
- Stanford Murder Mystery.

Shoham's Logic of Chronological Ignorance

– Shoham constructs a nonmonotonic logic called *chronological ignorance*(CI) which combines elements of

- Temporal logic
- Modal logic
- Nonmonmotonic logic

– Uses the strategy of taking an arbitrary logic and defining a preference ordering on its models. He then supplies an algorithm for computing preferred entailments directly.

– We will introduce Shoham's logic in two steps:

1. Describe a monotonic modal logic of temporal knowledge.
2. Describing a partial order on its models.

Preliminaries

– Recall Shoham's previous temporal logic STL. Temporal information is described in terms of logical combinations of atomic formulas of the following type

$$TRUE(t_1, t_2, r(\mu_1, \ldots, \mu_n),$$

where r is a relation symbol and $\mu_1, \ldots, \mu_n$ are atemporal terms.

– In CI, Shoham basically uses the same representation of temporal information but the propositional type is simplified to be a propositional constant rather than allowing arguments

$$TRUE(t_1, t_2, p).$$

– In addition to the temporal information, Shoham introduces a modal operator, $\Box$ which may loosely be understood as a *knowledge* operator.

$$\Box TRUE(t_1, t_2, p)$$

asserts that p is known to be true during the interval from t_1 to t_2.

– Shoham also assumes an integer time structure $\langle \mathcal{N}, \leq \rangle$.

Modal Logic and Kripke Frames

– Models or interpretations would then be similar to *Kripke interpretations* which are pairs $\langle W, M \rangle$ where W is a non-empty universe of (possible) worlds, and M is a meaning function.

– Possible worlds can be viewed as classical models with a integer time structure which is the same for all worlds. In other words, a set of infinite parallel time lines all sharing the same temporal interpretation but with different facts true at different points in time.

– Each possible world then represents a possible evolution in time or history of the universe.

TK: A Monotonic Logic of Temporal Knowledge

Syntax:

1. **P**: A set of primitive propositions.

2. **TV**: A set of temporal variables.

3. **TC**: The set $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$.

4. **U= TC**$\cup$ **TV**: The set of temporal terms.

The set of well-formed formulas is defined inductively as follows:

1. If $u_1 \in \mathbf{U}$ and $u_2 \in \mathbf{U}$ then $u_1 = u_2$ and $u_1 \preceq u_2$ are wffs.

2. If $u_1 \in \mathbf{U}$, $u_2 \in \mathbf{U}$, and $p \in \mathbf{P}$ then $TRUE(u_1, u_2, p)$ is a wff.

3. If α and β are wffs then so are $\alpha \wedge \beta$, $\neg\alpha$, and $\Box\alpha$.

4. If α is a wff and $v \in \mathbf{TV}$, then $\forall v.\alpha$ is also a wff.

In addition there are a number of abbreviations that may be found in the article.

Semantics:

Shoham assumes an S5 modal structure with a fixed interpretation of time across worlds. That fixed interpretation is the integer structure $\langle \mathcal{N}, \leq \rangle$.

Definition 2 (Kripke Interpretation) *A* Kripke Interpretation *is a pair* $\langle W, M \rangle$ *where* W *is a non-empty universe of possible worlds, and* M *is a meaning function* $M : \mathbf{P} \rightarrow 2^{W \times \mathcal{N} \times \mathcal{N}}$. *In addition,* $\langle w, t, t' \rangle \in M(p)$ *iff* $\langle w, t', t \rangle \in M(p)$.

Definition 3 (Variable Assignment) *A* variable assignment *is a function* $VA : \mathbf{TV} \rightarrow \mathcal{N}$.

If $u \in \mathbf{U}$ *then* $Val(u) = VA(u)$ *if* $u \in \mathbf{TV}$ *and the* standard interpretation *of* u *if* $u \in \mathbf{TC}$.

