

2. WYKŁAD 2: INDUKCJA

2.1. Definiowanie kategorii syntaktycznych przez indukcję. Istotnym składnikiem definicji języka programowania jest jego **składnia syntaktyczna**, dzięki której jesteśmy w stanie odpowiedzieć jaki program (rozumiany jako ciąg znaków) będzie rozpoznany przez parser, a jaki nie będzie. Na ogół składnia syntaktyczna jest określona za pomocą pewnej liczby **kategorii syntaktycznych** takich jak wyrażenia, typy i wzorce. Poniżej przedstawimy na kilku przykładach w jaki sposób można definiować kategorie syntaktyczne przez indukcję w kilku prostych językach.

Jako pierwszy przykład rozważmy następującą definicję kategorii syntaktycznej **nat** liczb naturalnych:

$$\text{nat} \quad n ::= 0 \mid S \ n$$

W przykładzie tym **nat** jest nazwą kategorii syntaktycznej, którą definiujemy, n jest zwany **nie-terminalem**, $::=$ czytamy “*jest definiowane jako*”, a I jako “*lub*”, 0 oznacza “*zero*”, zaś S oznacza “*następnik*”. Tym samym definicję można odczytać następująco:

liczba naturalna n jest definiowana jako 0 albo $S \ n'$, gdzie n' jest inną liczbą naturalną.

Zauważmy, że **nat** jest tu definiowane indukcyjnie: liczba naturalna $S \ n'$ jest zdefiniowana za pomocą innej liczby naturalnej n' i tym samym **nat** używa tej samej kategorii syntaktycznej w swej definicji. A zatem definicja **nat** produkuje nieskończony zbiór liczb naturalnych takich jak

$$0, S \ 0, S \ S \ 0, S \ S \ S \ 0, \dots$$

Kategoria syntaktyczna może odnosić się w swej definicji do innej kategorii syntaktycznej. Na przykład, mając daną zdefiniowaną powyżej kategorię **nat**, możemy indukcyjnie zdefiniować nową kategorię syntaktyczną **tree** jak następuje:

$$\text{tree} \quad t ::= \text{leaf } n \mid \text{node}(t, n, t)$$

Tu z kolei **leaf** n oznacza liść drzewa o numerze n , **node** (t_1, n, t_2) oznacza węzeł o numerze n z lewym potomkiem t_1 i prawym potomkiem t_2 . Zatem **tree** specyfikuje język drzewek binarnych liczb naturalnych takich jak

$$\text{leaf } n, \text{node}(\text{leaf } n_1, n, \text{leaf } n_2), \text{node}(\text{node}(\text{leaf } n_1, n, \text{leaf } n_2), n', n''), \dots$$

W podobny sposób możemy pokusić się o zdefiniowanie dwóch kategorii syntaktycznych które wzajemnie definiują się przez indukcję. Na przykład możemy równocześnie zdefiniować kategorie **par** i **npar** liczb parzystych i nieparzystych jak następuje

$$\begin{array}{ll} \text{par} & e ::= 0 \mid S \ o \\ \text{npar} & o ::= S \ e \end{array}$$

Tym samym **par** składa się z liczb parzystych takich jak

$$0, S \ S \ 0, S \ S \ S \ S \ 0, \dots$$

a **npar** z liczb nieparzystych takich jak

$$S \ 0, S \ S \ S \ S \ 0, S \ S \ S \ S \ S \ 0, \dots$$

Zauważmy też, że **par** oraz **npar** są **podkategoriami** kategorii **nat** ponieważ każda liczba parzysta e i każda liczba nieparzysta o są również liczbami naturalnymi. Tym samym możemy myśleć o **par** lub **npar** jako o **nat** spełniających pewne dodatkowe warunki.

Zadanie 1. Zdefiniować **par** oraz **npar** niezależnie od siebie.

