CSE 513T Theory of Artificial Intelligence & Machine Learning November 10, 2016

Lecture 21

Lecturer: Brendan Juba *Scribe: Philip Bonneville*

CSE 513T 11/10

---

1. Computing Well-Founded Semantics

2. (if time) Introduction to classical planning

Note: Tuesday–Dingwen's 2nd guest-lecture

Definitions from last lecture—

- A *partial model* is given by disjoint sets of literals $L_\top$ and $L_\bot$ ("provable" & "improvable"), s.t. for every rule in the KB $[\ell_1 \wedge \cdots \wedge \ell_r \wedge \neg\ell_{r+1} \wedge \cdots \wedge \neg\ell_s]$ if $\ell_1, \ldots, \ell_r, \neg\ell_{r+1}, \ldots, \neg\ell_s \in L_r$, then $\ell^* \in L_\top$ as well, and, if $\ell^* \in L_\bot$, at least one of $\ell_1, \ldots, \ell_r, \neg\ell_{r+1}, \ldots, \neg\ell_s \in L_\bot$. (All facts must be in $L_\top$.)

- A *least* partial model in a set M is $(L_T^*, L_\bot^*)$ s.t. for other $(L_\top, L_\bot) \in M$, $L_\bot \subseteq L_\bot^*$, and $L_\top^* \subseteq L_\top$.

- Given the pair of sets $(L_\top, L_\bot)$ and a KB, the *quotient of the KB modulo* $(L_\top, L_\bot)$, denoted $\frac{\text{KB}}{(L_\top, L_\bot)}$ is obtained by (1) substituting 1 for $\neg\ell$ when $\ell \in L_\bot$, (2) substituting 0 for $\neg\ell$ when $\ell \in L_\top$, and (3) deleting occurrences of 1 plus the rules with 0 in the body.

Main Theorem Consider the sequence $(L_\top^{(0)}, L_\bot^{(0)}) = (\emptyset, \emptyset)$, $(L_\top^{(i)}, L_\bot^{(i)})$ is the least partial model of our $\frac{\text{KB}}{(L_\top^{(i-1)}, L_\bot^{(i-1)})}$ (that does not place any ℓ in $L_\bot^{(1)}$). Then the sequence converges to $(L_\top^*, L_\bot^*)$ s.t. for each of the standard literals, (1) if there is a chaining proof of ℓ in $\frac{KB}{(L_\top^*, L_\bot^*)}$, then $\ell \in L_\top^*$, and (2) if $\ell \in L_\bot^*$, then there is no chaining proof of ℓ from $\frac{\text{KB}}{(L_\top^*, L_\bot^*)}$. Furthermore, there is an $O(n^2|\text{KB}|)$ algorithm for computing $(L_\top^*, L_\bot^*)$.

Main Lemma $L_\top^{(i)} \subseteq L_\top^{(i+1)}$ and $L_\bot^{(i)} \subseteq L_\bot^{(i+1)}$.

Proof: By induction on i. For the base case $i = 0$, $\emptyset \subseteq L_\top^{(i)}$ and $\emptyset \subseteq L_\bot^{(i)}$ regardless. We now proceed to the induction step ($i \geq 2$). (1) $L_\top^{(i-1)} \subseteq L_\top^{(i)}$: Consider any $\ell \in L_\top^{(-1)}$ by construction of $L_\top^{(i-1)}$ plus the LPM Theorem, there was a chaining proof of ℓ^* from $\frac{KB}{(L_\top^{(i-2)}, L_\bot^{(i-2)})}$. (As an interruption to our proof, we will state the following theorem.)

Theorem The least partial model among those for which no "not ℓ"$\in L_\bot$ is obtained as follows. Take $L_\top^*$ to be the set of all literals provable from KB when we treat every naf'd literal as a new literal; take $L_\bot^*$ to be the set of literals that remain unprovable when the naf'd literals are dropped from the bodies of rules in KB.

Since, by the induction hypothesis, $L_\top^{(i-2)} \subseteq L_\top^{(i-1)}$ and $L_\bot^{(i-2)} \subseteq L_\bot^{(i-1)}$, we find (1) any rule that has "not ℓ" in its body in $\frac{\text{KB}}{(L_\top^{(n-1)}, L_\bot^{(n-1)})}$ must also have had "not ℓ" appear in its body in $\frac{\text{KB}}{(L_\top^{(i-2)}, L_\bot^{(i-2)})}$. Hence, this rule can't have been used in the proof (since there was no way to derive any "not ℓ"), (2) any rule that appears in $\frac{\text{KB}}{(L_\top^{(i-2)}, L_\bot^{(i-2)})}$ but not in $\frac{\text{KB}}{(L_\top^{(i-2)}, L_\bot^{(i-1)})}$ must have had an occurrence of "not ℓ" in its body for the quotient to delete it, and hence couldn't have been used in the proof of ℓ^* from $\frac{KB}{(L_\top^{(i-2)}, L_\bot^{(i-2)})}$. Thus, we can invoke the same sequence of chaining steps to prove ℓ^* in $\frac{KB}{(L_\top^{(i-1)}, L_\bot^{(i-1)})}$. So, by construction, our LPM Theorem, $\ell^* \in L_\top^{(i)} \Diamond^{(1)}$.

