A fröccs szintaxisa és operációs szemantikája

Kaposi Ambrus
Eötvös Loránd Tudományegyetem
akaposi@inf.elte.hu

2017. október 2.

1. Bevezetés

Ipari partnerünk megkereste az Informatikai Kar Programozási Nyelvek és Fordítóprogramok Tanszékét abban a reményben, hogy segítünk neki jobb minőségű fröccsöt készíteni. Tanszékünkön programozási nyelvekkel foglalkozunk, melyek a természetes nyelvekhez hasonlóak, de szintaxisuk (mondattanuk) és szemantikájuk (jelentésük) teljesen precíz, nincs helye kétértelműségeknek.

2. Szintaxis

Ipari partnerünk képviselője elmagyarázta, mi a fröccs készítésének általános módszere: i) egy deciliteres adagokban rendelkezésre áll bor illetve szóda; ii) rendelkezésre állnak felülről nem korlátos méretű edények, melyek kezdetben üresek (tetszőleges méretű edény igény szerint legyártható); iii) bármely két edény tartalma összeönthető egy harmadik edénybe. A fröccskóstoló a iii) lépés során opcionálisan megkóstolja a kapott fröccsöt, ezzel elhanyagolható mennyiséget magához véve, majd, ha elégedett a kapott ízzel, a fröccsöt késznek nyilvánítja.

A fenti módszert egy programozási nyelvvel írjuk le, melynek szintaxisa a következő.

$$\begin{aligned} &\mathit{deci} && ::= \mathsf{bor} \,|\, \mathsf{szóda} \\ &\mathit{prefröccs} && ::= \mathit{deci} \,|\, \varepsilon \,|\, \mathit{prefröccs} + \mathit{prefröccs} \end{aligned} \tag{1}$$

Két **fajtánk** van, melyek metaváltozóit *deci*vel és *prefröccs*el, valamint ezeknek vesszőzött, indexelt változataival jelöljük (például deci', $\mathit{prefröccs}_1$). Egy *deci* kifejezést kétféle ún. **operátorral** tudunk létrehozni: bor és $\mathsf{szóda}$. Ezek nulláris operátorok, nincs egy **paraméterük** sem. Magát a *prefröccs*öt már háromféle módon létre tudunk hozni: i) kezdhetünk egy egy deciliteres adaggal (ez egy jelöletlen unáris operátor, paramétere egy *deci*); ii) egy üres edénnyel (ε, epszilon, nulláris operátor); iii) vagy két, már meglevő *prefröccs*ből a $+$ bináris operátorral kapunk egy újabbat.

1. Példa. Néhány példa *prefröccs*re:

a) ε

b) bor

c) szóda

d) bor + szóda

e) bor + (szóda + bor)

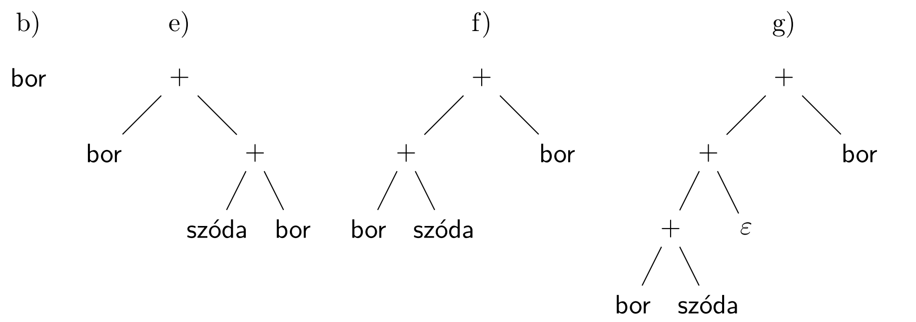
f) (bor + szóda) + bor

g) ((bor + szóda) + ε) + bor

h) ε + bor

i) ε + (ε + bor)

Figyeljük meg, hogy a kontextustól függően bor lehet *deci* és *prefröccs* is. A + operátor többször is alkalmazható, így jönnek létre például az e) és f) kifejezések, melyek különböznek egymástól: e)-ben egy bort öntünk hozzá egy szóda és bor keverékhez, míg f) esetében egy bor és szóda keveréket öntünk hozzá borhoz. Minden *prefröccs*nek megfeleltethető egy **szintaxisfa**, melynek gyökerét felülre rajzoljuk, lefele ágazik el, a fa levelei pedig vagy ε vagy egy *deci* lehetnek. Alább ábrázolunk néhányat a fenti példák közül.



A lineáris ábrázolás a számítógép számára egyszerűbben értelmezhető, és papíron is rövidebb, mint a szintaxisfa-változat, emiatt az előbbit fogjuk használni.

Azért a prefröccs kifejezést használjuk a szintaxisban leírt kifejezésekre, mert ezek között olyanok is vannak, melyek nem tartalmaznak szódát vagy nem tartalmaznak bort. Fröccsnek majd azokat a prefröccsöket szeretnénk nevezni, melyek mindkettő komponenst tartalmazzák. Erről szól a következő szakasz.

3. Típusrendszer

Az 1-ben megadott szintaxissal kifejezhető *prefröccs*ök közül soknak nincs gyakorlati jelentősége, ilyenek például az ε, $\varepsilon + (\varepsilon + \varepsilon)$, (bor + bor) + (bor + ε), szóda + szóda stb. A fröccsben ugyanis bornak és szódának is keverednie kell. A típusrendszer azokat az értelmetlen kifejezéseket szűri ki, melyek bár a szintaxissal kifejezhetők, mégsem szeretnénk, hogy a végleges programban megjelenjenek. Egy olyan programozási nyelvben, melyben például számok és szövegek vannak, és van egy összeadás operátorunk, nem szeretnénk megengedni azt a kifejezést, hogy 3 + `"hello"`, hiszen nem világos, mi legyen a jelentése.

