Azzolini Riccardo 2020-06-08

Teorema di Herbrand

1 Espansione di Herbrand

Definizione: Sia una formula chiusa e in forma di Skolem, cioè $\varphi = \forall x_1 \dots \forall x_n \psi$, dove in ψ non compaiono quantificatori (ovvero ψ è una matrice). L'**espansione di Herbrand** di φ è l'insieme

$$E(\varphi) = \big\{\psi[t_1/x_1, \dots, t_n/x_n] \,\big|\, t_1, \dots, t_n \in H(\varphi)\big\}$$

di tutte le formule chiuse ottenibili sostituendo alle variabili $x_1, \dots, x_n$ che compaiono nella matrice ψ dei termini chiusi appartenenti all'universo di Herbrand di φ.

1.1 Esempi

- Sia
$$\varphi = \forall y \underbrace{(M(c, y) \land \neg M(d, y))}_{\psi}$$
Questa formula contiene due costanti, c e d, quindi l'universo di Herbrand è $H(\varphi) = \{c, d\}$, da cui segue che l'espansione di Herbrand è:
$$\begin{aligned} E(\varphi) &= \{\psi[t/y] \mid t \in \{c, d\}\} = \{\psi[c/y],\ \psi[d/y]\} \\ &= \{M(c, c) \land \neg M(d, c),\ M(c, d) \land \neg M(d, d)\} \end{aligned}$$
In particolare, l'universo di Herbrand è finito, quindi anche l'espansione di Herbrand è finita.
- Si consideri la formula
$$\varphi = \forall x \underbrace{(P(x) \land \neg P(f(x)))}_{\psi}$$
Essa non contiene simboli di costante, ma contiene un simbolo di funzione $f^{(1)}$. Allora, aggiungendo una nuova costante c, si genera un universo di Herbrand infinito,
$$H(\varphi) = \{c, f(c), f(f(c)), \dots\}$$

e di conseguenza anche l'espansione di Herbrand è infinita:

$$\begin{aligned}E(\varphi) &= \{\psi[t/x] \mid t \in H(\varphi)\} \\ &= \{\psi[c/x],\ \psi[f(c)/x],\ \psi[f(f(c))/x], \dots\} \\ &= \{P(c) \land \neg P(f(c)),\ P(f(c)) \land \neg P(f(f(c))), \\ &\qquad P(f(f(c))) \land \neg P(f(f(f(c)))), \dots\}\end{aligned}$$

- Sia

  $$\varphi = \forall x \forall y \underbrace{(A(x, f(y)) \land B(y, f(x)))}_{\psi}$$

  L'universo di Herbrand è generato come nell'esempio precedente:

  $$H(\varphi) = \{c, f(c), f(f(c)), \dots\}$$

  L'espansione di Herbrand diventa invece più complessa, perché bisogna considerare le sostituzioni delle due variabili quantificate con tutte le possibili coppie di termini chiusi in $H(\varphi)$:

  $$\begin{aligned}E(\varphi) &= \{\psi[t_1/x, t_2/y] \mid t_1, t_2 \in H(\varphi) \\ &= \{\psi[c/x, c/y],\ \psi[c/x, f(c)/y],\ \psi[f(c)/x, c/y], \dots\} \\ &= \{A(c, f(c)) \land B(c, f(c)),\ A(c, f(f(c))) \land B(f(c), f(c)), \\ &\qquad A(f(c), f(c)) \land B(c, f(f(c))), \dots\}\end{aligned}$$

2 Espansioni di Herbrand e formule proposizionali

Per definizione, l'espansione di Herbrand è formata solo da formule chiuse senza quantificatori, la cui verità può quindi essere studiata senza bisogno di considerare assegnamenti. Allora, si potrebbe pensare di interpretare queste formule come se fossero formule proposizionali.

