

Forritunarmálið *Baldur*

Þættir

Notendahandbók

Verkefni í Þýðendum, haust 2003

ARNAR BIRGISSON

Þættir þýðendum

EFNISYFIRLIT

I	INNGANGUR	4
II	MÁLFRÆÐI	5
	II.1 Forgangur og tenging – 7	
III	MERKINGARFRÆÐI	8
	III.1 Atriði aðeins í Fjölni – 8 • III.2 Ný atriði í Baldri – 8	
IV	DÆMI	12
V	NOTKUN	14
	v.1 Uppsetning – 14 • v.2 Keyrsla – 14	
VI	HEIMILDIR	16

KAFLI I

INNGANGUR

Forritunarmálið *Baldur* var unnið sem verkefni í námskeiðinu *Þýðendur* við Háskóla Íslands haustið 2003. Málið er að mestu leyti byggt á forritunarmálinu Fjölni og leitast er við að útfæra sem flest atriði Fjölnis í Baldri.

Frá upphafi var markmið verkefnisins að bæta við virkni Fjölnis bálkmótun, þ.e. möguleika á földuðum innri stefjum. Auk þess var lagt upp með að stef í Baldri nýttu sér bestun halakalla (*tail-call optimization*) þar sem þess var kostur. Það er mat höfundar að báðum þessu markmiðum hafi verið náð á frumgerðarstigi.

Hugmyndir úr Baldri mætti í framtíðinni nota til að útvíkka síðari útgáfur Fjölnis, hugsanlega með möguleika á þróaðri fallsforritun þar sem vakningarfærslur yrðu geymdar í kös og notandanum gert kleyft að smíða og vinna með lokanir.

Nafn forritunarmálsins, *Baldur*, á sér enga aðra skýringu en að vera þjálft í notkun. Á forsíðu þessarar handbókar er nafnið ritað með Tengwar Sindaril lettri, ᚷᚢᚳᚢᚱ , en letrið er hugarsmíð J.R.R. Tolkiens.

KAFLI II

MÁLFRÆÐI

Málfræði Baldurs verður hér sett fram sem s.k. *PEG* mállýsing. Varðandi merkingarfræði hennar vísast í [1].

<i>Staffasti</i>	←	['] ([' \] . / <i>Stýristafur</i>) [']
<i>Strengur</i>	←	'" ([" \] . / <i>Stýristafur</i>) * '"
<i>Stýristafur</i>	←	'\\$' <i>Hexstafur</i> <i>Hexstafur</i> / '\'
<i>Hexstafur</i>	←	[0-9A-Fa-f]
<i>Fjöldatala</i>	←	'\$' <i>Hexstafur</i> ⁺ / [0-9] ⁺
<i>Heitala</i>	←	[-]* <i>Fjöldatala</i>
<i>Fleytitala</i>	←	[-]* [0-9] ⁺ ' . ' [0-9]* ([eE] [+ -]? [0-9] ⁺)?
<i>Aðgerðastafur</i>	←	[+ - * / % ! & = ? < > : ^ @]
<i>Aðgerð</i>	←	'\ ' <i>Nafn</i> / <i>Aðgerðastafur</i> ⁺
<i>Nafnstafur</i>	←	!(<i>Aðgerðastafur</i> / [' " \ 0-9] / [ASCII stafir 0-31]) .
<i>Nafn</i>	←	<i>Nafnstafur</i> (<i>Nafnstafur</i> / [0-9])*

Aftast í öllum reglum hér að ofan er ótilgreindur liður, *Hvítbil*. Hann er einnig ótilgreindur á eftir öllum lesföstum (auðkenndir með '...' eða [...]) hér eftir.

<i>Hvítbil</i>	←	([\ t \ n \ r] / <i>Athugasemd</i>)*
<i>Athugasemd</i>	←	' ; ; ' ([\ n \ r] .)* [\ n \ r]

Næsta tákn, *Forrit*, er byrjunartákn málfræðinnar. Tengingar og forgangur virkja og aðgerða ekki tiltekin í mállýsingunni heldur í lok kaflans.