Przyjrzyjmy się jeszcze jednemu przykładowi definiowania podkategorii syntaktycznej. Zdefiniujmy najpierw syntaktyczną kategorię **paren** ciągów nawiasów:

$$\text{paren} \quad s ::= \epsilon \mid (s) \mid s$$

ϵ będzie oznaczać pusty ciąg (a więc dla dowolnego ciągu s zachodzi $\epsilon s = s = s\epsilon$). Zauważmy, że **paren** definiuje język ciągów nawiasów zapisywanych bez żadnych ograniczeń. Teraz zdefiniujemy podkategorię **mparen** kategorii **paren** ciągów nawiasów, w których nawiasy “lewe” odpowiadają nawiasom “prawym”:

$$\text{mparen} \quad s ::= \epsilon \mid (s) \mid s s$$

mparen generuje ciągi takie jak

$$\epsilon, (), (()), (()), (())(), ()(), \dots$$

Zauważmy, że w myśl przyjętej definicji **mparen**, ciąg nawiasów z tej kategorii może nie dać się rozłożyć w jednoznaczny sposób: na przykład ciąg $()()$ może być równie dobrze otrzymany z dopisania do ciągu $()$ ciągu $()()$, jak i z dopisania do ciągu $()()$ ciągu $()$. Nietrudno zauważyć, że dwuznaczność ta jest spowodowana przez trzecią część definicji: dla danego ciągu podciągów nawiasów może istnieć więcej niż jeden sposób na rozdzielenie go na dwa podciągi nawiasów. Możemy wyeliminować tę dwuznaczność definiując kategorię **lparen**:

$$\text{lparen} \quad s ::= \epsilon \mid (s) s$$

Pomysł na definicję **lparen** jest następujący: pierwszy (licząc od lewej strony) nawias pojawiający się w niepustym ciągu nawiasów s jest lewy nawias “(”, któremu odpowiada dokładnie jedno wystąpienie prawego nawiasu “)”. Na przykład ciąg $s = (())()$ może być zapisany jako $(s_1)s_2$, gdzie $s_1 = ()$ oraz $s_2 = ()$ są ciągami odpowiadającymi sobie lewych i prawych nawiasów, każdy jednoznacznie wyznaczony przez s . Z drugiej strony ciągi $()$ oraz $(())$ nie są ciągami odpowiadającymi sobie lewych i prawych nawiasów i nie mogą być zapisane w postaci $(s_1)s_2$, gdzie s_1 i s_2 są ciągami odpowiadającymi nawiasów.

Definicja indukcyjna jest wygodnym sposobem specyfikowania języka. Nawet składnie syntaktyczne języków “na serio” (na przykład Standard ML) używają de facto takiej samej maszyny. Okazuje się jednak, że indukcyjnie definiowane kategorie syntaktyczne nie są najlepszym narzędziem do badania właściwości języka. Na przykład, w jaki w zasadzie sposób chcielibyśmy zapisać, że n należy do **nat** jeśli n należy do **nat** – a tym bardziej w jaki sposób chcielibyśmy coś takiego udowodnić? Podobnie, w jaki sposób chcielibyśmy sprawdzić, że dany ciąg nawiasów należy do **mparen**? Celem udzielenia odpowiedzi na takie pytania musimy najpierw poznać formalną definicję **osądu**.

2.2. Definiowanie osądów przez indukcję. Jako **osąd** będziemy rozumieli pewien fragment wiedzy, który może dać się dowieść, lub może nie dać się dowieść. Na przykład:

- “ $1 - 1$ jest równe 0” jest osądem, który zawsze daje się dowieść,
- “ 1 jest równe 0” jest osądem, którego nigdy nie daje się dowieść,
- “dzisiaj pada deszcz” jest osądem, który daje się dowieść w zależności od kontekstu :-)
- “**S S 0** należy do kategorii syntaktycznej **nat**” jest osądem, który daje się dowieść o ile **nat** jest kategorią zdefiniowaną tak, jak przed chwilą.

Jak wobec tego dowodzimy dany osąd? Na przykład, na jakiej podstawie twierdzimy, że “ $1 - 1$ jest równe 0” jest osądem, który zawsze daje się dowieść? Implicite chcielibyśmy użyć prostych reguł arytmetycznych, aby udowodnić, iż “ $1 - 1$ jest równe 0”, ale, chcąc być precyzyjnym, musimy wszak odnotować, że reguły arytmetyczne nie są nam dane – musimy je wpiąć sformułować jako **reguły wnioskowania**.