(2) $L_\bot^{(i-1)} \subseteq L_\bot^{(i)}$. We again have by induction hypothesis that $L_\top^{(i-2)} \subseteq L_\top^{(i-1)}$ and $L_\bot^{(i-2)} \subseteq L_\bot^{(i-1)}$. consider any $\ell^* \notin L_\bot^{(i)}$. By the LPM Theorem, there was a chaining proof of ℓ^* in KB$'$ obtained by dropping the remaining occurrences of naf'd literals from $\frac{\text{KB}}{(L_\top^{(i-1)}, L_\bot^{(i-1)})}$. Consider KB$''$ obtained by dropping the naf'd literals from $\frac{\text{KB}}{(L_\top^{(i-2)}, L_\bot^{(i-2)})}$. Notice every rule of KB$'$ is also in KB$''$ since (1) $L_\top^{(i-2)} \subseteq L_\top^{(i-1)}$ and (2) we delete rules iff some $\ell \in L_\top$ for a not ℓ in its body, and also we eliminated all remaining occurrences of "not ℓ" from either. Therefore, the same proof that establishes ℓ^* in KB$'$ also would establish it from KB$''$, hence $\ell^* \notin L_\bot^{(i-1)} \Diamond^{(2)}$.

Claim: Lemma ⇒ convergence (why?)

Dynamic Programming Algorithm: The idea is to iteratively compute $(L_\top^{(i)}, L_\bot^{(i)})$ from $(L_\top^{(i-1)}, L_\bot^{(i-1)})$, check for convergence.

On input KB, initialize two tables of size $2n$, $(L_\top^{(0)}, L_\bot^{(0)})$ containing all 0.
For $i = 1, 2, \ldots$
— Iterate over rules of KB
— — Delete occurrences of not ℓ for $\ell \in L_\bot^{(i-1)}$
— — Delete rule's that contain not ℓ $\ell \in L_\top^{(i-1)}$
— Initialize $L_\top^{(i)}$ to be all 0's, and $L_\bot^{(i)}$ to be all 1's
— Repeat
— — Interate over KB
— — — If the head is not in $L_\top^{(i)}$, but the body literals are
— — — — Add the head to $L_\top^{(i)}$
— — — If the head is in $L_\bot^{(i)}$ but none of the body literals are
— — — — Remove the head from $L_\bot^{(i)}$
— While $L_\top^{(i)}$ or $L_\bot^{(i)}$ was modified on the previous iteration
— If $L_\top^{(i-1)} = L_\top^{(i)}$ and $L_\bot^{(i-1)} = L_\bot^{(i)}$
— — Return $(L_\top^{(i)}, L_\top^{(i)})$

Claim (1): In the ith iteration, the algorithm uses KB = $\frac{\text{KB}}{(L_\top^{(i+1)}, L_\bot^{(i-1)})}$ and $(L_\top^{(i)}, L_\bot^{(i)})$.

By induction on i. Base case: $L_\top^{(0)}, L_\bot^{(0)}$ is correct immediately. On each (subsequent) iteration. Not that KB$/(L_\top^{(i-1)}, L_\bot^{(i-1)})$ is computed by deleting rules with "not ℓ" in their body for $\ell \in L_\top^{(i-1)}$, where

$L_\top^{(i-1)} \supseteq L_\top^{(i-2)}$ by the main lemma, and by dropping the occurrences of not ℓ for $\ell \in L_\bot^{(i-1)}$, where also $L_\bot^{(i-1)} \supseteq L_\bot^{(i-2)}$ by the main lemma. So, KB/$(L_\top^{(i-1)}, L_\bot^{(i-1)})$ is computed correctly in the first loop. We next argue—

Claim (2): $L_\top^{(i)}$ and $L_\bot^{(i)}$ are computed correctly within $O(n(\text{KB}))$ steps.

Proof by induction on length of the chaining proofswed. On iteration j, $L_\top^{(i)}$ contains every ℓ provable in at most j steps, and $L_\bot^{(i)}$ contains *none* of the literals provable in j steps (when we drop the naf'd literals). Base case trivially holds for $j = 0$. Given the claim for iteration $j-1$, in iteration j, when we consider the rule used to prove a literal ℓ (on the jth step of any chaining proof) all of the literals in its body must have been provable in $j' < j$ steps, and hence are present in $L_\top^{(i)}$ or absent from $L_\bot^{(i)}$, as appropriate. Therefore, we do ensure that ℓ is added to $L_\top^{(i)}$ or removed from $L_\bot^{(i)}$ on the jth iteration, as appropriate. This completes the induction step. Noting that there are only $2n$ literals and to continue to the next iteration, we must have eithere added a literal to $L_\top^{(i)}$ or deleted one from $L_\bot^{(i)}$, we find that there are at most $2n$ iterations, each taking $O(|\text{KB}|)$ time. After the final iteration, since any shortest proof of a literal ℓ^* must derive at least one new literal on each step, but no further literals are provable from $\text{KB}' = \frac{KB}{(L_\top^{(i-1)}, L_\bot^{(i-1)})}$, we find that $L_\top^{(i)}$ contains every provable literal in KB$'$ and $L_\bot^{(i)}$ omits precisely the literals provable when naf'd literals are dropped. Finally, by the main lemma, since we can only add literals to $L_\top^{(i)}$ and $L_\bot^{(i)}$ and there are at most $2n$ literals, the main loop runs for at most $2n$ iterations. Since each iteration takes $O(n|\text{KB}|)$ time, the algorithm overall runs in time $O(n^2|\text{KB}|)$.

Finally finally:

(1) $\ell \in L_\top^{(i+1)}$ if and only if ℓ is provable in KB/$(L_\top^{(i)}, L_\bot^{(i)})$ but at convergence $L_\top^{(i+1)} = L_\top^{(i)}$.

(2) Likewise, if $\ell \in L_\bot^{(i+1)}$, then it must be ℓ is not provable from KB/$(L_T^{(i)}, L_\bot^{(i)}$ by our LPM theorem (even if we disregard the remaining occurrences of not ℓ; which we cannot be included in $L_\top^{(i)}$), so ℓ is also not provable in KB/$(L_\top^{(i)}, L_\bot^{(i)})$, where $L_\bot^{(i+1)} = L_\bot^{(i)}$.