## 1 Craig Interpolation Theorem

**Remark 17.1.1** As a second serious application of the Model Existence Theorem we prove Craig's Interpolation Theorem for propositional logic. The first-order version of this result has many important consequences which we will consider more fully later. The proof of the first-order theorem is due to William Craig in 1957, and is probably the last deep elementary theorem about first-order logic. I am not sure if the interpolation theorem for propositional logic predates Craig's work or not.

The proof given here is due to Raymond Smullyan, modified a bit by Melvin Fitting.

**Definition 17.1.2** (Interpolant) A proposition $\gamma$ is an *interpolant* for the implication $\alpha \to \beta$ if every propositional symbol in $\gamma$ occurs in both $\alpha$ and $\beta$ and both $\models \alpha \to \gamma$ and $\models \gamma \to \beta$.

**Example 17.1.3** The implication $(P \vee (Q \wedge R)) \to (P \vee \neg\neg Q))$ has $P \vee Q$ as an interpolant. The implication $(P \wedge \neg P) \to Q$ has $\bot$ as an interpolant. The implication $Q \to (P \vee \neg P)$ has $\top$ as an interpolant.

**Theorem 17.1.4** (Craig Interpolation Theorem) If $\alpha \to \beta$ is a tautology, then it has an interpolant.

*Proof.* It is useful to introduce the following notation for this proof. If $S$ is a finite set of propositions, then $\langle S \rangle$ is the conjunction of the members of $S$. If $S = \{\gamma\}$ then $\langle S \rangle = \gamma$ and if $S = \emptyset$ take $\langle S \rangle$ to be $\top$. Call a finite set of propositions *Craig-consistent*, provided there is a partition of $S$ into two subsets $S_1$ and $S_2$ (that is, $S = S_1 \cup S_2$ and $\emptyset = S_1 \cap S_2$) such that $\langle S_1 \rangle \to \neg \langle S_2 \rangle$ has no interpolant. We will show that Craig-consistentcy is a propositional consistency property.

Suppose that Craig consistency is a propositional consistency property, and we will derive the theorem in the contrapositive form: if $\alpha \to \beta$ has no interpolant, then $\alpha \to \beta$ is not a tautology. Let $S = \{\alpha \to \neg\beta\}$, and consider the partition $S_1 = \{\alpha\}$ and $S_2 = \{\neg\beta\}$. If $\alpha \to \neg(\neg\beta)$, which is $\langle S_1 \rangle \to \neg \langle S_2 \rangle$, then this would be an interpolant for $\alpha \to \beta$ as well, hence it has no interpolant. So, $\{\alpha, \neg\beta\}$ is Craig-consistent, and thus satisfiable by the Model Existence Theorem 16.2.1. It follows that $\alpha \to \beta$ cannot be a tautology.

We complete the proof by showing that Craig-consistency is a propositional consistency property. It is more straightforward to show that Craig-inconsistency (the negation of Craig consistency) is a propositional inconsistency property. This is sufficient by Lemma 16.1.6. To show a finite set $S$ is Craig-inconsistency amounts to showing that any partition into $S_1$ and $S_2$ the implication $\langle S_1 \rangle \to \neg \langle S_2 \rangle$ has an interpolant. We will say that $S_1$ and $S_2$ has an interpolant if this is so.

Let $S$ be a finite set of propositions. Suppose $P, \neg P \in S$, for some propositional symbol $P$. Consider any partition $S_1$ and $S_2$ of $S$. If $P, \neg P$ is in $S_1$, then $\langle S_1 \rangle$ is tautologically equivalent to $\bot$, so we can take $\bot$ to be the interpolant. If $P, \neg P \in S_2$, then $\neg \langle S_2 \rangle$ is tautologically equivalent to $\top$, so we can take the interpolant to be $\top$. Otherwise, $P$ and $\neg P$ are in different partitions. If $P \in S_1$ take the interpolant to be $P$, since $\langle S_1 \rangle$ is tautologically equivalent to a conjunction with $P$ as one conjunction and $\neg \langle S_2 \rangle$ is tautologically equivalent to a disjunction with $\neg\neg P$ as a disjunct. And if $P \in S_2$, then take the interpolant to be $\neg P$, since $\langle S_1 \rangle$ is tautologically equivalent to a conjunction with $\neg P$ as a conjunct and $\neg \langle S_2 \rangle$ is is tautologically equivalent to a disjunction with $\neg P$ as a disjunct. In any case, $S$ is Craig-inconsistent, since for every way of partitioning $S$ into $S_1$ and $S_2$, $\langle S_1 \rangle \to \neg \langle S_2 \rangle$ has an interpolant. Very similar reasoning by cases shows that if $\bot \in S$ or $\neg\top \in S$ then $S$ will be Craig-inconsistent. So, Craig inconsistency satisfies condition (C1).