Természetesen adódik, hogy hasonló módszerrel szűrjük ki a nem megfelelő *prefröccs*öket. A fröccsök programozási nyelvében három típus van, a következő módon megadva.

$$\mathit{Típus} ::= \mathsf{Vanbor} \mid \mathsf{Vanszóda} \mid \mathsf{Fröccs} \tag{2}$$

A Vanbor típusú *prefröccs*ökről tudni fogjuk, hogy van bennük bor, a Vanszóda típusúakról, hogy van bennük szóda, míg a Fröccs típusúak biztosan valódi fröccsök lesznek: bort és szódát is tartalmaznak.

A típusrendszert **levezetési szabályokkal** (vagy egyszerűen: szabályokkal) adjuk meg, melyekkel **ítéleteket** lehet levezetni. A mi nyelvünkben egyféle ítélet lesz, ezt *prefröccs* : Fröccs-nek jelöljük. Ezzel azt fejezzük ki, hogy az adott *prefröccs* fröccs.

A fröccsök típusrendszerének levezetési szabályai a következők.

$$\frac{}{\mathsf{bor} : \mathsf{Vanbor}} \tag{3}$$

$$\frac{}{\mathsf{szóda} : \mathsf{Vanszóda}} \tag{4}$$

$$\frac{\mathit{prefröccs}_1 : \mathsf{Vanbor} \qquad \mathit{prefröccs}_2 : \mathsf{Vanszóda}}{\mathit{prefröccs}_1 + \mathit{prefröccs}_2 : \mathsf{Fröccs}} \tag{5}$$

$$\frac{\mathit{prefröccs}_1 : \mathsf{Vanszóda} \qquad \mathit{prefröccs}_2 : \mathsf{Vanbor}}{\mathit{prefröccs}_1 + \mathit{prefröccs}_2 : \mathsf{Fröccs}} \tag{6}$$

$$\frac{\mathit{prefröccs}_1 : \mathit{Típus}}{\mathit{prefröccs}_1 + \mathit{prefröccs}_2 : \mathit{Típus}} \tag{7}$$

$$\frac{\mathit{prefröccs}_2 : \mathit{Típus}}{\mathit{prefröccs}_1 + \mathit{prefröccs}_2 : \mathit{Típus}} \tag{8}$$

A vízszintes vonal fölött találhatók a szabály **feltételei** (premisszái), alatta a szabály **következménye** (konklúziója). A 16 és 17 szabályokat **axiómáknak** nevezzük, mert egyetlen feltételük sincs. Az 5–6 szabályoknak két feltételük van, a 7–8 szabályoknak egy.

A levezetési szabályok azt fejezik ki, hogy ha a feltételeket már levezettük, akkor a következmény is levezethető. A mi szabályaink mögött az alábbi intuíció áll. Egyrészt, bor tartalmaz bort, szóda tartalmaz szódát (16–17 szabályok), valamint ha van egy bort tartalmazó prefröccsünk és egy szódát tartalmazó prefröccsünk, ezekből készíthetünk fröccsöt kétféleképpen is (5–6 szabályok): vagy a bort tartalmazót öntjük a szódát tartalmazóba, vagy fordítva. Végül a 7–8 szabályok azt fejezik ki, hogy ha már van egy adott típusú keverékünk, akkor annak, bármilyen prefröccsel (jobbról vagy balról) összeöntve is megmarad a típusa. Például, ha egy bort tartalmazó prefröccshöz hozzáöntjük a $(\mathsf{szóda} + \mathsf{bor}) + \varepsilon$ keveréket, az eredmény továbbra is fog bort tartalmazni.

A fenti 5–8 szabályok tulajdonképpen szabály-sémák: bármely metaváltozó helyére kifejezéseket írva megkapjuk a szabály egy konkrét változatát. Például a 8 szabály egy konkrét változata

$$\frac{\mathsf{bor} + \mathsf{szóda} : \mathsf{Vanbor}}{\varepsilon + (\mathsf{bor} + \mathsf{szóda}) : \mathsf{Vanbor},}$$

de

$$\frac{\textsf{bor} + \textsf{szóda} : \textsf{Vanbor}}{\varepsilon + (\textsf{bor} + \textsf{szóda}) : \textsf{Fröccs}}$$

nem, hiszen *Típus* mindkét előfordulását ugyanarra a típusra kell helyettesítenünk.

A levezetési szabályokat **levezetési fává** (vagy egyszerűen: levezetéssé) kombinálhatjuk. Azt mondjuk, hogy a *prefröccs* : Fröccs ítélet levezethető, ha létezik olyan levezetési fa, melynek gyökerénél *prefröccs* : Fröccs van. A levezetési fák gyökerét alulra írjuk, és felfelé ágaznak el, leveleiknél axiómák vannak. Például azt, hogy $\textsf{szóda} + ((\textsf{bor} + \varepsilon) + (\textsf{szóda} + \textsf{szóda}))$ fröccs, a következőképp vezetjük le (a fa elágazásaihoz írjuk a használt szabályok sorszámát).