Si consideri ad esempio $\varphi = \forall y(M(c, y) \land \neg M(d, y))$, che, come già visto, genera l'universo di Herbrand $H(\varphi) = \{c, d\}$ e l'espansione di Herbrand

$$E(\varphi) = \{M(c, c) \land \neg M(d, c),\ M(c, d) \land \neg M(d, d)\}$$

Siccome le formule atomiche che compaiono in questo insieme,

$$M(c, c), M(d, c), M(c, d), M(d, d)$$

sono formule chiuse, il loro valore di verità in un modello è determinato semplicemente dall'interpretazione del predicato M in quel modello (e non, come già detto, dall'assegnamento). Allora, sostituendo a queste formule atomiche delle variabili proposizionali,

$$X_1 = M(c, c) \quad X_2 = M(d, c) \quad X_3 = M(c, d) \quad X_4 = M(d, d)$$

si può riscrivere l'espansione di Herbrand come un insieme di formule proposizionali:

$$E(\varphi) = \{X_1 \wedge \neg X_2,\ X_3 \wedge \neg X_4\}$$

Dato un modello $\mathcal{M}$, se si costruisce una valutazione proposizionale $v_\mathcal{M}$ che associa alla variabile X_1 il valore di verità della corrispondente formula atomica $M(c,c)$ in $\mathcal{M}$, e così anche per X_2, X_3, X_4, si ha immediatamente che $\mathcal{M} \models E(\varphi)$ se e solo se $v_\mathcal{M}$ verifica la "rappresentazione" di $E(\varphi)$ come insieme di formule proposizionali. Dunque, si può ridurre il problema di studiare la soddisfacibilità di $E(\varphi)$ a quello di studiare la soddisfacibilità di un insieme di formule proposizionali.

2.1 Notazione

Per passare da $E(\varphi)$ al corrispondente insieme di formule proposizionali, non è in realtà necessario introdurre dei nomi di variabili proposizionali (e, anzi, ciò complica inutilmente la presentazione dell'argomento).

Invece, posto $\Sigma(\varphi)$ l'insieme delle formule atomiche chiuse che occorrono in $E(\varphi)$, si possono vedere gli elementi di $\Sigma(\varphi)$ direttamente come nomi di variabili proposizionali, e quindi gli elementi di $E(\varphi)$ come formule proposizionali costruite a partire da tali variabili.

Per indicare quando si sta leggendo una formula atomica chiusa $M(c,d)$ come il nome di una variabile proposizionale, si introduce la notazione $\overline{M(c,d)}$. Analogamente, si indicano con:

- $\overline{\Sigma(\varphi)}$ l'insieme delle variabili proposizionali corrispondenti alle atomiche in $\Sigma(\varphi)$;
- $\overline{E(\varphi)}$ l'insieme di formule proposizionali corrispondenti alle formule in $E(\varphi)$;
- $\overline{\psi} \in \overline{E(\varphi)}$ la formula proposizionale corrispondente alla formula $\psi \in E(\varphi)$.

Rivedendolo con questa notazione, l'esempio precedente diventa:

$$\begin{aligned}
E(\varphi) &= \{\overbrace{M(c,c) \wedge \neg M(d,c)}^{\psi_1},\ \overbrace{M(c,d) \wedge \neg M(d,d)}^{\psi_2}\} \\
\Sigma(\varphi) &= \{M(c,c), M(d,c), M(c,d), M(d,d)\} \\
\overline{\Sigma(\varphi)} &= \{\overline{M(c,c)}, \overline{M(d,c)}, \overline{M(c,d)}, \overline{M(d,d)}\} \\
\overline{E(\varphi)} &= \{\underbrace{\overline{M(c,c)} \wedge \neg \overline{M(d,c)}}_{\overline{\psi_1}},\ \underbrace{\overline{M(c,d)} \wedge \neg \overline{M(d,d)}}_{\overline{\psi_2}}\}
\end{aligned}$$

2.2 Definizione formale

Dato un modello di Herbrand $\mathcal{H}$ per φ, si può definire una valutazione $v_\mathcal{H}$ per le variabili proposizionali in $\overline{\Sigma(\varphi)}$ in questo modo:

$$\widetilde{\forall} \bar{\theta} \in \overline{\Sigma(\varphi)} \quad v_\mathcal{H}(\bar{\theta}) = 1 \iff \mathcal{H} \models \theta$$

Si può dimostrare, mediante una banale induzione strutturale, che $\mathcal{H}$ soddisfa $E(\varphi)$ se e solo se $v_\mathcal{H}$ soddisfa $\overline{E(\varphi)}$. Formalmente:

Proposizione: Per ogni formula $\psi \in E(\varphi)$, $\mathcal{H} \models \psi$ se e solo se $v_\mathcal{H} \models \bar{\psi}$.

