知識工学　第5回

二宮　崇

教科書と資料

- 教科書
  - Artificial Intelligence: A Modern Approach (3rd Edition): Stuart Russell, Peter Norvig (著), Prentice Hall, 2009

- この講義のウェブサイト

http://aiweb.cs.ehime-u.ac.jp/~ninomiya/ke/

今日の内容

- 命題論理の推論
  - 証明の戦略その3
    - 融合法

命題論理における推論パターン(§7.5) 証明の戦略その3 (融合法)

- 証明の戦略その1やその2で証明できたときは、$KB \vDash \alpha$ となることがわかる

しかし…

- なかなか証明できないとき、証明が本当にできないのか否かわからない
- $KB \vDash \alpha$が成り立つのか成り立たないのかよくわからない
- どのような証明手続きを踏めば証明できるのか定かではない

そこで…

- 健全かつ完全な推論手続きである融合法

命題論理における推論パターン(§7.5)
証明の戦略その3 (融合法)

- 融合法
  - $KB \land \neg\alpha$が充足不能であることを示すことで、$KB \vDash \alpha$を示す(背理法)

命題論理における推論パターン(§7.5) 証明の戦略その3 (融合法)

- 融合法の手続き
  1. $KB \land \neg\alpha$を連言標準形(conjunctive normal form: CNF)に変換し、その結果を節集合KB'とする
  2. 節集合KB'に対し**融合規則**を適用し、$KB' \vDash R$となるRをみつけRを節集合KB'に追加する(KB'は更新される)。この操作(推論)を$KB' \vdash R$と書くことにする。
  3. 追加できる新しい節がない場合、KBからαが伴意しないことがわかる
  4. 偽$(= 0)$が得られたら、KBからαが伴意することがわかる
  5. 2.に戻る

命題論理における推論パターン(§7.5) 証明の戦略その3 (融合法)

- 連言標準形(conjunctive normal form: CNF)
  - $(l_{1,1} \vee l_{1,2} \vee \cdots) \wedge (l_{2,1} \vee l_{2,2} \vee \cdots) \wedge \cdots \wedge (l_{n,1} \vee l_{n,2} \vee \cdots)$
  - $l_{i,j}$はリテラルと呼ばれる
  - リテラルは命題記号または命題記号の否定
  - 命題記号は正リテラル、命題記号の否定は負リテラルと呼ばれる
  - $(l_{i,1} \vee l_{i,2} \vee \cdots)$は節と呼ばれる
- $KB \wedge \neg\alpha$をCNFに変換した結果の節集合をKB'とする

命題論理における推論パターン(§7.5) 証明の戦略その3 (融合法)

- 融合規則 (もっとも簡単な形、選言的三段論法)

$$\frac{l \vee m \qquad \neg m}{l}$$

- 導出 (または真理値表を作成)
  - 分配則より、$(l \vee m) \wedge \neg m \equiv (l \wedge \neg m) \vee 0 \equiv l \wedge \neg m$
  - 縮小律(And除去)により、$l \wedge \neg m \vDash l$
  - 従って、$(l \vee m) \wedge \neg m \vDash l$

命題論理における推論パターン(§7.5)
証明の戦略その3 (融合法)

- 融合規則(長さ2の節の場合)

$$\frac{l \lor m \qquad \neg m \lor n}{l \lor n}$$

- 導出(または真理値表を作成)
  - 分配則を2回適用
  - And除去を適用
  - 結論部が得られる

命題論理における推論パターン(§7.5) 証明の戦略その3 (融合法)

- 融合規則(一般)

$$\frac{l_1 \vee \cdots \vee l_i \vee m \qquad \neg m \vee n_1 \vee \cdots \vee n_j}{l_1 \vee \cdots \vee l_i \vee n_1 \vee \cdots \vee n_j}$$

- 導出
  - 長さ2の場合の融合規則におけるlやnに$l_1 \vee \cdots \vee l_i$や$n_1 \vee \cdots \vee n_j$を代入

命題論理における推論パターン(§7.5) 証明の戦略その3 (融合法)