Każda **reguła wnioskowania** składa się z osądów $J_1, \dots, J_n$ zwanych **przesłankami** i z osądu J zwanego **wnioskiem** i zapisywana jest w postaci

$$\frac{J_1 \ J_2 \ \dots \ J_n}{J} R$$

Regułę tę, którą nazwaliśmy tu R , odczytujemy w ten sposób, że jeśli zachodzą przesłanki $J_1, \dots, J_n$ to zachodzi również wniosek J . Jako szczególny przypadek reguła wnioskowania pozbawiona przesłanej będzie nazywana **aksjomatem**. Oto kilka przykładów reguł wnioskowania – aby nie zaciemniać obrazu, zrezygnowaliśmy z wprowadzania nazw poniższych reguł:

$$\frac{m \text{ jest równe } l \quad l \text{ jest równe } n}{m \text{ jest równe } n} \quad \frac{m \text{ jest równe } l}{m+1 \text{ jest równe } l+1} \quad \frac{m \text{ jest równe } m}{1 \text{ jest liczbą naturalną}} \quad \frac{m \text{ jest równe } l}{m+1 \text{ jest równe } l+1} \quad \frac{m \text{ jest równe } l}{m+1 \text{ jest równe } l+1} \quad \frac{m \text{ jest równe } l}{m+1 \text{ jest równe } l+1}$$

Koncepcja osądu jest na tyle ogólna, że w zasadzie stosuje się do dowolnego rodzaju wiedzy: do wiedzy o pogodzie, wiedzy o liczbach, wiedzy o programowaniu etc. Zauważmy wszelako, że same osądy nie są wystarczającym środkiem do uzasadniania fragmentów badanej wiedzy – równie ważne są tu reguły wnioskowania służące do dowodzenia lub obalania osądów. Innymi słowy, aby definicja osądu jest kompletna wtedy i tylko wtedy, gdy istnieją reguły wnioskowania pozwalające na jego udowodnienie bądź obalenie. Bez reguł wnioskowania nie sposób nadać osądowi znaczenie. Na przykład, bez reguł wnioskowania w arytmetyce, stwierdzenie “ $1 - 1$ jest równe 0 ” nie jest niczym więcej jak ciągiem znaków bez znaczenia i nie może być nazwane osądem.

Jak się domyślamy, osądy są narzędziem wystarczająco silnym aby rozstrzygać o przynależności do kategorii syntaktycznej. Jako przykład spójrzmy na indukcyjną definicję kategorii **nat** jako na zbiór pewnych osądów i reguł wnioskowania. Na początek wprowadzamy osąd $n \text{ nat}$:

$$n \text{ nat} \quad \Leftrightarrow \quad n \text{ jest liczbą naturalną}$$

oraz następujące dwie reguły wnioskowania, nazwane tu jako *Zero* oraz *Succ*:

$$\frac{}{n \text{ nat}} \text{ Zero} \quad \frac{n \text{ nat}}{S \ n \text{ nat}} \text{ Succ}$$

Występująca w regule *Succ* zmienna n jest zwana **metazmienną** i oznacza pewien ciąg złożony z elementów S oraz 0 i tym samym nie jest częścią języka składającego się z S i 0 . Innymi słowy, n jest tu metazmienną przyjmującą wartości ze zbioru ciągów złożonych z S i 0 i (przynajmniej przed zastąpieniem n przez pewien konkretny ciąg, na przykład $S \ 0$) nie jest testowana na przynależność do **nat**.