Definition 4 (Satisfaction Conditions) *A Kripke Interpretation* $KI = \langle W, M \rangle$ *and a world* $w \in W$ *satisfy a formula* α *under the variable assignment* VA*, written* $KI, w \models \alpha[VA]$*, under the following conditions:*

- $KI, w \models u_1 = u_2[VA]$ *iff* $VAL(u_1) = VAL(u_2)$.
- $KI, w \models u_1 \preceq u_2[VA]$ *iff* $VAL(u_1) \leq VAL(u_2)$.
- $KI, w \models TRUE(u_1, u_2, p)[VA]$ *iff* $\langle w, VAL(u_1), VAL(u_2) \rangle \in M(p)$.
- $KI, w \models \alpha \wedge \beta[VA]$ *iff* $KI, w \models \alpha[VA]$ *and* $KI, w \models \beta[VA]$
- $KI, w \models \neg\alpha[VA]$ *iff* $KI, w \not\models \alpha[VA]$.
- $KI, w \models (\forall v \alpha)[VA]$ *iff* $KI, w \models \alpha[VA']$ *for all* VA' *that agree with* VA *everywhere except possible on* v.
- $KI, w \models \Box\alpha[VA]$ *iff* $KI, w' \models \alpha[VA]$ *for all* $w \in W$.

Definition 5 (Semantic Entailment) *A Kripke interpretation* $KI = \langle W, M \rangle$ *and a world* $w \in W$ *are a* model *for a formula* α*, written* $KI, w \models \alpha$*, if* $KI, w \models \alpha[VA]$ *for any variable assignment* $[VA]$.

A wff is satisfiable *if it has a model, and* valid *if its negation has no model.*

$\alpha_1$entailsα_2*, written* $\alpha_1 \models \alpha_2$ *iff* α_2 *is satisfied by all models of* α_1.

A Scenario

– The syntax and model theory of TK will be used for representing and interpreting scenario descriptions.

– Scenario used by Shoham as a vehicle for discussing the frame and qualification problems.

Boundary Conditions

$$\Box(1, loaded) \tag{35}$$

$$\Box(5, fire) \tag{36}$$

Causal Rules

$$\begin{aligned} &\Box(t, loaded) \wedge \Box(t, fire) \wedge \\ &\Diamond(t, air) \wedge \Diamond(t, firingpin) \\ \wedge \ldots \Diamond \ldots &\text{ other mundane conditions} \\ &\supset \Box(t+1, noise), \text{ for all } t \in TC. \end{aligned} \tag{37}$$

Frame Axioms

$$\begin{aligned} &\Box(t, loaded) \wedge \Box(t, fire) \wedge \\ &\Diamond(t, \neg fire) \wedge \Diamond(t, \neg emptied - manually) \\ &\supset \Box(t+1, loaded), \text{ for all } t \in TC. \end{aligned} \tag{38}$$

An alternate axiomatization, based on abnormality theories:

Modified Causal Rules

$$\Box(t, loaded) \wedge \Box(t, fire) \wedge \\ \wedge \Diamond(t, sit - normal) \\ \supset \Box(t+1, noise), \text{ for all } t \in TC. \quad (39)$$

$$\Box(t, \neg air) \supset \Box(t, \neg sit - normal), \text{ for all } t \in TC. \quad (40)$$

$$\Box(t, \neg firingpin) \supset \Box(t, \neg sit - normal), \text{ for all } t \in TC \quad (41)$$

$$\ldots \Box \ldots \text{ other mundane conds.} \ldots \supset \Box(t, \neg sit - normal), \\ \text{for all } t \in TC \quad (42)$$

Modified Frame Axioms

$$\Box(t, loaded) \wedge \Box(t, fire) \wedge \\ \wedge \Diamond(t, sit - normal_1) \\ \supset \Box(t+1, loaded), \text{ for all } t \in TC. \quad (43)$$

$$\Box(t, \neg fire) \supset \Box(t, \neg sit - normal_1), \text{ for all } t \in TC. \quad (44)$$

$$\Box(t, \neg emptied - manually) \supset \Box(t, \neg sit - normal_1), \\ \text{for all } t \in TC. \quad (45)$$

– Note the explicit use of a frame axiom rather than a general frame rule.

– Note the use of schemas (to insure that we reason propositionally).

CI: A Nonmonotonic Logic of Temporal Knowledge

– CI, the logic of chronological ignorance, is a nonmonotonic version of TK with the same syntax, but also augmented with an ordering relation on Kripke structures.

Definition 6 (Base Wff) *A* base wff *is a wff containing no modal operator.*

Let $maxt(t_1, t_2)$ be the max function for time points relative to the ordering $\leq$.

Definition 7 (Latest Time Point) *The* latest time point (l.t.p) *of a base formula is the latest time point mentioned in it defined as follows.*

1. *The l.t.p of* $TRUE(t_1, t_2, p)$ *is* $max(t_1, t_2)$.

