SEMANTYKA I WERYFIKACJA - ĆW. 1

1. Po co jest semantyka i ten przedmiot?
   - *Co robi program?*
   - *Czy dwa programy robią to samo?*
     - w jednym języku: poprawność optymalizacji
     - w dwóch językach: poprawność tłumaczenia
   - **Jasna** definicja języka programowania
     - dla programistów (*Co się stanie, jeśli w procedurze obsługi wyjątku podniosę ten sam wyjątek?*)
     - dla twórców kompilatorów itd. (*Czy mój kompilator jest poprawny? I co to znaczy?*)
   - **Formalna** definicja języka programowania
     - dowodzenie własności programów
     - generowanie kompilatorów, interpreterów
     - automatyczna analiza programów
   - Napięcie między jasnością a formalnością
   - Wspólny formalizm dla wielu języków
     - Systematyzacja pojęć (*Co to jest typ? A klasa? A kontynuacja, wyjątek, leniwość?*)
2. Idealnie: semantyka zrozumiana tak dobrze jak składnia
   - zalety BNF: formalizm, jasność, modularność
   - masa narzędzi, np. generatory parserów
   - z semantyką jest gorzej, bo:
     - zachowanie programów jest bardziej różnorodne niż składnia,
     - trudne aspekty składni nazywamy semantyką (poprawność typowania, widoczność zmiennych)
   - Różne podejścia:
     - operacyjna (*Co robi program?*)
     - denotacyjna (*Czym jest program?*)
     - aksjomatyczna (*Jakie własności ma program?*)

   Semantyka jest nie tylko dla programów, ale dla dowolnych **przepisów na obliczenia**. Przykład: wyrażenia arytmetyczne.
3. **Zadanie:** Napisać gramatykę bezkontekstową dla wyrażeń arytmetycznych.

Rozwiązanie:

$$\begin{aligned} N &::= 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid \cdots \\ E &::= N \mid (E) \mid E\ Op\ E \\ Op &::= + \mid - \mid * \mid / \end{aligned}$$

Ta gramatyka nie jest jednoznaczna. Narysować parę drzew parsowania dla wyrażenia $4 * 2 - 1$.

4. **Zadanie:** Przepisać tę gramatykę na jednoznaczną. *Rozwiązanie:*

$$\begin{aligned} N &::= 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid \cdots \\ E &::= E1 \mid E1\ LOp\ E \\ E1 &::= E2 \mid E2\ HOp\ E1 \\ E2 &::= N \mid (E) \\ LOp &::= + \mid - \\ HOp &::= * \mid / \end{aligned}$$

Narysować drzewo parsowania dla $4 \times 2 - 1$, upewnić się przy tym że jest ono jedyne.

5. Składnia konkretna a abstrakcyjna

   - Składnia konkretna: zbiór słów, przydatna w parsowaniu.
   - Składnia abstrakcyjna: zbiór legalnych drzew parsowania.
   - Parsowanie: mapowanie składni konkretnej (słowa) na abstrakcyjną (drzewa).
   - Definiując semantykę, będziemy zawsze pracować na składni abstrakcyjnej (czyli parsowaniem nie zawracamy sobie głowy).
   - **Zadanie:** Opisać semantykę abstrakcyjną dla wyrażeń. Powinna być taka jak nasza pierwsza gramatyka, ale uwaga: nawiasy niepotrzebne!

6. Co powinno być semantyką wyrażenia: liczba, czy coś więcej? Na przykład, czy wyrażenia $4 * 2 - 1$ i $3 + 4$ powinny mieć taką samą, czy różną semantykę? Obie odpowiedzi mają sens, na początek powiedzmy, że semantyką jest liczba.

   - *Dziedzina semantyczna*: $\mathbb{Q}$.
   - *Funkcja semantyczna*: $[\![\_]\!] : Expr \to \mathbb{Q}$, zdefiniowana przez indukcję po budowie wyrażenia:

   $$\begin{aligned} [\![n]\!] &= n \\ [\![e_1 + e_2]\!] &= [\![e_1]\!] + [\![e_2]\!] \qquad \text{itd.} \end{aligned}$$

   (Plus po lewej stronie to składnia, a po prawej – operacja na dziedzinie semantycznej.)

7. **Problem:** dzielenie przez zero.

   Rozwiązanie 1: funkcja semantyczna może być funkcją częściową.

   Rozwiązanie 2: Rozszerzenie dziedziny semantycznej o dodatkowy element (np. $\sharp$) i odpowiednie rozszerzenie definicji operacji na dziedzinie semantycznej ($5 + \sharp = \sharp$ itd.)

8. **Problem:** Uzupełnijmy gramatykę o mechanizm rozpoznawania liczb w dziesiętnym systemie pozycyjnym:

$$\begin{aligned} N &::= D \mid D\ N \\ D &::= 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9 \end{aligned}$$

   Jak napisać kompozycjonalną funkcję semantyczną w dziedzinę semantyczną $\mathbb{N}$? To się nie da zrobić (przykład: drzewo parsowania liczby 1005). Dwa możliwe rozwiązania:

   - Wzbogacenie dziedziny semantycznej tak, by umożliwiała kompozycjonalną definicję (w tym przypadku: wartość i liczba zer na początku, czyli $\mathbb{N} \times \mathbb{N}$), albo
   - Przeprojektowanie składni abstrakcyjnej, w tym przypadku

$$N ::= D \mid N\ D$$

9. **Zadanie:** Uzupełnijmy abstrakcyjną składnię wyrażeń o zmienne:

$$E ::= \cdots \mid Var$$

   Jaka jest wtedy odpowiednia dziedzina semantyczna?

   Rozwiązanie: Semantyka wyrażenia to funkcja z wartościowań w liczby, tj:

$$[\![\_]\!] : Expr \to ((Var \to \mathbb{Q}) \to \mathbb{Q})$$

   Funkcję semantyczną definiujemy tak:

$$\begin{aligned} [\![x]\!] &= \lambda\rho.\rho(x) \\ [\![n]\!] &= \lambda\rho.n \\ [\![e_1 + e_2]\!] &= \lambda\rho.([\![e_1]\!]\rho + [\![e_2]\!]\rho) \end{aligned}$$

10. Żeby tych zmiennych można było sensownie użyć, dodajmy też wyrażenia `let`. Składnia konkretna:

$$Expr ::= \cdots \mid \texttt{let}\ Var\ \ =\ \ Expr\ \texttt{in}\ Expr$$

   Składnia abstrakcyjna:

$$Expr ::= \cdots \mid Var\ Expr\ Expr$$

Ewentualnie ze znacznikiem, żeby nam się nie pomyliły podobne konstrukcje językowe:

$$Expr ::= \cdots \mid \texttt{let}\ Var\ Expr\ Expr$$

Semantyka:

$$[\![\texttt{let}\ x\ e_1\ e_2]\!] = \lambda\rho.[\![e_2]\!](\rho[x \mapsto [\![e_1]\!]\rho])$$