<i>Forrit</i>	←	(<i>Veiting</i> ' ; ') * !.
<i>Veiting</i>	←	<i>Strengur</i> '<' <i>Nafn</i> <i>Eining</i> / <i>Strengur</i> '= ' <i>Eining</i> / <i>Nafn</i> '= ' <i>Eining</i>
<i>Eining</i>	←	<i>Grunneining</i> ([& : + *] <i>Eining</i>)? / ' ! ' <i>Eining</i>
<i>Grunneining</i>	←	' (' <i>Eining</i> ')' / <i>Strengur</i> / <i>Nafn</i> / ' { ' <i>Vörpun</i> * ' }'
<i>Vörpun</i>	←	(<i>Nafn</i> / <i>Aðgerð</i>) ' - > ' <i>Minnissvæði</i>

<i>Minnissvæði</i>	←	<i>Nafn</i> / <i>Aðgerð</i> / 'breyta' / <i>Stefskilgreining</i>
<i>Stefskilgreining</i>	←	'stef' '(' <i>Nafnaruna?</i> ';' <i>Nafnaruna?</i> ')' <i>Skilgreiningar</i> 'stofn' <i>Segðaruna</i> 'stofnlok' <i>InnriTextaeining?</i>
<i>Nafnaruna</i>	←	<i>Nafn</i> (',' <i>Nafn</i>)*
<i>Skilgreiningar</i>	←	('innflutt' <i>Nafnaruna</i> / 'staðvær' <i>Frumstillingaruna</i>)*
<i>Frumstillingaruna</i>	←	<i>Frumstilling</i> (',' <i>Frumstilling</i>)*
<i>Frumstilling</i>	←	<i>Nafn</i> (':'= <i>Segð</i>)?
<i>Segðaruna</i>	←	(<i>Segð</i> (',' <i>Segð</i>)*)?
<i>InnriTextaeining</i>	←	'{' (<i>InnraStef</i>)* '}'
<i>InnraStef</i>	←	<i>Nafn</i> '->' <i>Stefskilgreining</i>
<i>Segð</i>	←	<i>Aðgerðarsegð</i> (('og' / 'eða') <i>Segð</i>)? / 'ekki' <i>Segð</i>
<i>Aðgerðarsegð</i>	←	<i>Fylkissegð</i> (<i>Aðgerð</i> <i>Fylkissegð</i>)*
<i>Fylkissegð</i>	←	<i>Smásegð</i> '[' <i>Segðaruna</i> ']' / <i>Smásegð</i>
<i>Smásegð</i>	←	<i>Efsegð</i> / '[' <i>Segðaruna</i> ']' / <i>Lykkjusegð</i> / 'stofn' <i>Segðaruna</i> 'stofnlok' / 'út' / 'skila' <i>Segð</i> / '(' <i>Segð</i> ')' / <i>Nafnsegð</i> / <i>Aðgerð</i> '(' ';' <i>Segðaruna</i> ')' / <i>Aðgerð</i> <i>Segð</i> / <i>Strengur</i> / <i>Staffasti</i> / <i>Fleytitala</i> / <i>Fjöldatala</i> / <i>Heitala</i>
<i>Nafnsegð</i>	←	<i>Nafn</i> ':'= <i>Segð</i> / <i>Nafn</i> '(' <i>Nafnaruna?</i> ';' <i>Segðaruna</i> ')' / <i>Nafn</i>

Lykkjusegð ← 'lykkja' *Segðaruna* 'lykkjulok'
 / 'meðan' *Segð* 'lykkja' *Segðaruna* 'lykkjulok'
 / 'fyrir' '(' *Segðaruna* ';' *Segð* ';' *Segðaruna* ')'
 'lykkja' *Segðaruna* 'lykkjulok'

Efsegð ← 'ef' *Segð* 'þá' *Segðaruna*
 ('annarsef' *Segð* 'þá' *Segðaruna*) *
 ('annars' *Segðaruna*) ?
 'eflok'

II.1 Forgangur og tenging

Einingaraðgerðirnar hafa allar sömu merkingu og sama forgang eins og í Fjölni. Forgangur aðgerða ákvarðast hins vegar af fyrsta staf aðgerðarnafnsins, eins og eftirfarandi tafla sýnir.

FYRSTI STAFUR AÐGERÐAR	FORGANGUR
^	7
!, ? eða @	6
*, /, % eða ?	5
+, - eða	4
<, >, = eða @	3
:	2
Aðrir aðgerðastafir	1

TAFLA II.2: FORGANGUR AÐGERÐA

Allar aðgerðir eru tengdar til vinstri, nema þær aðgerðir sem hafa : sem fyrsta staf, þær eru tengdar til hægri.

MERKINGARFRÆÐI

Merkingarfræði Baldurs verður best skýrð með tilvísunum í forritunarmálið Fjölni. Allar venjulegar segðir, ásamt einingavinnslu, er nákvæmlega eins og í Fjölni og því vísast í viðeigandi efni og handbækur ([3] og [4]).