Let $\alpha$ is a type-$A$ proposition with components $\alpha_1$ and $\alpha_2$ and suppose that $S \cup \{\alpha_1, \alpha_2\}$ is Craig inconsistent. Consider any partition of $S$ into $S_1$ and $S_2$. Let $\gamma_1$ be an interpolant for $S_1 \cup \{\alpha_1, \alpha_2\}$ and $S_2$, let $\gamma_2$ be an interpolant for $S_1$ and $S_2 \cup \{\alpha_1, \alpha_2\}$ Since $\alpha \simeq \alpha \wedge \beta$ and $\neg\alpha \simeq \neg\alpha \vee \neg\beta$, it follows that

$\langle S_1 \cup \{\alpha_1, \alpha_2\}\rangle \simeq \langle S_1 \cup \{\alpha\}\rangle$ and $\neg \langle S_2 \cup \{\alpha_1, \alpha_2\}\rangle \simeq \neg \langle S_1 \cup \{\alpha\}\rangle$. So, $\gamma_1$ is an interpolant for $S_1 \cup \{\alpha\}$ and $S_2$ and $\gamma_2$ is an interpolant for $S_1$ and $S_2 \cup \{\alpha\}$. Since $S_1$ and $S_2$ was an arbitrary partition of $S$, $S \cup \{\alpha\}$ is Craig-inconsistent. So Craig-inconsistency satisfies (C2).

Let $\beta$ is a type-$A$ proposition with components $\beta_1$ and $\beta_2$ and suppose that both $S \cup \{\beta_1\}$ and $S \cup \{\beta_2\}$ are Craig inconsistent. Consider any partition of $S$ into $S_1$ and $S_2$. Let $\gamma_1$ be an interpolant for $S_1 \cup \{\beta_1\}$ and $S_2$ and $\delta_1$ be an interpolant for $S_1$ and $S_2 \cup \{\beta_1\}$. Similarly, let $\gamma_2$ be an interpolant for $S_1 \cup \{\beta_2\}$ and $S_2$ and $\delta_2$ be an interpolant for $S_1$ and $S_2 \cup \{\beta_2\}$. Note that $\beta \simeq \beta_1 \vee \beta_2$ and $\neg \beta \simeq \neg \beta_1 \wedge \neg \beta_2$. So,

$$\langle S_1 \cup \{\beta\}\rangle \simeq (\langle S_1 \cup \{\beta_1\}\rangle \vee \langle S_1 \cup \{\beta_2\}\rangle)$$

Thus, $\gamma_1 \vee \gamma_2$ is an interpolant for $S_1 \cup \{\beta\}$ and $S_2$, since

$$\models \langle S_1 \cup \{\beta_1\}\rangle \rightarrow \gamma_1 \vee \gamma_2 \quad \text{and} \quad \models \langle S_1 \cup \{\beta_1\}\rangle \rightarrow \gamma_1 \vee \gamma_2 \quad \text{and} \quad \models \gamma_1 \vee \gamma_2 \rightarrow \langle S_2\rangle.$$

Also,

$$\neg \langle S_2 \cup \{\beta\}\rangle \simeq (\neg \langle S_2 \cup \{\beta_1\}\rangle \wedge \neg \langle S_2 \cup \{\beta_2\}\rangle)$$

