

## **Universidade do Minho** Escola de Engenharia

# Cálculo de Programas

Trabalho Prático (2024/25)

Lic. em Ciências da Computação

# Grupo G15

a108473 André Filipe Dourado Pinheiroa105532 Killian Alexandre Ferreira Oliveiraa108398 Pedro Dong Mo

### Preâmbulo

Na UC de Cálculo de Programas pretende-se ensinar a programação de computadores como uma disciplina científica. Para isso parte-se de um repertório de *combinadores* que formam uma álgebra da programação (conjunto de leis universais e seus corolários) e usam-se esses combinadores para construir programas *composicionalmente*, isto é, agregando programas já existentes.

Na sequência pedagógica dos planos de estudo dos cursos que têm esta disciplina, opta-se pela aplicação deste método à programação em Haskell (sem prejuízo da sua aplicação a outras linguagens funcionais). Assim, o presente trabalho prático coloca os alunos perante problemas concretos que deverão ser implementados em Haskell. Há ainda um outro objectivo: o de ensinar a documentar programas, a validá-los e a produzir textos técnico-científicos de qualidade.

Antes de abordarem os problemas propostos no trabalho, os grupos devem ler com atenção o anexo A onde encontrarão as instruções relativas ao *software* a instalar, etc.

Valoriza-se a escrita de *pouco* código que corresponda a soluções simples e elegantes que utilizem os combinadores de ordem superior estudados na disciplina.

### Problema 1

Esta questão aborda um problema que é conhecido pela designação 'Container With Most Water' e que se formula facilmente através do exemplo da figura seguinte:



A figura mostra a sequência de números

$$hghts = [1, 8, 6, 2, 5, 4, 8, 3, 7]$$

representada sob a forma de um histograma. O que se pretende é obter a maior área rectangular delimitada por duas barras do histograma, área essa marcada a azul na figura. (A "metáfora" *container with most water* sugere que as barras selecionadas delimitam um *container* com água.)

Pretende-se definida como um catamorfismo, anamorfismo ou hilomorfismo uma função em Haskell

$$mostwater :: [Int] \rightarrow Int$$

que deverá dar essa área. (No exemplo acima tem-se mostwater [1,8,6,2,5,4,8,3,7]=49.) A resolução desta questão deverá ser acompanhada de diagramas elucidativos.



Figure 1: RNN vista como instância de um accumulating map [3].

#### Problema 2

Um dos problemas prementes da Computação na actualidade é conseguir, por engenharia reversa, interpretar as redes neuronais (RN) geradas artificialmente sob a forma de algoritmos compreensíveis por humanos.

Já foram dados passos que, nesse sentido, explicam vários padrões de RNs em termos de combinadores funcionais [3]. Em particular, já se mostrou como as RNNs (*Recurrent Neural Networks*) podem ser vistas como instâncias de *accumulating maps*, que em Haskell correspondem às funções *mapAccumR* e *mapAccumL*, conforme o sentido em que a acumulação se verifica (cf. figura 1).

A RNN que a figura 1 mostra diz-se 'one-to-one' [1]. Há contudo padrões de RNNs mais gerais: por exemplo, o padrão 'many-to-one' que se mostra na figura 2 extraída de [1].

Se *mapAccumR* e *mapAccumL* juntam *maps* com *folds*, pretendemos agora combinadores que a isso acrescentem *filter*, por forma a selecionar que etapas da computação geram ou não *outputs* – obtendo-se assim o efeito *'many-to-one'*. Ter-se-á, para esse efeito:

```
\begin{aligned} \mathit{mapAccumRfilter} &:: ((a,s) \rightarrow \mathit{Bool}) \rightarrow ((a,s) \rightarrow (c,s)) \rightarrow ([a],s) \rightarrow ([c],s) \\ \mathit{mapAccumLfilter} &:: ((a,s) \rightarrow \mathit{Bool}) \rightarrow ((a,s) \rightarrow (c,s)) \rightarrow ([a],s) \rightarrow ([c],s) \end{aligned}
```

Pretendem-se as implementações de *mapAccumRfilter* e *mapAccumLfilter* sob a forma de ana / cata ou hilomorfismos em Haskell, acompanhadas por diagramas.