- ファクタリング
  - 同じリテラルが複数個出現した場合それらを一つに簡単化
  - 例
    - $(A \vee B)$と$(A \vee \neg B)$に融合規則を適用すれば、$A \vee A$が得られる
    - $A \vee A$をAに簡単化

命題論理における推論パターン(§7.5)
証明の戦略その3 (融合法)

- 偽の出現の意味
  - $\dfrac{m \qquad\qquad \neg m}{0}$となる場合
  - $KB \land \neg\alpha \equiv R_1 \land \cdots \land R_n \land m \land \neg m \equiv 0$となることを意味しており、全体が充足不能となる

命題論理における推論パターン(§7.5)
証明の戦略その3 (融合法)

- $KB \vDash \alpha$の判定
  - 偽$(= 0)$が出現
    - 全体$(KB \wedge \neg\alpha)$が充足不能である→背理法により$KB \vDash \alpha$となる
  - 偽$(= 0)$が出現せず、融合規則をもうこれ以上適用できない
    - 全体$(KB \wedge \neg\alpha)$を真とする何らかのモデルを割り当てることができる→$KB \vDash \alpha$が成り立たない

命題論理における推論パターン(§7.5) 証明の戦略その3 (融合法)

- 連言標準形(CNF)への変換
  1. $\alpha \Leftrightarrow \beta$を$(\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \alpha)$に置き換えることで、$\Leftrightarrow$を除去
  2. $\alpha \Rightarrow \beta$を$\neg\alpha \vee \beta$に置き換えることで$\Rightarrow$を除去
  3. 次の論理的同値関係を使うことで否定($\neg$)を内側へいれる
     - $\neg\neg\alpha \equiv \alpha$　(二重否定除去)
     - $\neg(\alpha \wedge \beta) \equiv \neg\alpha \vee \neg\beta$　(ド・モルガンの法則)
     - $\neg(\alpha \vee \beta) \equiv \neg\alpha \wedge \neg\beta$　(ド・モルガンの法則)
  4. 次の論理的同値関係を使うことで、$\vee$を$\wedge$に対して可能な限り分配する
     - $\alpha \vee (\beta \wedge \gamma) \equiv (\alpha \vee \beta) \wedge (\alpha \vee \gamma)$

命題論理における推論パターン(§7.5)
証明の戦略その3 (融合法)

- 連言標準形(CNF)への変換の例

$$B_{1,1} \Leftrightarrow \left(P_{1,2} \vee P_{2,1}\right)$$

1. ⇔の除去

$$\left(B_{1,1} \Rightarrow \left(P_{1,2} \vee P_{2,1}\right)\right) \wedge \left(\left(P_{1,2} \vee P_{2,1}\right) \Rightarrow B_{1,1}\right)$$

2. ⇒の除去

$$\left(\neg B_{1,1} \vee P_{1,2} \vee P_{2,1}\right) \wedge \left(\neg\left(P_{1,2} \vee P_{2,1}\right) \vee B_{1,1}\right)$$

3. ¬を内側にいれる

$$\left(\neg B_{1,1} \vee P_{1,2} \vee P_{2,1}\right) \wedge \left(\left(\neg P_{1,2} \wedge \neg P_{2,1}\right) \vee B_{1,1}\right)$$

4. ∨を∧に対して分配

$$\left(\neg B_{1,1} \vee P_{1,2} \vee P_{2,1}\right) \wedge \left(\neg P_{1,2} \vee B_{1,1}\right) \wedge \left(\neg P_{2,1} \vee B_{1,1}\right)$$

命題論理における推論パターン(§7.5)
証明の戦略その3 (融合法)

- 融合法の完全性
  - $RC(S)$:節集合Sに対し、融合規則を可能な限り適用した結果の節の集合
  - 融合規則の適用は必ず終了する
    - 命題記号は有限
    - ファクタリングにより一つの節の中に同じリテラルは二度出現しない
    - $\rightarrow RC(S)$は有限のサイズ