$$\dfrac{\dfrac{\dfrac{\dfrac{}{\textsf{bor} : \textsf{Vanbor}}\,16}{\textsf{bor} + \varepsilon : \textsf{Vanbor}}\,7 \qquad \dfrac{\dfrac{}{\textsf{szóda} : \textsf{Vanszóda}}\,17}{\textsf{szóda} + \textsf{szóda} : \textsf{Vanszóda}}\,8}{(\textsf{bor} + \varepsilon) + (\textsf{szóda} + \textsf{szóda}) : \textsf{Fröccs}}\,5}{\textsf{szóda} + ((\textsf{bor} + \varepsilon) + (\textsf{szóda} + \textsf{szóda})) : \textsf{Fröccs}}\,8$$

Azt, hogy $(\textsf{szóda} + \varepsilon) + \textsf{szóda} : \textsf{Fröccs}$, nem tudjuk levezetni, mert bármelyik szabályt is alkalmazzuk az 5–8 közül, mindig elakadunk valahol (a 16–17 szabályokat nem tudjuk az első lépésben alkalmazni, hiszen nem + operátorral megadott kifejezéseket vezetnek le).

$$\dfrac{\dfrac{\dfrac{?}{\textsf{szóda} : \textsf{Vanbor}}}{\textsf{szóda} + \varepsilon : \textsf{Vanbor}}\,7 \qquad \dfrac{}{\textsf{szóda} : \textsf{Vanszóda}}\,17}{(\textsf{szóda} + \varepsilon) + \textsf{szóda} : \textsf{Fröccs}}\,5$$

$$\dfrac{\dfrac{\dfrac{?}{\varepsilon : \textsf{Vanbor}}}{\textsf{szóda} + \varepsilon : \textsf{Vanbor}}\,8 \qquad \dfrac{}{\textsf{szóda} : \textsf{Vanszóda}}\,17}{(\textsf{szóda} + \varepsilon) + \textsf{szóda} : \textsf{Fröccs}}\,5$$

$$\dfrac{\dfrac{\dfrac{}{\textsf{szóda} : \textsf{Vanszóda}}\,17}{\textsf{szóda} + \varepsilon : \textsf{Vanszóda}}\,7 \qquad \dfrac{?}{\textsf{szóda} : \textsf{Vanbor}}}{(\textsf{szóda} + \varepsilon) + \textsf{szóda} : \textsf{Fröccs}}\,6$$

$$\dfrac{\dfrac{\dfrac{?}{\textsf{szóda} : \textsf{Vanbor}} \qquad \dfrac{?}{\varepsilon : \textsf{Vanszóda}}}{\textsf{szóda} + \varepsilon : \textsf{Fröccs}}\,5}{(\textsf{szóda} + \varepsilon) + \textsf{szóda} : \textsf{Fröccs}}\,7$$

$$\dfrac{\dfrac{\dfrac{}{\textsf{szóda} : \textsf{Vanszóda}}\,17 \qquad \dfrac{?}{\varepsilon : \textsf{Vanbor}}}{\textsf{szóda} + \varepsilon : \textsf{Fröccs}}\,6}{(\textsf{szóda} + \varepsilon) + \textsf{szóda} : \textsf{Fröccs}}\,7$$

$$\dfrac{\dfrac{\dfrac{?}{\textsf{szóda} : \textsf{Fröccs}}}{\textsf{szóda} + \varepsilon : \textsf{Fröccs}}\,7}{(\textsf{szóda} + \varepsilon) + \textsf{szóda} : \textsf{Fröccs}}\,7$$

$$\dfrac{\dfrac{\dfrac{?}{\varepsilon : \textsf{Fröccs}}}{\textsf{szóda} + \varepsilon : \textsf{Fröccs}}\,8}{(\textsf{szóda} + \varepsilon) + \textsf{szóda} : \textsf{Fröccs}}\,7$$

$$\dfrac{\dfrac{?}{\textsf{szóda} : \textsf{Fröccs}}}{(\textsf{szóda} + \varepsilon) + \textsf{szóda} : \textsf{Fröccs}}\,8$$

Típusozhatónak nevezzük azt a *prefröccs*öt, melyre létezik olyan *Típus*, hogy *prefröccs* : *Típus*. Például az $\varepsilon + \varepsilon$ prefröccs nem típusozható.

2. Megjegyzés. Vegyük észre, hogy ha a 7–8 szabályok helyett az alábbi szabályt adtuk volna meg, nem tudnánk levezetni néhány fröccsöt.

$$\dfrac{\mathit{prefröccs}_1 : \mathit{Típus} \qquad \mathit{prefröccs}_2 : \mathit{Típus}}{\mathit{prefröccs}_1 + \mathit{prefröccs}_2 : \mathit{Típus}}$$

Például az $\varepsilon + (\textsf{bor} + \textsf{szóda}) : \textsf{Fröccs}$ nem lenne levezethető, hiszen az ε nem fröccs.

Ez a szabály azonban **megengedhető**, tehát minden olyan esetben, melyben a feltételei igazak, a mi levezetési szabályaink segítségével is levezethető a következmény (konkrétan egy lépésben, a 7 vagy a 8 szabály használatával).

3. Megjegyzés. Van olyan prefröccs, melyekre többféleképpen is levezethető, hogy fröccs. Például ilyen a $(\textsf{bor} + \textsf{szóda}) + (\textsf{bor} + \textsf{szóda})$.

A következő lemma azt fejezi ki, hogy a Fröccs altípusa a Vanbor és a Vanszóda típusoknak.

