Einführung in Wissensbasierte Systeme

Closed World Assumption

I leave out all the background information about the ideas behind the CWA, since it is covered in the slides. The aim of this document is to answer some frequently asked questions through examples.

Example from the slides

We start by looking at the theory

$$T = \{\forall x(p(x) \vee q(x)), p(a), q(b)\}$$

The predicates in T are $p/1$ and $q/1$, and the terms are a and b. Thus, the ground atoms in T are $\{p(a), p(b), q(a), q(b)\}$. (A combination of all predicates with all terms in T. It does not matter if they actually appear in T or not, what matters is that they can be constructed from the relevant predicates, terms or functions appearing in the theory.)

Computing T_{asm}**:**
For a theory T, we defined T_{asm} to be $T_{asm} := \{\neg P \mid P \text{ is a ground atom}, T \not\models P\}$.
Looking at our set of ground atoms again, we conclude:

	$p(a)$	$p(b)$	$q(a)$	$q(b)$
$T \models$	✓	×	×	✓

Therefore, $T_{asm} = \{\neg p(b), \neg q(a)\}$

Computing $CWA(T)$**:**
By definition $CWA(T) = \mathsf{Cn}(T \cup T_{asm})$. Furthermore, we say that $CWA(T)$ is consistent, because we can find a model for it, or equivalently, because it does not hold that $CWA(T) \models \perp$, or yet again equivalently, because there is no P for which $CWA(T) \models P$ and $CWA \models \neg P$ both hold.

This, however, is all under the assumption that a and b are the only constants in the language of the theory T.

Assume now, we have a theory T', defined exactly the same as our previous T, but its language also includes another constant – c, T'_{asm} would also include $\not p(c)$ and $\neg q(c)$. Thus, these literals would also be included in $CWA(T')$. But then, $\forall x(p(x) \vee q(x)), \neg p(c) \models q(c)$ and similarly $\forall x(p(x) \vee q(x)), \neg q(c) \models p(c)$.

Since all the formulas on the left hand side are indeed in this new $CWA(T')$, it means that $q(c)$ and $p(c)$ are also in $CWA(T')$. So, now that we know that $CWA(T') \models p(c)$ and $CWA(T') \models \neg p(c)$, we can easily see that it is inconsistent.

Practical example (WS 18):
Consider the theory

$$T = \{\forall x(\neg P(x) \to Q(x)), \neg P(b), R(a), \forall x(R(x) \to P(x))\}$$

1. Determine $CWA(T)$:

   We once again start by identifying all the ground atoms in the language of T, which we construct by combining all the predicate symbols ($P/1$, $Q/1$ and $R/1$) with all the terms (a and b). We check which of the ground atoms follow from our theory.

	$P(a)$	$P(b)$	$Q(a)$	$Q(b)$	$R(a)$	$R(b)$
$T \models$	✓	×	×	✓	✓	×

   Note that $T \models \neg P(b)$ and $T \models \neg R(b)$, but for us, it is only the positive literals that actually matter.

   Therefore, according to the definition $T_{asm} = \{\neg P(b), \neg Q(a), \neg R(b)\}$.

   *Note again that it is **wrong** to exclude $\neg P(b)$ and $R(b)$ from T_{asm} just because they are also in $\mathsf{Cn}(T)$. This is a common mistake.*

   For $CWA(T)$ it suffices to define $CWA(T) = \mathsf{Cn}(T \cup T_{asm})$ after already having defined T_{asm} properly.

2. Determine the generalization $CWA^Q(T)$:

   Here, we proceed exactly like in the previous subtask, except we only concern ourselves with atoms in which the predicate symbol Q is used. Thus $T^Q_{asm} = \{\neg Q(a)\}$ and $CWA^Q(T) = \mathsf{Cn}(T \cup T^Q_{asm})$.

3. Prove that $CWA(T)$ is consistent.

   First we argue that T is consistent. This can be done by providing a model for T and is omitted here. For the main proof, we use the theorem on Slide 19.

   Proof. ($\Longrightarrow$) If $CWA(T)$ is inconsistent, then there are ground atoms $A_1, \dots, A_n$ for which $T \models A_1 \vee \dots \vee A_n$, but $T \not\models A_i$ for all $i \in [1, n]$.

   The atoms that cannot be derived from T are $P(b), Q(a), R(b)$. Thus, we must show that no disjunction follows from T which consists of any two of these ground atoms. It suffices to show that $T \not\models P(b) \vee Q(a) \vee R(b)$, as if this is not the case, then T surely doesn't model a "stronger" statement, with any of the conjuncts removed. To do so, we find an interpretation $\mathcal{I}$ such that $\mathcal{I} \models T$, but $\mathcal{I} \not\models P(b) \vee Q(a) \vee R(b)$. Through that, we can show that $CWA(T)$ is consistent. □

   Lastly, given that we know that $CWA(T)$ is consistent, we can easily[1] come to the conclusion that $CWA^Q(T)$ also has to be consistent.

[1] "Easily" is defined as "by just reasoning about properties". Writing out this simple reasoning entirely could potentially spoil parts of the Übung, so I chose to omit it. In your own work, please also provide your reasoning. :)