Como caso de uso, sugere-se o que se dá no anexo E que, inspirado em [1], recorre à biblioteca Data. Matrix.

## Problema 3

A fornecer na segunda edição deste enunciado

## Problema 4

A fornecer na segunda edição deste enunciado

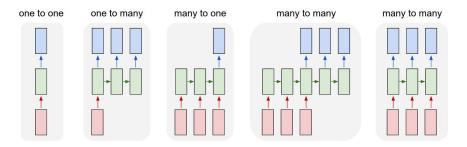


Figure 2: Várias tipologias de RNNs [1].

#### **Anexos**

#### A Natureza do trabalho a realizar

Este trabalho teórico-prático deve ser realizado por grupos de 3 alunos. Os detalhes da avaliação (datas para submissão do relatório e sua defesa oral) são os que forem publicados na página da disciplina na *internet*.

Recomenda-se uma abordagem participativa dos membros do grupo em **todos** os exercícios do trabalho, para assim poderem responder a qualquer questão colocada na *defesa oral* do relatório.

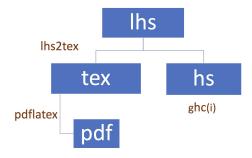
Para cumprir de forma integrada os objectivos do trabalho vamos recorrer a uma técnica de programação dita "literária" [2], cujo princípio base é o seguinte:

Um programa e a sua documentação devem coincidir.

Por outras palavras, o **código fonte** e a **documentação** de um programa deverão estar no mesmo ficheiro.

O ficheiro cp2425t.pdf que está a ler é já um exemplo de programação literária: foi gerado a partir do texto fonte cp2425t.lhs¹ que encontrará no material pedagógico desta disciplina descompactando o ficheiro cp2425t.zip.

Como se mostra no esquema abaixo, de um único ficheiro (*lhs*) gera-se um PDF ou faz-se a interpretação do código Haskell que ele inclui:



Vê-se assim que, para além do GHCi, serão necessários os executáveis pdflatex e lhs2TeX. Para facilitar a instalação e evitar problemas de versões e conflitos com sistemas operativos, é recomendado

<sup>&</sup>lt;sup>1</sup> O sufixo 'lhs' quer dizer *literate Haskell*.

o uso do Docker tal como a seguir se descreve.

#### **B** Docker

Recomenda-se o uso do container cuja imagem é gerada pelo Docker a partir do ficheiro Dockerfile que se encontra na diretoria que resulta de descompactar cp2425t.zip. Este container deverá ser usado na execução do GHCi e dos comandos relativos ao LTEX. (Ver também a Makefile que é disponibilizada.)

Após instalar o Docker e descarregar o referido zip com o código fonte do trabalho, basta executar os seguintes comandos:

```
$ docker build -t cp2425t .
$ docker run -v ${PWD}:/cp2425t -it cp2425t
```

**NB**: O objetivo é que o container seja usado *apenas* para executar o GHCi e os comandos relativos ao LaTeX. Deste modo, é criado um *volume* (cf. a opção -v \${PWD}:/cp2425t) que permite que a diretoria em que se encontra na sua máquina local e a diretoria /cp2425t no container sejam partilhadas.

Pretende-se então que visualize/edite os ficheiros na sua máquina local e que os compile no container, executando:

```
$ lhs2TeX cp2425t.lhs > cp2425t.tex
$ pdflatex cp2425t
```

lhs2TeX é o pre-processador que faz "pretty printing" de código Haskell em La que faz parte já do container. Alternativamente, basta executar

\$ make

para obter o mesmo efeito que acima.

Por outro lado, o mesmo ficheiro cp2425t.lhs é executável e contém o "kit" básico, escrito em Haskell, para realizar o trabalho. Basta executar

```
$ qhci cp2425t.lhs
```

Abra o ficheiro cp2425t.lhs no seu editor de texto preferido e verifique que assim é: todo o texto que se encontra dentro do ambiente

```
\begin{code}
...
\end{code}
```

é seleccionado pelo GHCi para ser executado.