Tornando all'esempio di prima,

$$\varphi = \forall y(M(c,y) \land \neg M(d,y)) \qquad H(\varphi) = \{c, d\}$$
$$E(\varphi) = \{\overbrace{M(c,c) \land \neg M(d,c)}^{\psi_1},\ \overbrace{M(c,d) \land \neg M(d,d)}^{\psi_2}\}$$
$$\overline{E(\varphi)} = \{\underbrace{\overline{M(c,c)} \land \neg\overline{M(d,c)}}_{\overline{\psi_1}},\ \underbrace{\overline{M(c,d)} \land \neg\overline{M(d,d)}}_{\overline{\psi_2}}\}$$

e considerando il modello di Herbrand $\mathcal{H} = (H(\varphi), J)$ tale che $J(M) = \{(c,c), (c,d)\}$, è immediato verificare che $\mathcal{H} \models E(\varphi)$. Costruendo poi la valutazione associata a questo modello,

$$\begin{aligned}
v_\mathcal{H}(\overline{M(c,c)}) &= 1 \qquad \text{perché } (c,c) \in J(M) \implies \mathcal{H} \models M(c,c) \\
v_\mathcal{H}(\overline{M(d,c)}) &= 0 \qquad (d,c) \notin J(M) \implies \mathcal{H} \not\models M(d,c) \\
v_\mathcal{H}(\overline{M(c,d)}) &= 1 \qquad (c,d) \in J(M) \implies \mathcal{H} \models M(c,d) \\
v_\mathcal{H}(\overline{M(d,d)}) &= 0 \qquad (d,d) \notin J(M) \implies \mathcal{H} \not\models M(d,d)
\end{aligned}$$

si verifica immediatamente anche che $v_\mathcal{H} \models \overline{E(\varphi)}$.

3 Teorema di Herbrand

Teorema (di Herbrand): Sia φ una formula chiusa e in forma di Skolem. φ è soddisfacibile se e solo se $E(\varphi)$ è soddisfacibile come insieme di formule proposizionali.

Dimostrazione: φ è in forma di Skolem (cioè $\varphi = \forall x_1 \dots \forall x_n \psi$, dove ψ è una matrice) ed è chiusa. Allora, ricordando che ogni tale formula è soddisfacibile se e solo se ha un modello di Herbrand (un teorema dimostrati in precedenza), si ha che:

$$\begin{aligned}
\varphi \text{ è soddisfacibile} &\iff \varphi \text{ ha un modello di Herbrand } \mathcal{H} \\
&\iff \mathcal{H} \models \forall x_1 \dots \forall x_n \psi \\
&\iff \widetilde{\forall} t_1, \dots, t_n \in H(\varphi) \quad (\mathcal{H}, [t_1/x_1, \dots, t_n/x_n]) \models \psi \\
&\iff \widetilde{\forall} t_1, \dots, t_n \in H(\varphi) \quad \mathcal{H} \models \underbrace{\psi[t_1/x_1, \dots, t_n/x_n]}_{\theta \in E(\varphi)} \\
&\iff \widetilde{\forall} \theta \in E(\varphi) \quad \mathcal{H} \models \theta \\
&\iff \widetilde{\forall} \bar{\theta} \in \overline{E(\varphi)} \quad v_{\mathcal{H}} \models \bar{\theta} \qquad \text{(proposizione precedente)} \\
&\iff v_{\mathcal{H}} \models \overline{E(\varphi)} \\
&\iff \overline{E(\varphi)} \text{ è soddisfacibile}
\end{aligned}$$

3.1 Combinazione con il teorema di compattezza

Combinando il teorema di Herbrand con il teorema di compattezza (un insieme potenzialmente infinito di formule proposizionali è soddisfacibile se e solo se ogni suo sottoinsieme finito è soddisfacibile), si ottiene:

Teorema: Una formula φ chiusa e in forma di Skolem è soddisfacibile se e solo se ogni sottoinsieme finito di $\overline{E(\varphi)}$ è soddisfacibile.