Þó að Baldur sé hannaður til að líkja sem mest eftir Fjölni með ákveðnum útvíkkunum, var óhjákvæmilega ekki hægt að útfæra öll atriði úr Fjölni. Ástæðan er fyrst og fremst að þýðandinn er unnin sem verkefni í einu námskeiði og því ekki nægur tími til né ábati af því að útfæra þau atriði sem krefjast fyrst og fremst handavinnu.

III.1 Atriði aðeins í Fjölni

Baldur er fyrst og fremst smíðaður út frá útgáfu 1 af Fjölni. Það þýðir að hlutbundin forritun er ekki möguleg (nema með sérstökum tilfæringum) í Baldri.

Í Fjölni er einnig að finna ákveðin gagnleg atriði sem ekki eru útfærð í Baldri. Ber þar helst að nefna valsegð og stefgildi. Að auki eru nokkur atriði í lesgreiningu og þáttun sem samræmast ekki Fjölni, t.d. er komma (,) ekki leyfð sem síðasti stafur í segðarunu í Baldri.

III.2 Ný atriði í Baldri

Baldur inniheldur tvær meginútvíkkunarir á virkni Fjölnis. Sú helsta er möguleiki á bálmótun, þ.e. notandanum er gert kleyft að skilgreina s.k. innri stef.

Innri stef eiga við eitt stef í forritinu, og er aðeins nýtanleg frá því stafi og öðrum innri stefjum þess. Í innri stefjum er hægt að vísa í breytur og viðföng stefja í efri földunarhæð.

Skilgreining innri stefja verður best sýnd með dæmi.

```

1 "forrit1" < aðalforrit
2 {
3   aðalforrit ->
4     stef(;)
5     staðvær x := 1, increment := "bull"
6     stofn
7       ;; Hér köllum við í innra stef til að hækka teljara
8       hækkateljara(;),
9       \skrifastreng "Teljari_eftir_upphækkun_er:_",
10      \skrifa x,
11      nýlína(;)
12   stofnlok
13   {
14     hækkateljara ->
15       stef(;)
16       staðvær increment := 1
17     stofn

```



```

18         x := x + increment
19     stofnlok
20 }
21 } * "GRUNNUR";

```

Það fyrsta sem við veitum athygli í þessu dæmi er að í línu 13 er opnuð s.k. *innri eining*. Innri eining getur innihaldið runu af innsetningum nafns og stefskilgreiningar. Í línu 14 bindum við svo stef við nafnið hækkateljara, en þessi binding er aðeins gild innan í stefinu aðalstef og í stefinu hækkateljara sjálfu.

Hinn punkturinn í þessu dæmi er að í línu 18 vísum við í breytuna *x*, jafnvel þó sú breyta sé ekki skilgreind í því stafi sem lína 18 tilheyrir. Á hinn bóginn skilgreina bæði stefin staðværa breytu með nafninu *increment*. Þegar faldaða stefið notar það nafn vísar það í þá breytu sem er næst í földunarhæð. Þannig felur breytan *increment* í stefinu hækkateljara breytuna með sama nafni í stefinu aðalforrit.

Skilgreina má innra stef í hvaða stafi sem er, óháð földunarhæð. Þannig má földunardýptin verða svo djúp sem hver vill.¹

Til að draga saman reglur um það hvar í forritstexta breytur og stefnöfn eru sýnileg, þá höfum við eftirfarandi.

Breytur og viðföng Þegar vísað er í breytu eða viðfang með nafni leitar þýðandinn fyrst í því stafi sem tilvísunin kemur fyrir. Ef nafnið er ekki þekkt sem staðvær breyta eða viðfang í því stafi er leitað í næstu földunarhæð fyrir ofan, þ.e. því stafi sem „á“ viðkomandi stef. Þannig er svo haldið áfram koll af kolli þar til skilgreining nafnsins finnst sem breyta eða viðfang. Ef nafnið er ekki skilgreint í neinu yfirstefi þess stefs sem tilvísunin kemur fyrir í fær notandinn villu við þýðingu forritsins.

Nöfn stefja Þegar kallað er í stef í forritstexta, er fyrst athugað hvort viðkomandi stefnafn er skilgreint sem undirstef þess stefs sem kallað er úr. Ef ekki er haldið áfram upp földunarkedjuna og athugað í næsta stafi fyrir ofan hvort nafnið er þekkt sem undirstef þar. Ef nafnið finnst hvergi sem undirstef er viðkomandi nafn merkt sem innflutningur í viðkomandi einingu og þannig höndlað sem venjulegt Fjölnisstef.

Óformlega má lýsa sýnileika stefnafna þannig að stef getur kallað í sjálft sig, öll börn foreldris síns, öll börn ömmu sinnar o.s.frv.