## C Em que consiste o TP

Em que consiste, então, o *relatório* a que se referiu acima? É a edição do texto que está a ser lido, preenchendo o anexo F com as respostas. O relatório deverá conter ainda a identificação dos membros do grupo de trabalho, no local respectivo da folha de rosto.

Para gerar o PDF integral do relatório deve-se ainda correr os comando seguintes, que actualizam a bibliografia (com BibT<sub>F</sub>X) e o índice remissivo (com makeindex),

```
$ bibtex cp2425t.aux
```

\$ makeindex cp2425t.idx

e recompilar o texto como acima se indicou. (Como já se disse, pode fazê-lo correndo simplesmente make no container.)

No anexo E disponibiliza-se algum código Haskell relativo aos problemas que são colocados. Esse anexo deverá ser consultado e analisado à medida que isso for necessário.

Deve ser feito uso da programação literária para documentar bem o código que se desenvolver, em particular fazendo diagramas explicativos do que foi feito e tal como se explica no anexo D que se seque.

## D Como exprimir cálculos e diagramas em LaTeX/lhs2TeX

Como primeiro exemplo, estudar o texto fonte (lhs) do que está a ler<sup>1</sup> onde se obtém o efeito seguinte:<sup>2</sup>

$$id = \langle f,g \rangle \\ \equiv \qquad \{ \text{ universal property } \} \\ \left\{ \begin{array}{l} \pi_1 \cdot id = f \\ \pi_2 \cdot id = g \end{array} \right. \\ \equiv \qquad \{ \text{ identity } \} \\ \left\{ \begin{array}{l} \pi_1 = f \\ \pi_2 = g \end{array} \right. \\ \Box$$

Os diagramas podem ser produzidos recorrendo à package xymatrix, por exemplo:

## E Código fornecido

Teste relativo à figura da página 1:

$$test_1 = mostwater\ hghts$$

#### Problema 2

Testes relativos a mapAccumLfilter e mapAccumRfilter em geral (comparar os outputs)

$$test_{2a} = mapAccumLfilter ((>10) \cdot \pi_1) f (odds 12, 0)$$
  
 $test_{2b} = mapAccumRfilter ((>10) \cdot \pi_1) f (odds 12, 0)$ 

<sup>1</sup> Procure e.g. por "sec:diagramas".

<sup>&</sup>lt;sup>2</sup> Exemplos tirados de [4].

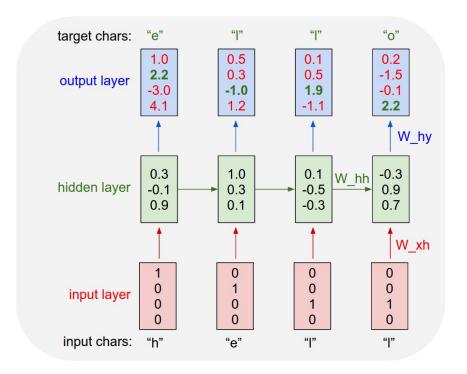


Figure 3: Exemplo *char seq* extraído de [1].

onde:

odds 
$$n = \text{map } ((1+) \cdot (2*)) [0..n-1]$$
  
 $f(a,s) = (s,a+s)$ 

Teste

$$test_{2c} = mapAccumLfilter true step ([x_1, x_2, x_3, x_4], h_0)$$

baseado no exemplo de Karpathy [1] que a figura 3 mostra, usando os dados seguintes:

• Estado inicial:

$$h_0 = fromList \ 3 \ 1 \ [1.0, 1.0, 1, 0]$$

• Step function:

$$step (x,h) = (\alpha (wy * h), \alpha (wh * h + wx * x))$$

• Função de activação:

$$\alpha = \text{fmap } \sigma \text{ where } \sigma x = (tanh x + 1) / 2$$

• Input layer:

$$\begin{aligned} &inp = [x_1, x_2, x_3, x_4] \\ &x_1 = fromList \ 4 \ 1 \ [1.0, 0, 0, 0] \\ &x_2 = fromList \ 4 \ 1 \ [0, 1.0, 0, 0] \\ &x_3 = fromList \ 4 \ 1 \ [0, 0, 1.0, 0] \\ &x_4 = x_3 \end{aligned}$$