Thus, $\delta_1 \wedge \delta_2$ is an interpolant for $S_1$ and $S_2 \cup \{\beta\}$ since

$$\models \langle S_1\rangle \rightarrow \delta_1 \wedge \delta_2 \quad \text{and} \quad \models (\delta_1 \wedge \delta_2) \rightarrow \neg \langle S_2 \cup \{\beta_1\}\rangle \quad \text{and} \quad (\delta_1 \wedge \delta_2) \rightarrow \neg \langle S_2 \cup \{\gamma_2\}\rangle$$

and $\langle S_1\rangle \rightarrow (\delta_1 \wedge \delta_2)$. Since $S_1$ and $S_2$ was an arbitrary partition of $S$, $S \cup \{\beta\}$ is Craig-inconsistent. So Craig-inconsistency satisfies (C3).

□

**Example 17.1.5** The proof of the Craig Interpolation Theorem suggests an algorithm for finding the interpolant using the unified notation. Consider the tautology $(P \rightarrow (Q \rightarrow R)) \rightarrow (Q \rightarrow (P \rightarrow R))$.

1. Start with the set $S = \{P \rightarrow (Q \rightarrow R), \neg (Q \rightarrow (P \rightarrow R))\}$ with the partition $S_1 = \{P \rightarrow (Q \rightarrow R)\}$ and $S_2 = \{\neg (Q \rightarrow (P \rightarrow R))\}$ in the notation of the theorem. We are looking for an interpolant for this partition.

2. It is sufficient to find an interpolant for $\{P \rightarrow (Q \rightarrow R)\}$ and $\{Q, \neg (P \rightarrow R)\}$. This is a type-$A$ case.

3. It is sufficient to find an interpolant for $\{P \rightarrow (Q \rightarrow R)\}$ and $\{Q, P, \neg R\}$. This is a type-$A$ case.

4. The proposition in $\{P \rightarrow (Q \rightarrow R)\}$ is type-$B$, so if $\gamma_1$ is an interpolant for $\{\neg P\}$ and $\{Q, P, \neg R\}$ and $\gamma_2$ is an interpolant for $\{Q \rightarrow R\}$ and $\{Q, P, \neg R\}$, then $\gamma_1 \vee \gamma_2$ is the desired interpolant.

   (a) $\gamma_1 = \neg P$ is an interpolant for $\{\neg P\}$ and $\{Q, P, \neg R\}$. It is worth verifying this:

   $$\models \neg P \rightarrow \neg P \quad \text{and} \quad \models \neg P \rightarrow (\neg Q \vee \neg P \vee \neg\neg R).$$

   (b) We now need an interpolant for $\{Q \rightarrow R\}$ and $\{Q, P, \neg R\}$. It is sufficient to find an interpolant $\gamma_3$ for $\{\neg Q\}$ and $\{Q, P, \neg R\}$ and an interpolant $\gamma_4$ for $\{R\}$ and $\{Q, P, \neg R\}$.

      i. $\gamma_3 = \neg Q$ is an interpolant for $\{\neg Q\}$ and $\{Q, P, \neg R\}$.

      ii. $\gamma_4 = R$ is an interpolant for $\{R\}$ and $\{Q, P, \neg R\}$. We verify this:

      $$\models R \rightarrow R \quad \text{and} \quad \models R \rightarrow (\neg Q \vee \neg P \vee \neg\neg R).$$

   (c) We have $\gamma_3 = \neg Q$ and $\gamma_4 = R$, so the required interpolant for step b is $\neg Q \vee R$.

5. We now have $\gamma_1 = \neg P$ and $\gamma_2 = \neg Q \vee R$, so the required interpolant for step 4 is $\neg P \vee \neg Q \vee R$. This is also the required interpolant for $(P \rightarrow (Q \rightarrow R)) \rightarrow (Q \rightarrow (P \rightarrow R))$.

So, $\neg P \vee \neg Q \vee R$ is an interpolant for $(P \rightarrow (Q \rightarrow R)) \rightarrow (Q \rightarrow (P \rightarrow R))$. We verify this:

$$\models (P \rightarrow (Q \rightarrow R)) \rightarrow (\neg P \vee \neg Q \vee R) \quad \text{and} \quad \models (\neg P \vee \neg Q \vee R) \rightarrow (Q \rightarrow (P \rightarrow R)).$$