4. Lemma. Ha *prefröccs* : Fröccs levezethető, akkor *prefröccs* : Vanbor és *prefröccs* : Vanszóda is levezethető.

Bizonyítás. *prefröccs* : Fröccs levezetése szerinti strukturális indukció. □

4. Szemantika

A szemantika azzal foglalkozik, hogy a szintaxisban megadott kifejezéseknek jelentést adjon.

A jelentést egy szemantikus függvénnyel adjuk meg, mely egy *prefröccs* kifejezéshez egy $[\![\mathit{prefröccs}]\!]$ jelentést társít.

A jelentésről ipari partnerünktől a következő információkat kaptuk.

– Ha először összeöntünk két keveréket, majd hozzá egy harmadikat, az ugyanaz, mintha először a másodikat és a harmadikat öntöttük volna össze, majd az elsőt ezzel öntöttük volna össze. Tehát például

$$[\![(\textsf{bor} + \textsf{szóda}) + \textsf{bor}]\!] = [\![\textsf{bor} + (\textsf{szóda} + \textsf{bor})]\!],$$

általánosságban pedig

$$\begin{aligned} &[\![(\mathit{prefröccs}_1 + \mathit{prefröccs}_2) + \mathit{prefröccs}_3]\!] \\ &= [\![\mathit{prefröccs}_1 + (\mathit{prefröccs}_2 + \mathit{prefröccs}_3)]\!]. \end{aligned}$$

Ezt **asszociativitásnak** (átcsoportosíthatóság) nevezzük.

– Ha üreset öntünk hozzá egy keverékhez, az olyan, mintha nem csináltunk volna semmit: $[\![\varepsilon + \mathit{prefröccs}]\!] = [\![\mathit{prefröccs}]\!]$. Ezt **identitás tulajdonságnak** nevezzük.

– Az összeöntés sorrendje nem számít, tehát például $[\![\mathsf{bor}+\mathsf{szóda}]\!] = [\![\mathsf{szóda}+\mathsf{bor}]\!]$, általánosságban $[\![\mathit{prefröccs}_1+\mathit{prefröccs}_2]\!] = [\![\mathit{prefröccs}_2+\mathit{prefröccs}_1]\!]$. Ezt a tulajdonságot **kommutativitásnak** (felcserélhetőség) nevezzük.

Ezek a tulajdonságok nem a szintaxisról mondanak el valamit, hanem a szintaktikus kifejezések jelentéséről.

4.1. Denotációs szemantika

A denotációs szemantika valamilyen matematikai struktúrában adja meg a szintaktikus kifejezések jelentését. A *prefröccs* strukturája szerinti rekurzióval adjuk meg, bármely *prefröccs*re $[\![\mathit{prefröccs}]\!] \in \mathbb{N} \times \mathbb{N}$ lesz, tehát a jelentés két természetes számból álló rendezett pár lesz: az első azt adja meg, hogy hány deci bor van az adott prefröccsben, a második azt, hogy hány deci szóda.

$$
\begin{aligned}
[\![\mathsf{bor}]\!] &:= (1{,}0)\\
[\![\mathsf{szóda}]\!] &:= (0{,}1)\\
[\![\varepsilon]\!] &:= (0{,}0)\\
[\![\mathit{prefröccs}_1 + \mathit{prefröccs}_2]\!] &:= (\mathsf{proj}_1\,[\![\mathit{prefröccs}_1]\!] + \mathsf{proj}_1\,[\![\mathit{prefröccs}_2]\!],\\
&\qquad \mathsf{proj}_2\,[\![\mathit{prefröccs}_1]\!] + \mathsf{proj}_2\,[\![\mathit{prefröccs}_2]\!])
\end{aligned}
$$

A proj_1 és a proj_2 az első és a második projekció (vetítés) függvények, melyek egy rendezett pár első illetve második elemét adják meg, például $\mathsf{proj}_1\,(3{,}4) = 3$ és $\mathsf{proj}_2\,(3{,}4) = 4$. Vegyük észre, hogy a +-ot itt kétféle értelemben is használjuk: egyrészt egy operátor, mely két prefröccsből készít egy újabb prefröccsöt, másrészt a metaelméleti összeadás természetes számokra.

Ez a definíció egy eljárást (funkcionális programot [2]) ad meg, mellyel ki tudjuk számolni tetszőleges prefröccs jelentését a fenti egyenlőségek mentén. Az eljárás a következő: a fenti := egyenlőségeket alkalmazhatjuk a kifejezés tetszőleges helyén. A sorrend nincs meghatározva, egy lehetséges számítási sorozat például a következő.

