知識工学（第５回）

二宮 崇（ninomiya@cs.ehime-u.ac.jp）

論理的エージェント(7 章のつづき)

〇証明の戦略その 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、たしかに$KB \vDash \alpha$となることがわかるが、なかなか証明できないときや、証明が本当にできないときには、$KB \vDash \alpha$が成り立つのか成り立たないのかわからない。また、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、必ず有限時間で終了する完全な推論の方法である融合法を紹介する。

融合法はその基本戦略として背理法を用いる。つまり、$KB \wedge \neg\alpha$が充足不能であることを示すことで、$KB \vDash \alpha$を示す。

1. $KB \wedge \neg\alpha$を連言標準形(conjunctive normal form: CNF)に変換する。連言標準形は次の形をした文である。

$$(l_{1,1} \vee l_{1,2} \vee \cdots) \wedge (l_{2,1} \vee l_{2,2} \vee \cdots) \wedge \cdots \wedge (l_{n,1} \vee l_{n,2} \vee \cdots)$$

ただし、$l_{i,j}$はリテラルと呼ばれ、リテラルは命題記号または命題記号の否定のいずれかである。命題記号は正リテラル、命題記号の否定は負リテラルとも呼ばれる。$(l_{i,1} \vee l_{i,2} \vee \cdots)$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。

2. 節集合KB'（知識ベース+$\neg\alpha$）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、$KB \nvDash \alpha$が証明されたことになる。

4. $KB' \vdash 0$となれば、$KB \vDash \alpha$が証明されたことになる。

5. 2.に戻る

・融合規則（もっとも簡単な形，選言的三段論法)

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

これについては、$(l \vee m) \wedge \neg m$が推論の前提部になり、$\neg m$であるから、mは常に偽となることがわかり、$l \vee m$はlと等しくなることがわかる。機械的には、分配則より、$(l \vee m) \wedge \neg m \equiv (l \wedge \neg m) \vee 0 \equiv l \wedge \neg m$であり、縮小律(And 除去)により、$l \wedge \neg m \vDash l$が成り立つ。従って、$(l \vee m) \wedge \neg m \vDash l$となる。または真理値表で確認しても成り立つことがわかる。

・融合規則(長さ 2 の節の場合)

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

これについては分配則を 2 回適用し､縮小律(And 除去)を適用すれば､結論部が得られる｡もしくは、真理値表を書けばこの推論が成り立つことがわかる。

・融合規則(一般)

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

これについては長さ 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$n_1 \vee \cdots \vee n_j$を代入すれば成り立つことがわかる。

融合規則においては、二つの節が選言で融合することになるがその場合に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できる(簡単化しなければいけない)。例えば、$(A \vee B)$と$(A \vee \neg B)$を融合すれば、$A \vee A$となるが、これはAに簡単化される。この操作をファクタリング(factoring)という。

融合規則をKB'に対し適用し、偽$(= 0)$が出現すれば、全体$(KB \wedge \neg\alpha)$が充足不能である、ということが証明できたことになり、背理法により$KB \vDash \alpha$となることがわかる。偽が出現せず、融合規則をもう適用できなくなれば、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、その場合は$KB \vDash \alpha$が成り立たないことがわかる。

偽の出現については、例えば、

$$\frac{m \qquad \neg m}{0}$$

となる場合であり、これは$KB \wedge \neg\alpha \equiv R_1 \wedge \cdots \wedge R_n \wedge m \wedge \neg m \equiv 0$となることを意味しており、全体が充足不能となる。

・連言標準形(CNF)への変換

次に、最初のステップである CNF への変換について説明する。任意の論理式は CNF に変換可能である。

次の CNF への変換手続きは次のようになる。

1. $\alpha \Leftrightarrow \beta$を$(\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \alpha)$に置き換えることで、⇔を除去する
2. $\alpha \Rightarrow \beta$を$\neg\alpha \vee \beta$に置き換えることで⇒を除去する
3. 次の論理的同値関係を使うことで否定(¬)を内側へいれていく

   $\neg\neg\alpha \equiv \alpha$　　(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、∨を∧に対して可能な限り分配する

$$\alpha \vee (\beta \wedge \gamma) \equiv (\alpha \vee \beta) \wedge (\alpha \vee \gamma)$$