4 Problema della soddisfacibilità

Data una formula φ, supponendo che $\overline{E(\varphi)} = \{\overline{\varphi_1}, \overline{\varphi_2}, \overline{\varphi_3}, \dots\}$ sia soddisfacibile, per capire se φ è soddisfacibile usando il teorema precedente bisogna controllare la soddisfacibilità di un numero infinito di sottoinsiemi finiti:

$$\{\overline{\varphi_1}\},\ \{\overline{\varphi_1}, \overline{\varphi_2}\},\ \{\overline{\varphi_1}, \overline{\varphi_2}, \overline{\varphi_3}\},\ \dots$$

se uno di questi è insoddisfacibile, allora si può concludere che la formula φ è insoddisfacibile; altrimenti, bisogna continuare a controllare, potenzialmente all'infinito.

Si dice allora che stabilire se una formula sia insoddisfacibile (o, analogamente, se essa sia valida) è un **problema semidecidibile**:

- se la formula è insoddisfacibile, esiste un sottoinsieme finito di $\overline{E(\varphi)}$ insoddisfacibile, e prima o poi lo si trova;
- se la formula è invece soddisfacibile, non esiste un sottoinsieme finito di $\overline{E(\varphi)}$ insoddisfacibile, e quindi la procedura continua all'infinito.

4.1 Esempio

Sia $\varphi = \forall x(P(x) \wedge \neg P(f(x)))$. Come visto in uno degli esempi precedenti, questa formula genera l'universo di Herbrand

$$H(\varphi) = \{c, f(c), f(f(c)), \ldots\}$$

da cui:

$$\begin{aligned}E(\varphi) &= \{P(c) \wedge \neg P(f(c)),\ P(f(c)) \wedge \neg P(f(f(c))), \ldots\} \\ \overline{E(\varphi)} &= \{\overline{P(c)} \wedge \neg \overline{P(f(c))},\ \overline{P(f(c))} \wedge \neg \overline{P(f(f(c)))}, \ldots\}\end{aligned}$$

Per rendere più evidente la struttura di quest'ultimo insieme, si possono sostituiscono le atomiche chiuse con nomi di variabili proposizionali "semplici":

$$X_1 = \overline{P(c)} \quad X_2 = \overline{P(f(c))} \quad X_1 = \overline{P(f(f(c)))}$$
$$\overline{E(\varphi)} = \{X_1 \wedge \neg X_2,\ X_2 \wedge \neg X_3, \ldots\}$$

Considerando i sottoinsiemi finiti, secondo la procedura descritta prima:

1. $\{X_1 \wedge \neg X_2\}$ è soddisfacibile;
2. $\{X_1 \wedge \neg X_2,\ X_2 \wedge \neg X_3\}$ è *insoddisfacibile*.

Quindi, attraverso il teorema di Herbrand, si deduce che φ è insoddisfacibile.

4.2 Casi in cui il problema è decidibile

Ci sono dei casi particolari nei quali cui la procedura termina sempre, ovvero nei quali il problema della soddisfacibilità di una formula diventa decidibile: se $H(\varphi)$ è finito, allora anche $\overline{E(\varphi)}$ è finito, quindi si può decidere se è soddisfacibile o no in un numero finito di passi.

$H(\varphi)$ è finito quando non ci sono funzioni: dunque, per le formule chiuse in forma di Skolem che non contengono simboli di funzione, il problema della soddisfacibilità è decidibile.

È importante ricordare che ciò vale, appunto, solo per le formula che non contengono simboli di funzione *in forma di Skolem*, e non, in generale, per quelle che non contengono funzioni prima di essere skolemizzate. Ad esempio, la formula $\varphi = \forall x \exists y M(x, y)$ non contiene funzioni, ma la sua forma di Skolem $\varphi^S = \forall x M(x, f(x))$ contiene invece il simbolo $f^{(1)}$, quindi essa *non* rientra nei casi in cui il problema della soddisfacibilità è decidibile