$$
\begin{aligned}
&[\![(\mathsf{bor} + \mathsf{szóda}) + \varepsilon]\!]\\
={}&(\mathsf{proj}_1\,[\![\mathsf{bor} + \mathsf{szóda}]\!] + \mathsf{proj}_1\,[\![\varepsilon]\!], \mathsf{proj}_2\,[\![\mathsf{bor} + \mathsf{szóda}]\!] + \mathsf{proj}_2\,[\![\varepsilon]\!])\\
={}&(\mathsf{proj}_1\,[\![\mathsf{bor} + \mathsf{szóda}]\!] + \mathsf{proj}_1\,(0{,}0), \mathsf{proj}_2\,[\![\mathsf{bor} + \mathsf{szóda}]\!] + \mathsf{proj}_2\,(0{,}0))\\
={}&(\mathsf{proj}_1\,[\![\mathsf{bor} + \mathsf{szóda}]\!] + 0, \mathsf{proj}_2\,[\![\mathsf{bor} + \mathsf{szóda}]\!] + 0)\\
={}&(\mathsf{proj}_1\,[\![\mathsf{bor} + \mathsf{szóda}]\!], \mathsf{proj}_2\,[\![\mathsf{bor} + \mathsf{szóda}]\!])\\
={}&(\mathsf{proj}_1\,(\mathsf{proj}_1\,[\![\mathsf{bor}]\!] + \mathsf{proj}_1\,[\![\mathsf{szóda}]\!], \mathsf{proj}_2\,[\![\mathsf{bor}]\!] + \mathsf{proj}_2\,[\![\mathsf{szóda}]\!]),\\
&\ \mathsf{proj}_2\,(\mathsf{proj}_1\,[\![\mathsf{bor}]\!] + \mathsf{proj}_1\,[\![\mathsf{szóda}]\!], \mathsf{proj}_2\,[\![\mathsf{bor}]\!] + \mathsf{proj}_2\,[\![\mathsf{szóda}]\!]))\\
={}&(\mathsf{proj}_1\,(\mathsf{proj}_1\,(1{,}0) + \mathsf{proj}_1\,(0{,}1), \mathsf{proj}_2\,(1{,}0) + \mathsf{proj}_2\,(0{,}1)),\\
&\ \mathsf{proj}_2\,(\mathsf{proj}_1\,(1{,}0) + \mathsf{proj}_1\,(0{,}1), \mathsf{proj}_2\,(1{,}0) + \mathsf{proj}_2\,(0{,}1)))\\
={}&(\mathsf{proj}_1\,(1 + 0{,}0 + 1), \mathsf{proj}_2\,(1 + 0{,}0 + 1))\\
={}&(\mathsf{proj}_1\,(1{,}1), \mathsf{proj}_2\,(1{,}1))\\
={}&(1{,}1)
\end{aligned}
$$

Sokféle módon el lehet készíteni a nagyfröccsöt, de ezeknek a jelentése mind meg fog egyezni. Néhány példa:

$$[\![\mathsf{bor}+(\mathsf{bor}+\mathsf{szóda})]\!] = [\![\mathsf{bor}+(\mathsf{szóda}+(\mathsf{bor}+\varepsilon))]\!] = [\![\mathsf{szóda}+(\mathsf{bor}+\mathsf{bor})]\!] = (2{,}1).$$

5. Lemma. A fenti szemantikára teljesül az asszociativitás, identitás és kommutativitás tulajdonság.

Bizonyítás. Közvetlenül a $[\![-]\!]$ függvény definíciójából. $\square$

4.2. Operációs szemantika

Az operációs szemantika a denotációs szemantikánál konkrétabb. Utóbbinál a szintaxis (prefröccsök) jelentését egy metaelméleti strukturában (természetes számok párjai) adtuk meg. Előbbiben viszont a szintaxis jelentése szintén szintaxis: megadjuk, mely szintaktikus kifejezések a **normál formák**, és megadunk egy **átírási rendszert**, mely leírja, hogyan kell bármely prefröccsöt normál formára írni. Abból a szempontból is konkrétabb az operációs szemantika, hogy a jelentés kiszámításának pontos sorrendjét is tartalmazza, míg az előző alfejezetben láttuk, hogy a számítási sorrend denotációs szemantika esetén nincs megkötve (pontosabban a metanyelvünkből öröklődik).

A normál formából leolvasható, hogy pontosan mennyi bor illetve szóda van az adott prefröccsben. A normál formákat háromféle ítélettel adjuk meg: bizonyos prefröccsöket borreceptnek ($\mathit{prefröccs}\ \mathsf{borrecept}$), másokat szódareceptnek ($\mathit{prefröccs}\ \mathsf{szódarecept}$) nevezünk. Egy normál forma (recept) két prefröccsből áll ($\mathit{prefröccs}_1, \mathit{prefröccs}_2\ \mathsf{recept}$). Ezek az ítéletek a következő levezetési szabályokkal vannak megadva.

$$\frac{}{\varepsilon\ \mathsf{borrecept}} \tag{9}$$

$$\frac{\mathit{prefröccs}\ \mathsf{borrecept}}{\mathsf{bor} + \mathit{prefröccs}\ \mathsf{borrecept}} \tag{10}$$

$$\frac{}{\varepsilon\ \mathsf{szódarecept}} \tag{11}$$

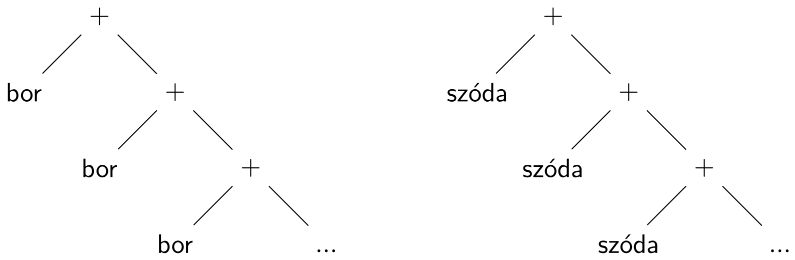
$$\frac{\mathit{prefröccs}\ \mathsf{szódarecept}}{\mathsf{szóda} + \mathit{prefröccs}\ \mathsf{szódarecept}} \tag{12}$$

$$\frac{\mathit{prefröccs}_1\ \mathsf{borrecept} \qquad \mathit{prefröccs}_2\ \mathsf{szódarecept}}{\mathit{prefröccs}_1, \mathit{prefröccs}_2\ \mathsf{recept}} \tag{13}$$

Borrecept vagy egy üres prefröccs (9 szabály), vagy egy bort hozzáadhatunk egy már létrehozott borrecepthez (10 szabály). A 11–12 szabályok ugyanezt fejezik ki a szódareceptekre. Tehát a következő prefröccsökre tudjuk levezetni, hogy ők borreceptek illetve szódareceptek.