この手続きによって CNF となる。

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

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

・融合法の完全性

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を$RC(S)$とする。命題記号は有限であり、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ため、$RC(S)$は有限のサイズとなる。従って、融合規則の適用は必ず終了する。

基礎融合定理 (ground resolution theorem)

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

右から左は明らかに成り立つので、

$$\text{節集合}S\text{が充足不能} \Rightarrow RC(S)\text{が偽を含む}$$

を証明すれば良い。これは、この対偶を証明すれば良い。

$$RC(S)\text{が偽を含まない} \Rightarrow \text{節集合}S\text{は充足可能}$$

$RC(S)$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し、そのため、上の基礎融合定理が成り立つ。

実際にワンパスワールドの例を融合法で解いてみる。

〇ワンパスワールドの例

R_1:　$\neg P_{1,1}$

R_2:　$B_{1,1} \Leftrightarrow \left(P_{1,2} \vee P_{2,1}\right)$

R_3:　$B_{2,1} \Leftrightarrow \left(P_{1,1} \vee P_{2,2} \vee P_{3,1}\right)$

R_4:　$\neg B_{1,1}$

R_5:　$B_{2,1}$

$KB = \{\mathrm{R}_1, \mathrm{R}_2, \mathrm{R}_3, \mathrm{R}_4, \mathrm{R}_5\}$, $Q: \neg P_{1,2} \wedge \neg P_{2,1}$としたとき、$KB \vDash Q$が成り立つか否か？

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

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

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

続いて⇒を除去し、否定を内側に入れ、分配則を適用する。

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

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

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

続いて⇒を除去し、否定を内側に入れ、分配則を適用する。

$$\begin{aligned} & \left[\neg B_{2,1} \vee P_{1,1} \vee P_{2,2} \vee P_{3,1}\right] \wedge \left[\neg\left(P_{1,1} \vee P_{2,2} \vee P_{3,1}\right) \vee B_{2,1}\right] \\ &\quad \equiv \left[\neg B_{2,1} \vee P_{1,1} \vee P_{2,2} \vee P_{3,1}\right] \wedge \left[\left(\neg P_{1,1} \wedge \neg P_{2,2} \wedge \neg P_{3,1}\right) \vee B_{2,1}\right] \\ &\quad \equiv \left(\neg B_{2,1} \vee P_{1,1} \vee P_{2,2} \vee P_{3,1}\right) \wedge \left(\neg P_{1,1} \vee B_{2,1}\right) \wedge \left(\neg P_{2,2} \vee B_{2,1}\right) \wedge \left(\neg P_{3,1} \vee B_{2,1}\right) \quad \cdots \ ② \end{aligned}$$

クエリーの否定を中にいれ、

$$\neg\left(\neg P_{1,2} \wedge \neg P_{2,1}\right) = P_{1,2} \vee P_{2,1} \quad \cdots \ ③$$

①②③より、

$$\begin{aligned} KB \wedge \neg Q \equiv\ & \neg P_{1,1} \wedge \left(\neg B_{1,1} \vee P_{1,2} \vee P_{2,1}\right) \wedge \left(\neg P_{1,2} \vee B_{1,1}\right) \wedge \left(\neg P_{2,1} \vee B_{1,1}\right) \wedge \left(\neg B_{2,1} \vee P_{1,1} \vee P_{2,2} \vee P_{3,1}\right) \\ & \wedge \left(\neg P_{1,1} \vee B_{2,1}\right) \wedge \left(\neg P_{2,2} \vee B_{2,1}\right) \wedge \left(\neg P_{3,1} \vee B_{2,1}\right) \wedge \neg B_{1,1} \wedge B_{2,1} \wedge \left(P_{1,2} \vee P_{2,1}\right) \end{aligned}$$

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。

付録: 融合法の完全性の証明の詳細

$$RC(S)\text{が偽を含まない} \Rightarrow \text{節集合}S\text{は充足可能}$$

このことを証明する。$RC(S)$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する。その手続きは次のようになる。

Sに出現する命題記号を$P_1, \cdots, P_k$とする。1からkまでiを動かしながら、

- $\neg P_i$を含み、かつそれまでに選ばれた$P_1, \cdots, P_{i-1}$への真理値割り当てのもとで他のリテラルがすべて偽となるような節が$RC(S)$に存在する場合、偽をP_iに割り当てる。
- そのような節が存在しなければ、P_iに真を割り当てる。