Tym samym osąd $n \text{ nat}$ jest teraz zdefiniowany indukcyjnie przez dwie reguły wnioskowania. Reguła *Zero* jest aksjomatem i gra tu rolę przypadku bazowego, zaś reguła *Succ* jest przypadkiem indukcyjnym, albowiem jej przesłanka zawiera osąd o mniejszym rozmiarze niż osąd (tego samego rodzaju) występujący we wniosku. Teraz możemy udowodnić, na przykład, że osąd $S \ S \ 0 \text{ nat}$ ma dowód: robimy to wskazując następujące **drzewo wnioskowania**, w którym $S \ S \ 0 \text{ nat}$ jest korzeniem, a 0 nat jedynym liściem:

$$\frac{\frac{\frac{}{0 \text{ nat}} \text{ Zero}}{S \ 0 \text{ nat}} \text{ Succ}}{S \ S \ 0 \text{ nat}} \text{ Succ}$$

W podobny sposób możemy zapisać definicję kategorii syntaktycznej **tree** przy użyciu osądów i reguł wnioskowania:

$$t \text{ tree} \quad \Leftrightarrow \quad t \text{ jest drzewem binarnym liczb naturalnych}$$

$$\frac{n \text{ nat}}{\text{leaf } n \text{ tree}} \text{ Leaf} \quad \frac{t_1 \text{ tree} \quad n \text{ nat} \quad t_2 \text{ tree}}{\text{node } (t_1, n, t_2) \text{ tree}} \text{ Node}$$

W troszkę bardziej skomplikowany, ale w sumie też podobny sposób możemy zapisać definicję kategorii syntaktycznej $\text{ctree}\langle d \rangle$ pełnych drzew o danej wysokości:

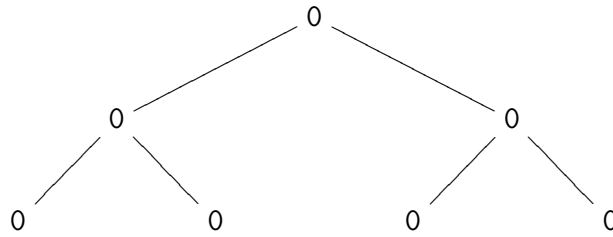
$$t \text{ ctree}\langle d \rangle \quad \Leftrightarrow \quad t \text{ jest pełnym drzewem binarnym liczb naturalnych o wysokości } d$$

$$\frac{n \text{ nat}}{\text{leaf } n \text{ ctree}\langle 0 \rangle} \text{ Cleaf} \quad \frac{t_1 \text{ ctree}\langle d \rangle \quad n \text{ nat} \quad t_2 \text{ ctree}\langle d \rangle}{\text{node } (t_1, n, t_2) \text{ ctree}\langle S \ d \rangle} \text{ Cnode}$$

Rozważmy następujące drzewo:

$$\text{node } (\text{node } (\text{leaf } 0, 0, \text{leaf } 0), 0, \text{node } (\text{leaf } 0, 0, \text{leaf } 0))$$

i zobaczymy, że jest do istotnie pełne drzewko binarne liczb naturalnych o wysokości $S \ S \ 0$. Oczywiście widać to na rysunku:



ale spróbujmy jednak zapisać formalny dowód:

$$\frac{\frac{\overline{0 \text{ nat}} \text{ Zero}}{\text{leaf } 0 \text{ ctree}\langle 0 \rangle} \text{ Cleaf} \quad \frac{\overline{0 \text{ nat}} \text{ Zero}}{\text{leaf } 0 \text{ ctree}\langle 0 \rangle} \text{ Cleaf}}{\text{node } (\text{leaf } 0, 0, \text{leaf } 0) \text{ ctree}\langle S \ 0 \rangle} \text{ Cnode} \quad \frac{\overline{0 \text{ nat}} \text{ Zero}}{\text{leaf } 0 \text{ ctree}\langle 0 \rangle} \text{ Cleaf} \quad \dots \quad \frac{\overline{0 \text{ nat}} \text{ Zero}}{\text{leaf } 0 \text{ ctree}\langle 0 \rangle} \text{ Cleaf}}{\text{node } (\text{node } (\text{leaf } 0, 0, \text{leaf } 0), 0, \text{node } (\text{leaf } 0, 0, \text{leaf } 0)) \text{ ctree}\langle S \ S \ 0 \rangle} \text{ Cnode} \quad \dots$$