ε	ε
$\mathsf{bor} + \varepsilon$	$\mathsf{szóda} + \varepsilon$
$\mathsf{bor} + (\mathsf{bor} + \varepsilon)$	$\mathsf{szóda} + (\mathsf{szóda} + \varepsilon)$
$\mathsf{bor} + (\mathsf{bor} + (\mathsf{bor} + \varepsilon))$	$\mathsf{szóda} + (\mathsf{szóda} + (\mathsf{szóda} + \varepsilon))$
...	...

A szintaxisfák tehát mindig csak jobbra ágaznak tovább, általánosságban a következőképp ábrázolhatók.



A 13 szabály azt mondja ki, hogy egy recept egy borreceptből és egy szódareceptből áll. A vessző a szabályban csak jelölés, nem egy operátor. Az alábbi egy példa levezetés, melynek konklúziója $(\textsf{bor}+(\textsf{bor}+\varepsilon)),(\textsf{szóda}+\varepsilon)\ \textsf{recept}$.

$$\dfrac{\dfrac{\dfrac{\dfrac{}{\varepsilon\ \textsf{borrecept}}\ 9}{\textsf{bor}+\varepsilon\ \textsf{borrecept}}\ 10}{\textsf{bor}+(\textsf{bor}+\varepsilon)\ \textsf{borrecept}}\ 10 \qquad \dfrac{\dfrac{}{\varepsilon\ \textsf{szódarecept}}\ 11}{\textsf{szóda}+\varepsilon\ \textsf{szódarecept}}\ 12}{(\textsf{bor}+(\textsf{bor}+\varepsilon)),(\textsf{szóda}+\varepsilon)\ \textsf{recept}}\ 13$$

Az átírási rendszert kétféle ítélettel adjuk meg. Egyrészt megadjuk a szódareceptek illetve borreceptek összefűzését a $\mathit{prefröccs}_1 \mathbin{+\!\!+} \mathit{prefröccs}_2 \Downarrow \mathit{prefröccs}$ formájú ítélettel, melyet úgy olvasunk, hogy $\mathit{prefröccs}_1$ és $\mathit{prefröccs}_2$ összefűzése *prefröccs*öt eredményez. Másrészt a normalizálás ítélete $\mathit{prefröccs} \Downarrow \mathit{prefröccs}_1, \mathit{prefröccs}_2$ alakú, ezt úgy olvassuk, hogy *prefröccs* normál formája a $\mathit{prefröccs}_1$ borreceptből és $\mathit{prefröccs}_2$ szódareceptből álló pár. A levezetési szabályok a következők (a 19 szabály feltételeit több sorban írtuk az olvashatóság kedvéért).

$$\dfrac{}{\varepsilon \mathbin{+\!\!+} \mathit{prefröccs} \Downarrow \mathit{prefröccs}} \tag{14}$$

$$\dfrac{\mathit{prefröccs}_2 \mathbin{+\!\!+} \mathit{prefröccs}_3 \Downarrow \mathit{prefröccs}}{(\mathit{prefröccs}_1 + \mathit{prefröccs}_2) \mathbin{+\!\!+} \mathit{prefröccs}_3 \Downarrow \mathit{prefröccs}_1 + \mathit{prefröccs}} \tag{15}$$

$$\dfrac{}{\textsf{bor} \Downarrow \textsf{bor} + \varepsilon, \varepsilon} \tag{16}$$

$$\dfrac{}{\textsf{szóda} \Downarrow \varepsilon, \textsf{szóda} + \varepsilon} \tag{17}$$

$$\dfrac{}{\varepsilon \Downarrow \varepsilon, \varepsilon} \tag{18}$$

$$\dfrac{\begin{array}{l} \mathit{prefröccs}_1 \Downarrow \mathit{prefröccs}_{11}, \mathit{prefröccs}_{12} \\ \mathit{prefröccs}_2 \Downarrow \mathit{prefröccs}_{21}, \mathit{prefröccs}_{22} \\ \mathit{prefröccs}_{11} \mathbin{+\!\!+} \mathit{prefröccs}_{12} \Downarrow \mathit{prefröccs}'_1 \\ \mathit{prefröccs}_{21} \mathbin{+\!\!+} \mathit{prefröccs}_{22} \Downarrow \mathit{prefröccs}'_2 \end{array}}{\mathit{prefröccs}_1 + \mathit{prefröccs}_2 \Downarrow \mathit{prefröccs}'_1, \mathit{prefröccs}'_2} \tag{19}$$

A 14–15 szabályokat mindig úgy fogjuk alkalmazni, hogy a *prefröccs*ök mind borreceptek vagy szódareceptek. A 14 szabály azt fejezi ki, hogy ha egy üres prefröccsöt összefűzünk bármely prefröccsel, az nem változik meg. A 15 szabály azt mondja, hogy ha egy összetett receptet fűzünk össze egy prefröccsel (nota bene $\mathit{prefröccs}_1$ szóda- vagy borrecept), akkor először $\mathit{prefröccs}_2$-t fűzzük össze $\mathit{prefröccs}_3$-mal, majd ennek az eredménye elé rakjuk be $\mathit{prefröccs}_1$-et.