命題論理における推論パターン(§7.5) 証明の戦略その3 (融合法)

- 融合法の完全性
  - 基礎融合定理 (ground resolution theorem)

節集合Sが充足不能 $\Leftrightarrow$ $RC(S)$が偽を含む

  - 右から左は明らか
  - 左から右を証明したい
  - 対偶を証明: $RC(S)$が偽を含まない $\Rightarrow$ 節集合Sは充足可能
    - $RC(S)$が求まり、 偽が含まれなければ、全体を真とするモデル(各命題記号に対する真理値の割り当て)を必ず与える手続きが存在
  - よって、基礎融合定理が成り立つ

命題論理における推論パターン(§7.5) 融合法の例

- ワンパスワールド
- R_1:　$\neg P_{1,1}$
- R_2:　$B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$
- R_3:　$B_{2,1} \Leftrightarrow (P_{1,1} \vee P_{2,2} \vee P_{3,1})$
- R_4:　$\neg B_{1,1}$
- R_5:　$B_{2,1}$

2	1,2	2,2	3,2
1	1,1 ¬穴 ¬風	2,1 風	3,1
	1	2	3

- $KB = \{R_1, R_2, R_3, R_4, R_5\}$, $Q: \neg P_{1,2} \wedge \neg P_{2,1}$としたとき、$KB \vDash Q$が成り立つか否か？

命題論理における推論パターン(§7.5) 融合法の例

$$
\begin{aligned}
&KB \land \neg Q \\
&\equiv \neg P_{1,1} \land \left[B_{1,1} \Leftrightarrow \left(P_{1,2} \lor P_{2,1}\right)\right] \land \left[B_{2,1} \Leftrightarrow \left(P_{1,1} \lor P_{2,2} \lor P_{3,1}\right)\right] \\
&\land \neg B_{1,1} \land B_{2,1} \land \neg\left(\neg P_{1,2} \land \neg P_{2,1}\right)
\end{aligned}
$$

$B_{1,1} \Leftrightarrow \left(P_{1,2} \lor P_{2,1}\right)$に対し、$\Leftrightarrow$ の除去を適用し、

$$\left[B_{1,1} \Rightarrow \left(P_{1,2} \lor P_{2,1}\right)\right] \land \left[\left(P_{1,2} \lor P_{2,1}\right) \Rightarrow B_{1,1}\right]$$

続いて⇒を除去し、否定を内側に入れ、分配則を適用する。

$$
\begin{aligned}
&\left[\neg B_{1,1} \lor P_{1,2} \lor P_{2,1}\right] \land \left[\neg\left(P_{1,2} \lor P_{2,1}\right) \lor B_{1,1}\right] \\
&\equiv \left[\neg B_{1,1} \lor P_{1,2} \lor P_{2,1}\right] \land \left[\left(\neg P_{1,2} \land \neg P_{2,1}\right) \lor B_{1,1}\right] \\
&\equiv \left(\neg B_{1,1} \lor P_{1,2} \lor P_{2,1}\right) \land \left(\neg P_{1,2} \lor B_{1,1}\right) \land \left(\neg P_{2,1} \lor B_{1,1}\right) \quad \cdots \quad ①
\end{aligned}
$$

命題論理における推論パターン(§7.5)
融合法の例

$$KB \land \neg Q$$
$$\equiv \neg P_{1,1} \land \left[B_{1,1} \Leftrightarrow \left(P_{1,2} \lor P_{2,1}\right)\right] \land \left[B_{2,1} \Leftrightarrow \left(P_{1,1} \lor P_{2,2} \lor P_{3,1}\right)\right]$$
$$\land \neg B_{1,1} \land B_{2,1} \land \neg\left(\neg P_{1,2} \land \neg P_{2,1}\right)$$

同様に$\left[B_{2,1} \Leftrightarrow \left(P_{1,1} \lor P_{2,2} \lor P_{3,1}\right)\right]$に対し、$\Leftrightarrow$ の除去を適用し、

$$\left[B_{2,1} \Rightarrow \left(P_{1,1} \lor P_{2,2} \lor P_{3,1}\right)\right] \land \left[\left(P_{1,1} \lor P_{2,2} \lor P_{3,1}\right) \Rightarrow B_{2,1}\right]$$