...gdzie w miejsce (...) trzeba wstawić taki sam osąd jak pierwsza z przesłanek w ostatnim wnioskowaniu. Podobnie możemy sprawdzić, że

$$t = \text{node } (\text{leaf } 0, 0, \text{node } (\text{leaf } 0, 0, \text{leaf } 0))$$

nie jest pełnym drzewem binarnym, albowiem nie potrafimy udowodnić, iż $t \text{ ctree}\langle d \rangle$ dla żadnej liczby naturalnej d :

$$\frac{\frac{\overline{0 \text{ nat}} \text{ Zero}}{\text{leaf } 0 \text{ ctree}\langle d' \rangle} \text{ Cleaf} \quad \frac{d' = 0}{\text{leaf } 0 \text{ ctree}\langle d' \rangle} \text{ Cleaf}}{\text{node } (\text{leaf } 0, 0, \text{leaf } 0) \text{ ctree}\langle S \ d' \rangle} \text{ Cnode} \quad \frac{\overline{0 \text{ nat}} \text{ Zero}}{\text{leaf } 0 \text{ ctree}\langle d' \rangle} \text{ Cleaf} \quad \dots \quad \frac{d' = S \ d''}{\text{node } (\text{leaf } 0, 0, \text{leaf } 0) \text{ ctree}\langle d' \rangle} \text{ Cnode}}{\text{node } (\text{leaf } 0, 0, \text{node } (\text{leaf } 0, 0, \text{leaf } 0)) \text{ ctree}\langle S \ d' \rangle} \text{ Cnode}$$

Łatwo zauważyć, dlaczego dowód nie działa: lewe poddrzewo t wymaga, aby $d' = 0$, tymczasem prawe poddrzewo t wymaga, aby $d' = S d''$, co oczywiście jest niemożliwe.

Dokładnie tak jak to miało miejsce w przypadku kategorii syntaktycznych **par** i **npar**, różne osądy mogą być definiowane równocześnie. Na przykład, definicje **par** i **npar** można przedstawić w postaci osądów i reguł wnioskowania w następujący sposób:

$$\begin{array}{lll} n \text{ par} & \Leftrightarrow & n \text{ jest liczbą parzystą} \\ n \text{ npar} & \Leftrightarrow & n \text{ jest liczbą nieparzystą} \\ \frac{}{0 \text{ par}} \text{ ZeroE} & \frac{n \text{ npar}}{S \ n \text{ par}} \text{ SuccE} & \frac{n \text{ par}}{S \ n \text{ npar}} \text{ SuccO} \end{array}$$

W szczególności dowód tego, że $S S 0$ jest liczbą parzystą wyglądałby następująco:

$$\frac{\frac{0 \text{ par}}{S 0 \text{ npar}} \text{ZeroE}}{S S 0 \text{ par}} \text{SuccE}$$

Zadanie 2. Zapisać definicje paren , mparen i lparen przy użyciu osądów i reguł wnioskowania.

2.3. Reguły wyprowadzalne i dopuszczalne. Tak jak przed chwilą zobaczyliśmy, osady są definiowane wewnątrz pewnego ustalonego systemu reguł wnioskowania. Zauważmy, że dzięki istniejącym regułom wnioskowania możemy tworzyć nowe reguły wnioskowania, które można dodać do systemu. Tak otrzymane nowe reguły nie zmieniają w żaden sposób charakterystyki systemu, jako że wszystkie mogą zostać uzasadnione przy użyciu oryginalnych reguł, ale często mogą pomóc nam w rozwiązywaniu problemów. Na przykład mnożąc dwie liczby naturalne rzadko kiedy używamy podstawowych reguł arytmetyki, o których moglibyśmy myśleć jak o podstawowym zestawie reguł wnioskowania – zamiast tego powszechnie stosujemy tabliczkę mnożenia, o której możemy teraz myśleć jak o zestawie nowych reguł wnioskowania.