• Matrizes exemplo:

```
 \begin{array}{l} \textit{wh} = \textit{fromList} \ 3\ 3\ [0.4, -0.2, 1.6, -3.1, 1.4, 0.1, 5.4, -2.7, 0.1] \\ \textit{wy} = \textit{fromList} \ 4\ 3\ [2.1, 1.1, 0.8, 1.3, -6.4, -3.4, -2.7, -3.8, -1.3, -0.5, -0.9, -0.4] \\ \textit{wx} = \textit{fromLists} \ [[0.0, -51.9, 0.0, 0.0], [0.0, 26.6, 0.0, 0.0], [-16.7, -5.5, -0.1, 0.1]] \\ \end{array}
```

**NB**: Podem ser definidos e usados outros dados em função das experiências que se queiram fazer.

## F Soluções dos alunos

Os alunos devem colocar neste anexo as suas soluções para os exercícios propostos, de acordo com o "layout" que se fornece. Não podem ser alterados os nomes ou tipos das funções dadas, mas pode ser adicionado texto ao anexo, bem como diagramas e/ou outras funções auxiliares que sejam necessárias.

**Importante**: Não pode ser alterado o texto deste ficheiro fora deste anexo.

#### Problema 1

Vamos utilizar a estratégia de *divide and conquer*, isto é, vamos construir um anamorfismo que devolve uma lista com as maiores áreas possíveis e um catamorfismo para obter a maior dessas áreas.

Para fazer isto de forma otimal, começemos por calcular a área com base no primeiro e último elemento de uma lista não vazia *l*.

Sendo assim, o cálculo da área é dado pela função:

```
area \ l = (min \ (head \ l) \ (last \ l)) * ((length \ l) - 1)
```

A seguir, descartamos o menor elemento entre primeiro e último elemento da lista, de forma a encontrar um valor maior. Para isso temos dois casos:

- 1. Se head  $l \geqslant \text{last } l$ , então isto significa que calculámos a água armazenada para o recipiente de altura lastl, por isso descartamos o lastl.
- 2. Se head l < last l, então isto significa que calculámos a água armazenada para o recipiente de altura head l, por isso descartamos o head l.

Recursivamente, isto pode ser escrito como

```
 \begin{cases} \textit{bestAreas} \ [\ ] = [\ ] \\ \textit{bestAreas} \ l = (\textit{area} \ l) : (\textbf{if} \ \textit{head} \ l < \textit{last} \ l \ \textbf{then} \ \textit{bestAreas} \ (\textit{tail} \ l) \ \textbf{else} \ \textit{bestAreas} \ (\textit{init} \ l)) \end{cases}
```

Tornando esta definição point-free, temos

```
 \begin{cases} \textit{bestAreas} \ [\ ] = [\ ] \\ \textit{bestAreas} \ l = i_2 \ (\textit{area} \ l) : (\textbf{if} \ \textit{head} \ l < \textit{last} \ l \ \textbf{then} \ \textit{bestAreas} \ (\textit{tail} \ l) \ \textbf{else} \ \textit{bestAreas} \ (\textit{init} \ l)) \end{cases}   \equiv \qquad \{ \ \textit{definição} \ \textit{pointwise}, \ 72, \ 73 \ \}   \begin{cases} \textit{bestAreas} \cdot \textit{nil} = \textit{nil} \\ \textit{bestAreas} \cdot \textit{id} = \textit{cons} \cdot \langle \textit{area}, \textit{cond} \ \widehat{((<)} \cdot \langle \textit{head}, \textit{last} \rangle) \ (\textit{bestAreas} \cdot \textit{tail}) \ (\textit{bestAreas} \cdot \textit{init}) \rangle \end{cases}
```