この手続きにより、全体を真とするモデルが与えられるため、基礎融合定理が成り立つ。

最後に、この手続きで与えられる真理値割り当てがSのモデルとなることを数学的帰納法で証明する。$P_1, \cdots, P_{i-1}$まで真理値割り当てが終わっていて、その割り当てにより真偽が決定される節には偽となる節は存在しないと仮定する。今、P_iの真理値を決めようとしているとする。このとき、ある節Cが偽になる場合は、Cが$(0 \vee 0 \vee \cdots \vee 0 \vee P_i)$となるか、$(0 \vee 0 \vee \cdots \vee 0 \vee \neg P_i)$となるかのどちらかであるが、$RC(S)$にこの一方しか含まれない場合は、$C$が真になるよう$P_i$の値を決めれば良い。問題は$RC(S)$にこの両方が含まれる場合であるが、融合法により融合規則を可能な限り適用した後であるため、これらの二つを融合した$(0 \vee 0 \vee \cdots \vee 0) = 0$となる節が存在していることになる。しかし、これは仮定に反する。よって、$RC(S)$に上記の二つの節が同時に含まれることはない。従って、数学的帰納法により、この真理値割り当ての手順でS全体を真とするモデルを作ることができる。 (証明終)

第４回と第５回のまとめ

〇証明の戦略その１(普通の証明)

1. $KB \vDash R$となるRをみつけRを知識ベースに追加する(知識ベースは更新される)。この操作(推論)を$KB \vdash R$と書くことにする。

2. $KB \vdash \alpha$となれば、$KB \vDash \alpha$が証明されたことになる。

3. 1.に戻る。

〇証明の戦略その２(背理法による証明)

1. $KB' = KB \wedge \neg\alpha$とする。

2. $KB' \vDash R$となるRをみつけRを知識ベースKB'に追加する(知識ベース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. $KB' \vdash 0$となれば、$KB \vDash \alpha$が証明されたことになる。

4. 2.に戻る。

〇推論規則 ($KB \vdash R$)

推論規則$KB \vdash R$には、伴意関係$KB \vDash R$や論理的同値関係$KB \equiv R$を用いることができる。

○証明の戦略その 3 (融合法)

1. $KB \land \neg\alpha$を連言標準形(CNF)に変換し、その節集合をKB'とする。連言標準形は次の形をした論理式である。

$$(l_{1,1} \lor l_{1,2} \lor \cdots) \land (l_{2,1} \lor l_{2,2} \lor \cdots) \land \cdots \land (l_{n,1} \lor l_{n,2} \lor \cdots)$$

ただし、$l_{i,j}$は正リテラル(命題記号)か負リテラル(否定のついた命題記号)である。$(l_{i,1} \lor l_{i,2} \lor \cdots)$は節と呼ばれる。

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、$KB \nvDash \alpha$が証明されたことになる。

4. $KB' \vdash 0$となれば、$KB \vDash \alpha$が証明されたことになる。

5. 2.に戻る

○融合法での推論規則 ($KB \vdash R$)

融合法で用いる推論規則は融合規則ただ一つのみ

$$\frac{l_1 \lor \cdots \lor l_i \lor m \qquad \neg m \lor n_1 \lor \cdots \lor n_j}{l_1 \lor \cdots \lor l_i \lor n_1 \lor \cdots \lor n_j} \quad \text{(融合規則)}$$

○連言標準形(CNF)への変換

1. $\alpha \Leftrightarrow \beta$を$(\alpha \Rightarrow \beta) \land (\beta \Rightarrow \alpha)$に置き換えることで、$\Leftrightarrow$を除去する

2. $\alpha \Rightarrow \beta$を$\neg\alpha \lor \beta$に置き換えることで$\Rightarrow$を除去する

3. 次の論理的同値関係を使うことで否定($\neg$)を内側へいれていく

$\neg\neg\alpha \equiv \alpha$　(二重否定除去)

$\neg(\alpha \land \beta) \equiv \neg\alpha \lor \neg\beta$　(ド・モルガンの法則)

$\neg(\alpha \lor \beta) \equiv \neg\alpha \land \neg\beta$　(ド・モルガンの法則)

4. 次の論理的同値関係を使うことで、$\lor$を$\land$に対して可能な限り分配する

$$\alpha \lor (\beta \land \gamma) \equiv (\alpha \lor \beta) \land (\alpha \lor \gamma)$$