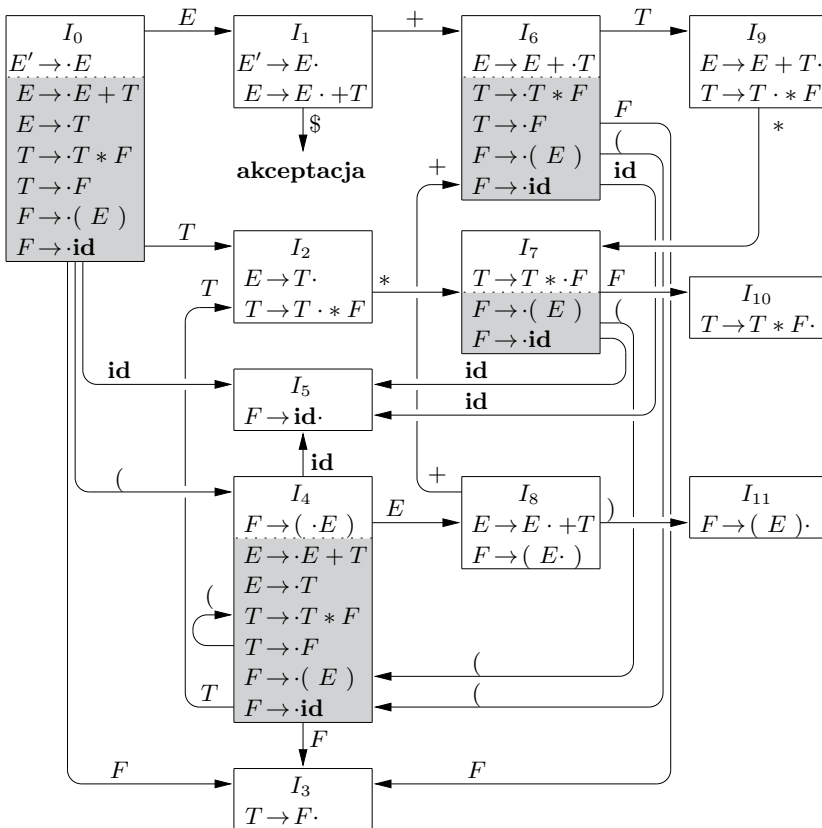


Reprezentowanie zbiorów sytuacji

Generator tworzący parser wstępujący może wymagać wygodnej reprezentacji sytuacji i ich zbiorów. Zauważmy, że sytuacja może być przedstawiona jako para liczb całkowitych, z których pierwsza jest numerem jednej z produkcji przetwarzanej gramatyki, druga zaś jest pozycją kropki w ciele tej produkcji. Zbiory sytuacji można przedstawiać jako listy takich par. Jednak, jak zobaczymy, potrzebne do zaprojektowania parsera zbiory sytuacji często zawierają sytuacje „domykające”, w których kropka znajduje się na początku ciała. Sytuacje takie zawsze można zrekonstruować z innych sytuacji w zbiorze i nie musimy ich dołączać do listy.



RYSUNEK 4.31: Automat LR(0) dla gramatyki wyrażeń (4.1)

Intuicyjnie fakt, że $A \rightarrow \alpha \cdot B\beta$ należy do $\text{CLOSURE}(I)$, sygnalizuje, że w tym punkcie procesu analizy, w którym się znajdujemy, możemy zobaczyć na wejściu jako następny podciąg, który można wyprowadzić z $B\beta$. Podciąg wyprowa-

dzalny z $B\beta$ będzie miał prefiks wyprowadzalny z B przez zastosowanie jednej z B -produkcji. Z tego względu dodajemy wszystkie sytuacje dla wszystkich B -produkcji; innymi słowy, jeśli istnieje produkcja $B \rightarrow \gamma$, dołączamy $B \rightarrow \cdot \gamma$ do $\text{CLOSURE}(I)$.

Przykład 4.40: Rozważmy uzupełnioną gramatykę wyrażeń:

$$\begin{aligned} E' &\rightarrow E \\ E &\rightarrow E + T \mid T \\ T &\rightarrow T * F \mid F \\ E &\rightarrow (E) \mid \text{id} \end{aligned}$$

Jeśli I jest zbiorem zawierającym jedną sytuację $\{[E' \rightarrow \cdot E]\}$, wówczas $\text{CLOSURE}(I)$ będzie zawierać zbiór sytuacji I_0 pokazany na rysunku 4.31.

Zobaczmy, jak wyliczane jest domknięcie. $E' \rightarrow \cdot E$ jest umieszczane w $\text{CLOSURE}(I)$ zgodnie z regułą (1). Ponieważ mamy tu E bezpośrednio na prawo od kropki, dodajemy E -produkcyjne z kropkami na lewym końcu: $E \rightarrow \cdot E + T$ oraz $E \rightarrow \cdot T$. Teraz mamy T bezpośrednio po kropce w drugiej sytuacji, zatem dodajemy $T \rightarrow \cdot T * F$ oraz $T \rightarrow \cdot F$. Następnie F na prawo od kropki skłania nas do dodania sytuacji $F \rightarrow \cdot (E)$ oraz $F \rightarrow \cdot \text{id}$, ale żadnych dalszych sytuacji nie trzeba już dodawać. ■

Domknięcie można wyliczyć w sposób pokazany na rysunku 4.32. Wygodnym sposobem implementacji funkcji *closure* jest utrzymywanie logicznej tablicy *added*, indeksowanej nieterminalami z gramatyki G , takiej, że *added*[B] jest ustawiane jako **true** wtedy, gdy dodajemy sytuację $B \rightarrow \cdot \gamma$ dla każdej B -produkcji $B \rightarrow \gamma$.

```

SetOfItems CLOSURE(I) {
    J = I;
    repeat
        for ( dla każdej sytuacji  $A \rightarrow \alpha \cdot B\beta$  należącej do J )
            for ( dla każdej produkcji  $B \rightarrow \gamma$  z G )
                if (  $B \rightarrow \cdot \gamma$  nie należy do J )
                    dodaj  $B \rightarrow \cdot \gamma$  to J;
    until nie ma więcej elementów do dodania do J w iteracji;
    return J;
}

```

RYSUNEK 4.32: Obliczanie funkcji CLOSURE

Zauważmy, że jeśli pewna B -produkcja zostanie dodana do domknięcia zbioru I z kropką po lewej stronie, wówczas dodane zostaną wszystkie B -produkcyjne w analogiczny sposób. Tym samym w niektórych sytuacjach nie ma potrzeby rzeczywistego wyznaczania sytuacji $B \rightarrow \cdot \gamma$ dodawanych do I przez CLOSURE . Wystarczająca będzie lista nieterminali B , których produkcje zostały tak dodane. Możemy podzielić wszystkie interesujące nas zbiory sytuacji na dwie klasy:

1. Sytuacje bazowe: początkowa sytuacja $S' \rightarrow \cdot S$ oraz wszystkie te sytuacje, w których kropki nie znajdują się na lewym końcu ciała produkcji.