2. *The l.t.p of* $\alpha \wedge \beta$ *is the maxt of the l.t.p's of* α *and* β.

3. *The l.t.p of* $\neg\alpha$ *is the l.t.p of* α.

4. *The l.t.p of* $\forall v.\alpha$ *is the earliest among the l.t.p's of all* α' *which result from substituting in* α *a time point symbol for all free occurrences of* v, *or* $-\infty$ *if there is no such earliest point.*

Preference Ordering on Models

Definition 8 (Chronologically more ignorant) *A Kripke interpretation M_2 is* chronologically more ignorant *than a Kripke interpretation M_1, written $M_1 \sqsubseteq M_2$, if there exists a time t_0 such that*

1. *for any base sentence α whose l.t.p. $\leq t_0$, if $M_2 \models \Box\alpha$ then also $M_1 \models \Box\alpha$, and*

2. *there exists some base sentence α whose l.t.p. is t_0 such that $M_1 \models \Box\alpha$ but $M_2 \not\models \Box\alpha$.*

Definition 9 (Chronologically Maximally Ignorant) *M is said to be a* chronologically maximally ignorant*(c.m.i.) model of α if $M \models_{\sqsubset_{ci}} \alpha$, that is, if $M \models \alpha$ and there is no other M' such that $M' \models \alpha$ and $M \sqsubset_{ci} M'$.*

Definition 10 (CI) *The logic of chronological ignorance, CI, is the nonmonotonic logic $TK_{\sqsubset_{ci}}$.*

Format of Axioms

Axioms in scenario descriptions like the YSP are generally restricted to two forms:

1. Boundary conditions represented as knowledge about atomic sentences.

$$\Box(1, loaded)$$

2. Causal rules and Frame axioms represented as implications from both knowledge $\Box$ and ignorance ($\Diamond \equiv \neg\Box\neg$) of past atomic sentences to knowledge of future atomic sentences.

$$\begin{aligned} &\Box(t, loaded) \wedge \Box(t, fire) \wedge \\ &\Diamond(t, air) \wedge \Diamond(t, firingpin) \\ \wedge \ldots \Diamond \ldots &\text{ other mundane conditions} \\ \supset &\Box(t+1, noise), \text{ for all } t \in TC. \end{aligned} \tag{46}$$

– The advantage of limiting axioms to these formats is that for any theory, the base sentences will be the same in any c.m.i. model for the theory.

Causal Theories

– Lets provide a formal definition for theories constructed from the restricted language. Shoham calls these *causal theories*.

Definition 11 (Causal Theory) *A* causal theory Ψ *is a theory in CI, in which all sentences have the form*

$$\Phi \wedge \Theta \supset \Box\alpha$$

where

1. α *is a (positive or negative) atomic base sentence.*

2. Φ *is a conjunction of sentences* $\Box\alpha_i$*, where* α_i *is a (positive or negative) atomic base sentence whose l.t.p.* $<$ *the l.t.p. of* $alpha$*.*

3. Θ *is a conjunction of sentences* $\Diamond\alpha_i$*, where* α_i *is a (positive or negative) atomic base sentence whose l.t.p.* $<$ *the l.t.p. of* $alpha$*.*

4. *Either or both* Φ *and* Θ *may be empty (true). A sentence in which* Φ *is empty is called a* boundary condition*. Other sentences are called* causal rules*.*

5. *There is a time point* $t_0 \neq -\infty$ *preceding the l.t.p. of the r.h.s. in all boundary conditions (that is, if* $\Theta \supset \Box\alpha$ *is a boundary condition, then* $t_0 <$ *the l.t.p. of* α*.*

6. For any p, t_1, t_2, there do not exist two sentences in Ψ such that one contains $\Diamond(t_1, t_2, p)$ on its l.h.s. and the other contains $\Diamond(t_1, t_2, \neg p)$ on its l.h.s.

7. If $\Phi_1 \wedge \Theta_1 \supset \Box(t_1, t_2, p)$ and $\Phi_2 \wedge \Theta_2 \supset \Box(t_1, t_2, \neg p)$ are two sentences in Ψ, then $\Phi_1 \wedge \Theta_1 \wedge \Phi_2 \wedge \Theta_2$ is inconsistent.

– Why are causal theories so nice?

– They have a number of properties that make it easy to construct c.m.i. models.