Þannig getur í eftirfarandi dæmi stef A kallað í stef B1, en ekki C. Stefið B2 getur kallað í stefin A, B1 og sjálft sig og stefin B1 og C geta kallað í öll stefin sem hér eru skilgreind.

```

A -> stef(;)
    stofn

```

¹ Að vísu verður raunverulegt nafn stefs lengra eftir því sem földunardýpt þess er meiri, og tengir2.exe ræður aðeins við endanlega löng nöfn.

```

...
stofnlok
{
B1 -> stef(;)
    stofn
    ...
    stofnlok
    {
        C -> stef(;)
        stofn
        ...
        stofnlok
    }
B2 -> stef(;)
    stofn
    ...
    stofnlok
}

```

Annað atriði sem er frábrugðið frá Fjölni er hvernig halaköll er meðhöndluð. Í Baldri er virkri vakningarfærslu stefs hent rétt fyrir síðasta kall í stefinu, svo framarlega sem eftirfarandi skilyrði séu uppfyllt:

1. Stefið sem kallað er í er ekki lægra í földunarhæð en stefið sem kallað er úr, og því er ekki hætt á að vísað verði í breytur eða viðföng þess síðarnefnda.
2. Hvorki stefið sem kallað er í eða sem kallað er úr nota afritsviðföng.

Þetta þýðir að við getum nýtt okkur halaendurkvæmni, hvort sem er beina eða óbeina, til að skrifa endurkvæm föll sem haga sér eins og lykkjur:

```

1 "forrit2" < aðalforrit
2 {
3   aðalforrit ->
4     stef(;)
5     staðvær n
6     stofn(;)
7     \skrifastreng "Sláðu_inn_jákvæða_heiltölu:_",
8     n := lesa(;),
9     \skrifa n,
10    \skrifastreng "!_er_",
11    \skrifa hröp(;n,1),
12    nýlína(;)
13   stofnlok
14   {
15     hröp ->
16       stef(;n, uppsafnað)
17       stofn
18       ef n < 2 þá
19         uppsafnað
20         annars
21         hröp(;n-1, n*uppsafnað)
22       eflok
23     stofnlok
24   }
25 } * "GRUNNUR";

```

Í línum 15-23 er skilgreint stef sem beitir halaendurkvæmni til þess að reikna út $n!$. Það síðasta sem þetta stef gerir þegar breytan n er stærri eða jöfn 2 er að

kalla í sjálf sig (lína 21). Við það er núverandi vakningarfærslu stefsins eytt áður en farið er inn í kallið að nýju. Seinni vakning stefsins skilar gildi sínu svo beint til þess er kallaði í fyrri vakninguna.

Hér verður ekki farið nánar út í kosti, galla eða notkun halaendurkvæmni og bestun halakalla. Hvað viðkemur Baldri er aðeins mikilvægt að notandinn geri sér grein fyrir því hvaða skilyrði þurfa að vera uppfyllt svo hægt sé að beita slíkri bestun og með því móti getur hann skrifað stef sem nýta sér þennan eiginleika.

KAFLI IV

DÆMI

Best er að kynnst nýju forritunarmáli með dæmum, við skulum líta á nokkur.

Halaendurkvæmni án földunar Halaendurkvæmni virkar í öllum föllum, óháð földunardýpt þeirra. Eftirfarandi eining flytur út stef sem finnur stærsta stak í lista m.t.t. innfluttu samanburðaraðgerðarinnar >.

```
1 "maximal" =
2 !{
3 stærst ->
4   ;; Notkun: m := stærst(l)
5   ;; Fyrir: l er listi löglegra viðfanga í >
6   ;; Eftir: m er stak úr l þ.a. m > x eða m = x f.öll x úr l
7   stef(l)
8   stofn
9   ef ekki \hali l þá
10      ;; Aðeins er eitt stak eftir svo það er stærst
11      \haus l
12      annarsef \haus l > \haus \hali l þá
13      ;; Stak nr. 1 í l er stærra en stak nr. 2, hendum nr. 2
14      stærst(\haus l : \hali \hali l)
15      annars
16      ;; Hendum staki nr. 1
17      stærst(\hali l)
18      eflok
19      stofnlök
20 };
```