Nowe reguły wnioskowania mogą być wprowadzane na dwa sposoby: jako **reguły wyprowadzalne** lub jako **reguły dopuszczalne**. Regułą wyprowadzalną nazywamy taką regułę, w której luka pomiędzy przesłankami a wnioskiem może być wypełniona przez wklejenie określonego drzewa wnioskowania. Innymi słowy, zawsze jesteśmy w stanie wskazać ciąg reguł wnioskowania startujący od wyjściowych przesłanek, a kończący na ostatecznym wniosku. Jako przykład rozważmy następującą regułę orzekającą, że jeśli n jest liczbą naturalną, to jest nią również $S S n$:

$$\frac{n \text{ nat}}{S S n \text{ nat}} \text{Succ2}$$

Jest to reguła wyprowadzalna, albowiem możemy ją uzasadnić przy użyciu następującego drzewa dowodowego:

$$\frac{\frac{n \text{ nat}}{S n \text{ nat}} \text{Succ}}{S S n \text{ nat}} \text{Succ}$$

Możemy teraz używać reguły Succ2 tak, jakby to była oryginalna reguła wnioskowania – w razie czego zawsze możemy podać jej dowód.

Reguła dopuszczalna to taka, w której przesłanka pociąga wniosek. Innymi słowy, jeżeli spełniona jest przesłanka, to spełniony jest też wniosek. Oczywiście każda reguła wyprowadzalna jest dopuszczalna dzięki możliwości formalnego udowodnienia wniosku na podstawie przesłanki. Istnieją wszelako reguły dopuszczalne, które nie są wyprowadzalne. Rozważmy na przykład następującą regułę wnioskowania orzekającą, że jeśli $S n$ jest liczbą naturalną, to jest nią też n :

$$\frac{S n \text{ nat}}{n \text{ nat}} \text{Succ}^{-1}$$

Zauważmy po pierwsze, że nie jest to reguła wyprowadzalna: jedyny sposób, aby otrzymać $n \text{ nat}$ z $S n \text{ nat}$ to zastosowanie reguły Succ , ale przesłanka reguły Succ jest mniejsza niż jej wniosek, podczas gdy $S n \text{ nat}$ jest większe niż $n \text{ nat}$.

Przypuśćmy wszakże, że dysponujemy dowodem przesłanki $S n \text{ nat}$. Jako że jedynym sposobem udowodnienia $S n \text{ nat}$ jest zastosowanie reguły Succ , dowód $S n \text{ nat}$ musi wyglądać jakoś tak:

$$\frac{\vdots}{S n \text{ nat}} \text{Succ}$$

Możemy zatem wyekstrachować mniejsze drzewo dowodowe $\frac{\vdots}{n \text{ nat}}$ i tym samym udowodnić $n \text{ nat}$. Tym samym pokazaliśmy, że $Succ^{-1}$ jest regułą dopuszczalną.

Istotną cechą reguł wyprowadzalnych jest to, że pozostają one prawdziwe nawet wówczas, gdy wzbogacimy nasz system o nowe reguły. Na przykład reguła $Succ2$ pozostanie prawdziwa niezależnie od tego ile nowych reguł wnioskowania dołożymy, ponieważ dowód osądu $S \text{ } S \text{ } n \text{ nat}$ z osądu $n \text{ nat}$ jest zawsze możliwy w oparciu o regułę $Succ$.

Nie można, niestety, powiedzieć tego samego o regułach dopuszczalnych. Innymi słowy, reguły dopuszczalne mogą przestać obowiązywać po wzbogaceniu systemu o nowe reguły wnioskowania. Na przykład wyobraźmy sobie, że wyposażymy nasz system w nową regułę

$$\frac{n \text{ tree}}{S \text{ } n \text{ nat}} ZKosmosu$$

Reguła $ZKosmosu$ psuje poprzednio dopuszczalną regułę $Succ^{-1}$ ponieważ zastosowanie $Succ$ nie jest teraz jedynym sposobem udowodnienia $S \text{ } n \text{ nat}$ i tym samym $S \text{ } n \text{ nat}$ nie gwarantuje obecnie $n \text{ nat}$. Tym samym prawdziwość każdej z reguł dopuszczalnych musi być weryfikowana za każdym razem, gdy wzbogacamy system o nowe reguły.

Zadanie 3. Czy reguła $\frac{n \text{ par}}{S \text{ } S \text{ } n \text{ par}} SuccE^2$ jest wyprowadzalna, czy dopuszczalna? A reguła $\frac{S \text{ } S \text{ } n \text{ par}}{n \text{ par}} SuccE^{-2}$?