Portanto, o *divide* desse anamorfismo é  $(id + \langle area, cond \ \widehat{(<)} \cdot \langle head, last \rangle) \ tail \ init \rangle) \cdot out$ . Resultando assim no diagrama do anamorfismo

$$\begin{array}{c} \mathbb{N}_0^* \xrightarrow{divide} 1 + \mathbb{N}_0 \times \mathbb{N}_0^* \\ \textit{bestAreas} = [(\textit{divide})] \Big | \qquad \qquad & |\textit{id} + \textit{id} \times \textit{bestAreas} \\ \mathbb{N}_0^* \xleftarrow{\quad \quad } 1 + \mathbb{N}_0 \times \mathbb{N}_0^* \end{array}$$

Além disso, já vimos nas aulas que o catamorfismo de listas

$$maxList = ([zero, \widehat{max}])$$

em que  $conquer = [zero, \widehat{max}]$ , é responsável por obter o maior valor numa lista de números não negativos, podendo ser representado pelo seguinte diagrama

$$\begin{array}{c} \mathbb{N_0}^* \xrightarrow{out} 1 + \mathbb{N_0} \times \mathbb{N_0}^* \\ \mathit{maxList} = ( \mathit{conquer} ) \bigvee_{} & \bigvee_{} \mathit{id} + (\mathit{id} \times \mathit{maxList}) \\ \mathbb{N_0} \xleftarrow{conquer} 1 + \mathbb{N_0} \times \mathbb{N_0} \end{array}$$

Para evitar conflito entre os tipos Int e Integer, modificamos o conquer para

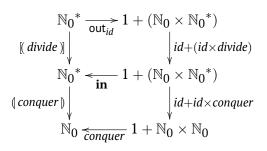
$$conquer = [zero, \widehat{max} \cdot (fromIntegral \times fromIntegral)]$$

Portanto, temos

$$\begin{aligned} \textit{mostwater} &= [\![ \textit{conquer}, \textit{divide} ]\!] \\ \textit{divide} &= (\textit{id} + \langle \textit{area}, \textit{cond} \ (\widehat{(<)} \cdot \langle \textit{head}, \textit{last} \rangle) \ \textit{tail init} \rangle) \cdot \mathsf{out}_{\textit{id}} \end{aligned}$$

```
\begin{split} & conquer = [zero, \widehat{max} \cdot (fromIntegral \times fromIntegral)] \\ & area \ l = (min \ (head \ l) \ (last \ l)) * ((length \ l) - 1) \\ & \mathsf{out}_{id} \ [\,] = i_1 \ () \\ & \mathsf{out}_{id} \ (h:t) = i_2 \ (h:t) \end{split}
```

Com diagrama final



#### Problema 2

 $mapAccumRfilter\ p\ f = \bot$  $mapAccumLfilter\ p\ f = \bot$ 

## Index

```
₽T<sub>E</sub>X, 4
    bibtex, 4
    lhs2TeX, 3-5
    makeindex, 4
    pdflatex, 3
    xymatrix, 5
Combinador "pointfree"
    cata
      Naturais, 5
    mapAccumL, 2
    mapAccumR, 2
    split, 5
Cálculo de Programas, 1, 3
    Material Pedagógico, 3
Docker, 4
    container, 4, 5
Função
    \pi_1, 5
    \pi_2, 5
    map, 6
Haskell, 1-5
    Hackage
      Data.Matrix, 2
    interpretador
      GHCi, 3, 4
    Literate Haskell, 3
Números naturais (N), 5
Programação
    literária, 3, 5
Rede neuronal, 2
    RNN, 2, 3
```

#### References

- [1] A. Karpathy. The unreasonable effectiveness of recurrent neural networks, 2015. Blog: http://karpathy.github.io/2015/05/21/rnn-effectiveness, last read: June 7, 2025.
- [2] D.E. Knuth. *Literate Programming*. CSLI Lecture Notes Number 27. Stanford University Center for the Study of Language and Information, Stanford, CA, USA, 1992.
- [3] C. Olah. Neural networks, types, and functional programming, 2015. Blog: http://colah.github.io/posts/2015-09-NN-Types-FP/, last read: June 7, 2025.
- [4] J.N. Oliveira. Program Design by Calculation, 2024. Draft of textbook in preparation. First version: 1998. Current version: Sep. 2024. Informatics Department, University of Minho (pdf).