Snúningur lista Eftirfarandi forrit beitir halaendurkvæmni og földun til þess að útfæra fall sem snýr við gefnum lista.

```
1 "snua" =
2 {
3 snúa ->
4   ;; Notkun: l2 := snúa(l1)
5   ;; Fyrir: l1 er listinn [x1,x2,...,xN]
6   ;; Eftir: l2 er listinn [xN,...,x2,x1]
7   stef(l)
8   stofn
9   snúaskeyta(l,[])
10  stofnlök
11  {
12  snúaskeyta ->
13    ;; Notkun: l := snúaskeyta(x,y)
14    ;; Fyrir: x er listinn [x1,x2,...,xN]
15    ;;        y er listinn [y1,y2,...,yM]
16    ;; Eftir: l er listinn [xN,...,x2,x1,y1,y2,...,yM]
17    stef(x,y)
18    stofn
19    ef x þá
20      snúaskeyta(\hali x, \haus x : y)
21    annars
```

```
22         y
23         eflok
24         stofnlok
25     }
26 };
```

NOTKUN

v.1 Uppsetning

Þýðandinn, ásamt fylgiskrá og hluta af Fjölnisþýðandanum, kemur í vöndlinum `baldur.zip`. Uppsetning felst aðeins í að afpakka öllum skráum úr vöndlinum í eitthvert skráarsafn, t.d. `c:\baldur`.

v.2 Keyrsla

Í öllum tilfellum er nauðsynlegt að vera staddur í því skráarsafni þar sem þýðandinn var settur upp.

Til að þýða einingar eða keyrsluskrár er hægt að nota skipanaskrána `baldur.bat`, en hana má keyra upp með einu viðfangi. Viðfangið er tekið sem nafn skrár, sem lesin er með ISO-8859-1 táknöflunni, þýdd og keyrð gegnum smala og Fjölnis-tenginn.

Þýðandinn sjálfur heitir hins vegar `falskur_fjolnir` (sem er vinnuheiti Baldurs). Hann er unnt að keyra með nokkrum rofum sem stýra virkni hans. Þeir eru `l`, `p`, `d`, `i so`, `n` og `o` og merking hvers þeirra er eftirfarandi:

-`l` Lesgreinirinn er keyrður á inntakið og úttakið er listi yfir les og tók hvers staks í táknastrumnum. Eftir lesgreiningu hættir þýðandinn keyrslu.

-`p` Úttak þýðandans verður máltré (e. *syntax tree*) fyrir inntakið á LISP/Scheme formi. Eftir þáttun hættir þýðandinn. Tvö máltré verða skrifuð, eitt fyrir umbreytingarfasa og annað eftir þann fasa. Formi úttaksins má breyta með `d-rofanum`.

-`d1` eða -`d2` Úttak þýðandans verður máltré fyrir inntakið á formi sem hentar sem inntak í forritið `dot`¹. Talan í rofanum segir til um hvort prenta eigi máltré fyrir umbreytingarfasa (-`d1`) eða eftir (-`d2`).

-`i so` Ef þessi rofi er gefinn upp skrifar þýðandinn úttak sitt eftir ISO-8859-1 táknöflunni. Sjálfgefið er að skrifa úttak í CP-861. Villuboð (*stderr*) eru alltaf skrifuð í CP-861.

¹`dot` er forrit sem teiknar tré og net. Það er hluti af *graphviz* pakkanum sem má nálgast á <http://www.research.att.com/sw/tools/graphviz/>

-n Þessi rofi gefur til kynna að inntak þýðandans sé í CP-861. Sjálfgefið er að lesa inntakið sem ISO-8859-1.

-o *skrú* Með þessum rofa má beina úttaki þýðandans (öðru en villum) í tilgreinda skrá. Skráin verður yfirskrifuð ef hún er þegar til. Sjálfgefið er að skrifa úttak í *stdout*.

Ef þýðandinn finnur streng á skipanalínunni sem byrjar ekki á bandstriki (-), þá er sá strengur túlkaður sem nafn skrár sem skal lesa inntakið úr. Aðeins er hægt að lesa úr einni skrá í einu, og séu margir slíkir strengir er sá síðasti notaður. Sé enginn slíkur strengur á skipanalínunni er sjálfgefin hegðun að lesa úr *stdin*.

HEIMILDIR

- [1] Ford, Bryan. „Parsing Expression Grammars: A Recognition-Based Syntactic Foundation“ *Symposium on Principles of Programming Languages*, 14.-16. janúar 2004, Feneyjum, Ítalíu
- [2] Aho, Sethi, Ullman. *Compilers – Principle, Techniques, and Tools*, Addison-Wesley, 1986.
- [3] Snorri Agnarsson, Páll Björnsson, Jón Harðarson. *Forritunarmálið Fjölnir, Notendahandbók* Reykjavík, Snorri Agnarsson, 2001.
- [4] Snorri Agnarsson. *08.71.33 Forritunarmál, Vikublað 13b*, Háskóli Íslands, 2003.