Theorem 1 *If Ψ is any causal theory, then*

1. *Ψ has a c.m.i. model,*

2. *If M_1 and M_2 are both c.m.i. models of Ψ, and α any base sentence, then $M_1 \models \Box\alpha$ iff $M_2 \models \Box\alpha$, and*

3. *If M is a c.m.i. of Ψ, then the set of known base sentences consists of the tautological consequences of the (positive and negative) known atomic base sentences.*

– What this theorem states is that for any causal theory Ψ, its preferred models will have the same set of base sentences, the theory will be consistent, and the models may be built by moving forward in time, adding only those sentences that we know about and those implied by what is already known. We can also do this efficiently for each causal theory.

The YSP Example

Boundary Conditions

$$\Box(0, loaded) \tag{47}$$

$$\Box(0, alive) \tag{48}$$

$$\Box(2, fire) \tag{49}$$

Causal Rules

$$\Box(t, loaded) \wedge \Box(t, fire) \wedge \supset \Box(t+1, \neg alive), \text{ for all } t \in TC. \tag{50}$$

Frame Axioms

$$\Box(t, loaded) \wedge \Diamond(t, \neg fire) \wedge \Diamond(t, \neg emptied - manually) \supset \Box(t+1, loaded), \text{ for all } t \in TC. \tag{51}$$

It is easy to see that, if starting at time point 0 and reasoning forwards, all c.m.i models of the theory will satisfy $\Box(3, \neg alive)$. We can show this by constructing the sentences known at each time point.

- **Tp 0:** Boundary conditions: $\Box(0, loaded)$, $\Box(0, alive)$.

- **Tp 1:** Instantiate axiom (**??**)

  $$\Box(0, loaded) \wedge \Diamond(0, \neg fire) \wedge \Diamond(0, \neg emptied - manually) \supset \Box(1, loaded) \quad (52)$$

  or equivalently,

  $$\Box(0, loaded) \wedge \neg\Box(0, fire) \wedge \neg\Box(0, emptied - manually) \supset \Box(1, loaded) \quad (53)$$

  By minimizing $\Box$ at Tp 1, we can derive $\neg\Box(0, fire)$, $\neg\Box(0, emptied - manually)$. Consequently, we can derive $\Box(1, loaded)$ using modus ponens.

  Proof: Suppose $M \models \Box(1, loaded)$ and $M' \not\models \Box(1, loaded)$. By axiom (**??**), either $M' \models \Box(0, fire)$ or $M' \models \Box(0, emptied - manually)$. Based on the c.m.i. definition, it must be the case that $M' \sqsubseteq M$ which implies that M is a c.m.i. model and $\Box(1, loaded)$ is true in it.

- **Tp 3:** A similar argument can be used to derive $\Box(2, loaded)$ (d1). (d1) together with axioms (**??**) and (**??**) permit us to derive $\Box(3, \neg alive)$ in all c.m.i. models as desired.

Intuitions behind CI

–As Shoham states, the intuition behind his logic and causal theories is the "what you do not know will not hurt you" principle.

– Note that $\Diamond\alpha \equiv \neg\Box\neg\alpha$, $\Diamond\alpha$ atoms are used to qualify causal rules, and we minimize $\Box$.

– The net result is that unless we are explicitly aware that certain conditions hold, we may assume we do not know about them and the causal rules will have there normal effects. Only include qualifications we explicitly know about.

– Shoham's logic exploits the partiality of knowledge contexts. Since the logic is S5-based, it is the case that

$$\models \Box(t,p) \vee \neg\Box(t,p)$$

but ,

$$\not\models \Box(t,p) \vee \Box(t,\neg p).$$

– Minimizing knowledge ($\Box$), does minimize all propositions simultaneously within the knowledge context. If one minimizes knowledge about p, then $\neg\Box(t,p)$ does not imply that $\Box(t,\neg p)$, it may also be the case that $\neg\Box(t,\neg p)$.

– Minimizing *knowledge* of the truth or falsity of a proposition does not affect knowledge about any other propositions, except indirectly, where knowledge about propositions later in time can be derived as an indirect result of applying the causal rules together with our conjectures of ignorance.

Conclusions

Chronological Minimization is not without problems:

1. Restricted at best to certain types of temporal projection.

2. Not good for postdiction, or explanation.

3. A number of counterexamples. Are there more?

4. Circumscriptive based approaches are computationally unfeasible?

A better Solution: For what classes of reasoning problems are each of the proposals suitable. What are the underlying assumptions upon which the proposals are based?