A 16–18 szabályok megadják az egyszerű prefröccsök normalizálását: egy borból egy egy bort tartalmazó borreceptet és egy ε-t kapunk, mely a nulla szódát szimbolizálja. Ennek ellenkezőjét kapjuk szóda normalizálásakor, az üres fröccs normalizálásakor pedig két üres komponenst kapunk. A 19 szabály azt fejezi ki, hogy $+$-al megadott prefröccsök esetén először mindkét komponenst ($\mathit{prefröccs}_1, \mathit{prefröccs}_2$) normalizáljuk, majd ezek eredményét összefűzzük, és itt felhasználjuk a $\mathbin{+\!\!+}$ ítéletet.

Példaképp levezetjük a $((\mathsf{bor} + \varepsilon) + \mathsf{bor}) + \mathsf{szóda}$ kifejezés normalizálását.

Az alábbi levezetést elnevezzük p_1-nek.

$$\dfrac{\dfrac{}{\mathsf{bor} \Downarrow \mathsf{bor} + \varepsilon, \varepsilon}\,16 \quad \dfrac{}{\varepsilon \Downarrow \varepsilon, \varepsilon}\,18 \quad \dfrac{\dfrac{}{\varepsilon \mathbin{+\!\!+} \varepsilon \Downarrow \varepsilon}\,14}{(\mathsf{bor} + \varepsilon) \mathbin{+\!\!+} \varepsilon \Downarrow \mathsf{bor} + \varepsilon}\,15 \quad \dfrac{}{\varepsilon \mathbin{+\!\!+} \varepsilon \Downarrow \varepsilon}\,14}{\mathsf{bor} + \varepsilon \Downarrow \mathsf{bor} + \varepsilon, \varepsilon}\,19$$

Az alábbi levezetést elnevezzük p_2-nek (felhasználjuk benne p_1-et).

$$\dfrac{p_1 \quad \dfrac{}{\mathsf{bor} \Downarrow \mathsf{bor} + \varepsilon, \varepsilon}\,16 \quad \dfrac{\dfrac{}{\varepsilon \mathbin{+\!\!+} (\mathsf{bor} + \varepsilon) \Downarrow \mathsf{bor} + \varepsilon}\,14}{(\mathsf{bor} + \varepsilon) \mathbin{+\!\!+} (\mathsf{bor} + \varepsilon) \Downarrow \mathsf{bor} + (\mathsf{bor} + \varepsilon)}\,15 \quad \dfrac{}{\varepsilon \mathbin{+\!\!+} \varepsilon \Downarrow \varepsilon}\,14}{((\mathsf{bor} + \varepsilon) + \mathsf{bor}) \Downarrow \mathsf{bor} + (\mathsf{bor} + \varepsilon), \varepsilon}\,19$$

Az alábbi levezetést elnevezzük p_3-nak.

$$\dfrac{\dfrac{\dfrac{}{\varepsilon \mathbin{+\!\!+} \varepsilon \Downarrow \varepsilon}\,14}{(\mathsf{bor} + \varepsilon) \mathbin{+\!\!+} \varepsilon \Downarrow \mathsf{bor} + \varepsilon}\,15}{(\mathsf{bor} + (\mathsf{bor} + \varepsilon)) \mathbin{+\!\!+} \varepsilon \Downarrow \mathsf{bor} + (\mathsf{bor} + \varepsilon)}\,15$$

A kívánt levezetést a következőképp kapjuk meg.

$$\dfrac{p_2 \quad \dfrac{}{\mathsf{szóda} \Downarrow \varepsilon, \mathsf{szóda} + \varepsilon}\,17 \quad p_3 \quad \dfrac{}{\varepsilon \mathbin{+\!\!+} (\mathsf{szóda} + \varepsilon) \Downarrow \mathsf{szóda} + \varepsilon}\,14}{((\mathsf{bor} + \varepsilon) + \mathsf{bor}) + \mathsf{szóda} \Downarrow \mathsf{bor} + (\mathsf{bor} + \varepsilon), \mathsf{szóda} + \varepsilon}\,19$$

Az operációs szemantika úgy van megadva, hogy minden *prefröccs*öt normalizálni tudunk egy másik prefröccsre.

6. Lemma. Minden $\mathit{prefröccs}_1, \mathit{prefröccs}_2$-re létezik olyan *prefröccs*, hogy

$$\mathit{prefröccs}_1 \mathbin{+\!\!+} \mathit{prefröccs}_2 \Downarrow \mathit{prefröccs}.$$

Bizonyítás. A $\mathit{prefröccs}_1$ szerkezete szerinti indukcióval. □

7. Tétel. Minden *prefröccs*re igaz, hogy létezik olyan $\mathit{prefröccs}_1$ és $\mathit{prefröccs}_2$, hogy $\mathit{prefröccs} \Downarrow \mathit{prefröccs}_1, \mathit{prefröccs}_2$.

Bizonyítás. A *prefröccs* szerkezete szerinti indukcióval. □

A normalizálás végeredménye recept.

8. Lemma. Ha $\mathit{prefröccs}_1$ $\textsf{borrecept}$, $\mathit{prefröccs}_2$ $\textsf{borrecept}$ és $\mathit{prefröccs}_1 + \mathit{prefröccs}_2 \Downarrow \mathit{prefröccs}$, akkor $\mathit{prefröccs}$ $\textsf{borrecept}$, és hasonlóan szódára is.