続いて⇒を除去し、否定を内側に入れ、分配則を適用する。

$$\left[\neg B_{2,1} \lor P_{1,1} \lor P_{2,2} \lor P_{3,1}\right] \land \left[\neg\left(P_{1,1} \lor P_{2,2} \lor P_{3,1}\right) \lor B_{2,1}\right]$$
$$\equiv \left[\neg B_{2,1} \lor P_{1,1} \lor P_{2,2} \lor P_{3,1}\right] \land \left[\left(\neg P_{1,1} \land \neg P_{2,2} \land \neg P_{3,1}\right) \lor B_{2,1}\right]$$
$$\equiv \left(\neg B_{2,1} \lor P_{1,1} \lor P_{2,2} \lor P_{3,1}\right) \land \left(\neg P_{1,1} \lor B_{2,1}\right) \land \left(\neg P_{2,2} \lor B_{2,1}\right)$$
$$\land \left(\neg P_{3,1} \lor B_{2,1}\right) \quad \cdots \ ②$$

命題論理における推論パターン(§7.5) 融合法の例

$$
\begin{aligned}
&KB \land \neg Q \\
&\equiv \neg P_{1,1} \land \left[B_{1,1} \Leftrightarrow \left(P_{1,2} \lor P_{2,1}\right)\right] \land \left[B_{2,1} \Leftrightarrow \left(P_{1,1} \lor P_{2,2} \lor P_{3,1}\right)\right] \\
&\land \neg B_{1,1} \land B_{2,1} \land \neg\left(\neg P_{1,2} \land \neg P_{2,1}\right)
\end{aligned}
$$

クエリーの否定を中にいれ、

$$\neg\left(\neg P_{1,2} \land \neg P_{2,1}\right) \equiv P_{1,2} \lor P_{2,1} \quad \cdots \quad ③$$

命題論理における推論パターン(§7.5) 融合法の例

$$KB \wedge \neg Q$$
$$\equiv \neg P_{1,1} \wedge \left[B_{1,1} \Leftrightarrow \left(P_{1,2} \vee P_{2,1}\right)\right] \wedge \left[B_{2,1} \Leftrightarrow \left(P_{1,1} \vee P_{2,2} \vee P_{3,1}\right)\right]$$
$$\wedge \neg B_{1,1} \wedge B_{2,1} \wedge \neg\left(\neg P_{1,2} \wedge \neg P_{2,1}\right)$$

①②③より、

$KB \wedge \neg Q = \neg P_{1,1} \wedge \left(\neg B_{1,1} \vee P_{1,2} \vee P_{2,1}\right) \wedge \left(\neg P_{1,2} \vee B_{1,1}\right) \wedge \left(\neg P_{2,1} \vee B_{1,1}\right) \wedge \left(\neg B_{2,1} \vee P_{1,1} \vee P_{2,2} \vee P_{3,1}\right) \wedge \left(\neg P_{1,1} \vee B_{2,1}\right) \wedge \left(\neg P_{2,2} \vee B_{2,1}\right) \wedge \left(\neg P_{3,1} \vee B_{2,1}\right) \wedge \neg B_{1,1} \wedge B_{2,1} \wedge \left(P_{1,2} \vee P_{2,1}\right)$

ここで、$KB \wedge \neg Q$が矛盾することを証明する。

$$\neg P_{1,2} \vee B_{1,1}, \neg B_{1,1} \vdash \neg P_{1,2}$$
$$\neg P_{2,1} \vee B_{1,1}, \neg B_{1,1} \vdash \neg P_{2,1}$$
$$P_{1,2} \vee P_{2,1}, \neg P_{1,2} \vdash P_{2,1}$$
$$P_{2,1}, \neg P_{2,1} \vdash 0$$

よって矛盾することが証明された。従って背理法より、$KB \vDash Q$が成り立つ。