Bizonyítás. $\mathit{prefröccs}_1 + \mathit{prefröccs}_2 \Downarrow \mathit{prefröccs}$ levezetése szerinti indukció. □

9. Tétel. Ha $\mathit{prefröccs} \Downarrow \mathit{prefröccs}_1, \mathit{prefröccs}_2$, akkor $\mathit{prefröccs}_1, \mathit{prefröccs}_2$ $\textsf{recept}$.

Bizonyítás. $\mathit{prefröccs} \Downarrow \mathit{prefröccs}_1, \mathit{prefröccs}_2$ levezetése szerinti indukció az előző lemmát felhasználva. □

A normalizálás egyértelmű.

10. Lemma. Ha $\mathit{prefröccs}_1 + \mathit{prefröccs}_2 \Downarrow \mathit{prefröccs}$ és $\mathit{prefröccs}_1 + \mathit{prefröccs}_2 \Downarrow \mathit{prefröccs}'$, akkor $\mathit{prefröccs} = \mathit{prefröccs}'$.

Bizonyítás. $\mathit{prefröccs}_1 + \mathit{prefröccs}_2 \Downarrow \mathit{prefröccs}$ és $\mathit{prefröccs}_1 + \mathit{prefröccs}_2 \Downarrow \mathit{prefröccs}'$ levezetése szerinti indukció. □

11. Tétel. Ha $\mathit{prefröccs} \Downarrow \mathit{prefröccs}_1, \mathit{prefröccs}_2$ és $\mathit{prefröccs} \Downarrow \mathit{prefröccs}'_1, \mathit{prefröccs}'_2$, akkor $\mathit{prefröccs}_1 = \mathit{prefröccs}'_1$ és $\mathit{prefröccs}_2 = \mathit{prefröccs}'_2$.

Bizonyítás. A $\mathit{prefröccs} \Downarrow \mathit{prefröccs}_1, \mathit{prefröccs}_2$ és $\mathit{prefröccs} \Downarrow \mathit{prefröccs}'_1, \mathit{prefröccs}'_2$ levezetése szerinti indukció. □

A fenti tételek ismeretében megadhatjuk a $[\![-]\!]$ szemantikus függvényt: ez egy prefröccshöz prefröccsöt fog rendelni a következőképp.

$$[\![\mathit{prefröccs}]\!] := \mathit{prefröccs}_1 + \mathit{prefröccs}_2, \text{ ahol } \mathit{prefröccs} \Downarrow \mathit{prefröccs}_1, \mathit{prefröccs}_2$$

Ilyen $\mathit{prefröccs}_1$ és $\mathit{prefröccs}_2$ a 7. tétel alapján létezik, és a 11. tétel alapján egyértelmű. Mivel bizonyításaink konstruktívak [3], ezért egy futtatható algoritmust is megadnak, mely kiszámítja a függvényt.

12. Feladat. Mutassuk meg, hogy erre a $[\![-]\!]$ függvényre igazak az identitás, asszociativitás, kommutativitás tulajdonságok.

Az operációs szemantika úgy van megadva, hogy nem rontja el a prefröccs típusát (típusmegőrzés tétele).

13. Lemma. Ha $\mathit{prefröccs}_1 + \mathit{prefröccs}_2 \Downarrow \mathit{prefröccs}$ és $\mathit{prefröccs}_1 + \mathit{prefröccs}_2 : A$, akkor $\mathit{prefröccs} : A$.

Bizonyítás. $\mathit{prefröccs}_1 + \mathit{prefröccs}_2 \Downarrow \mathit{prefröccs}$ levezetése szerinti indukció. □

14. Tétel. Ha $\mathit{prefröccs} \Downarrow \mathit{prefröccs}_1, \mathit{prefröccs}_2$ és $\mathit{prefröccs} : A$, akkor $\mathit{prefröccs}_1 + \mathit{prefröccs}_2 : A$.

Bizonyítás. A $\mathit{prefröccs} \Downarrow \mathit{prefröccs}_1, \mathit{prefröccs}_2$ levezetése szerinti indukció. □

5. Összefoglalás

A fentiekben a fröccs példáján illusztráltuk, hogy hogyan épülnek fel a programozási nyelvek. Akik további részletek iránt érdeklődnek, azoknak ajánljuk Bob Harper könyvét [1], ahol hasonló módszerrel valódi programozási nyelvek vannak felépítve. A fent említett szerkezeti rekurzió és levezetés szerinti indukció elvét a könyv 2. fejezete tárgyalja.

A típusrendszerek és a matematika kapcsolatát kiaknázó programozási nyelv a típuselmélet (type theory), ennek megismeréséhez ajánljuk a [4] könyvet és a típuselmélet Agda nevű implementációját [5]. Az összes fenti definíciót és tételt formalizáltuk Agdában. A formalizálás elérhető a szerző honlapján.

Köszönetnyilvánítás

Köszönöm a javaslatokat Kovács Andrásnak, Kozsik Tamásnak és Poór Artúrnak.

Hivatkozások

[1] Robert Harper. *Practical Foundations for Programming Languages. Second edition.* Cambridge University Press, New York, NY, USA, 2016.

[2] John Hughes. Why functional programming matters. *The Computer Journal*, 32(2):98–107, April 1989.

[3] Per Martin-Löf. Constructive mathematics and computer programming. In *Proc. Of a Discussion Meeting of the Royal Society of London on Mathematical Logic and Programming Languages*, pages 167–184, Upper Saddle River, NJ, USA, 1985. Prentice-Hall, Inc.

[4] The Univalent Foundations Program. Homotopy type theory: Univalent foundations of mathematics. Technical report, Institute for Advanced Study, 2013.

[5] The Agda development team. Agda wiki, 2017.