# Memory access

Luca Blum

May 2020

## 1 Introduction

We evaluated the memory accesses for the stable baum-welch algorithm. The reading accesses $M_{read}$ and writing accesses $M_{write}$ are separated. The timing consists following steps:

- forward step
- backward step
- update step
- finishing criteria

We define N as the number of hidden states, T as the number of time steps or equivalent as the number of observations made and K as the number of observables.

## 2 Standard Version

### 2.1 forward step

- init ct[0]

  $$M_1(n) = 1 * M_{write}$$

- compute $\alpha[0]$ and scaling factor ct[0]

  $$M_2(n) = (3 * N + 1) * M_{read} + (2 * N + 1) * M_{writes}$$

  Can easily be reduced by scalar replacement

- scale alpha[0]

  $$M_3(n) = N * M_{read} + N * M_{write}$$

  Can also be reduced to $N * C_{write}$

- for T timesteps:

- init ct[t]

$$M_4(n) = 1 * M_{write}$$

- compute $\alpha[s,t]$

$$M_5(n) = (2 * N * N + N) * M_{read} + (N * N + 2 * N) * M_{write}$$

- computing scaling factor ct[t]

$$M_6(n) = (N + 1) * M_{read} + (N + 1) * M_{write}$$

- scaling alpha[s,t]

$$M_7(n) = N * M_{read} + N * M_{write}$$

Overall

$$\begin{aligned}
M_{forward}(n) =&(3 * N + 1 + N + 2 * N * N * T + N * T + T + N * T) * M_{read} \\
&+ (1 + 2 * N + 1 + N + T + N * N * T + 2 * N * T + N * T + T + N * T) * M_{write} \\
=&(4 * N + 1 + 2 * N * N * T + 2 * N * T + T) * M_{read} \\
&+ (2 + 3 * N + 2 * T + N * N * T + 4 * N * T) * M_{write}
\end{aligned}$$

## 2.2 backward step

- scale $\beta[T-1]$

$$M_1 = N * M_{read} + N * Mwrite$$

- for T-1 timesteps:
  - init $\beta[s,t]$

  $$M_2 = N * M_{write}$$

  - compute $\beta[s,t]$

  $$M_3(n) = 3 * N * N * M_{read} + N * N * M_{write}$$

  - scaling $\beta[s,t]$

  $$M_4(n) = N * M_{read} + N * M_{write}$$

Overall

$$\begin{aligned}
M_{backward}(n) =&(N + 3 * N * N * (T - 1) + N * (T - 1)) * M_{read} \\
&+ (N + N * (T - 1) + N * N * (T - 1) + N * (T - 1)) * M_{write} \\
=&(N + 3 * N * N * (T - 1) + N * (T - 1)) * M_{read} \\
&+ (N + 2 * N * (T - 1) + N * N * (T - 1)) * M_{write}
\end{aligned}$$

## 2.3 update step

- compute gamma

$$M_1(n) = 2 * T * N * M_{read} + T * N * M_{write}$$

- compute xi

$$M_2(n) = 4 * (T - 1) * N * N * M_{read} + (T - 1) * N * N * M_{write}$$

- compute new pi

$$M_3(n) = 2 * N * M_{read} + N * M_{write}$$

- compute gamma sum denominator

$$M_4(n) = 2 * T * N * N * M_{read}$$

- new transition matrix

$$M_5(n) = N * N * M_{write}$$

- compute gamma sum numerator

$$M_6(n) = 2 * T * K * N * M_{read}$$

- new emission matrix

$$M_7(n) = N * K * M_{write}$$

Overall

$$\begin{aligned} M_{update} =& (2 * T * N + 4 * (T - 1) * N * N + 2 * N + 2 * T * N * N + 2 * T * K * N) * M_{read} \\ &+ (T * N + (T - 1) * N * N + N + N * N + N * K) * M_{write} \end{aligned}$$

## 2.4 finishing criteria

- for every timestep

$$M_1(n) = 1 * M_{read}$$

Overall

$$M_{finished}(n) = T * M_{read}$$

### 2.5 Total

$$
\begin{aligned}
M_{stable}(n) =& M_{forward}(n) + M_{backward}(n) + M_{update}(n) + M_{finished}(n) \\
M_{stable}(n) =& (4 * N + 1 + 2 * N * N * T + 2 * N * T + T) * M_{read} \\
& + (2 + 3 * N + 2 * T + N * N * T + 4 * N * T) * M_{write} \\
& + (N + 3 * N * N * (T - 1) + N * (T - 1)) * M_{read} \\
& + (N + 2 * N * (T - 1) + N * N * (T - 1)) * M_{write} \\
& + (2 * T * N + 4 * (T - 1) * N * N + 2 * N + 2 * T * N * N + 2 * T * K * N) * M_{read} \\
& + (T * N + (T - 1) * N * N + N + N * N + N * K) * M_{write} \\
& + T * M_{read} \\
=& (7 * N + 1 + 4 * N * N * T + 4 * N * T + 2 * T + 7 * N * N * (T - 1) \\
& + N * (T - 1) + 2 * T * K * N) * M_{read} \\
& + (2 + 5 * N + 2 * T + N * N * T + 5 * N * T + 2 * N * (T - 1) \\
& + 2 * N * N * (T - 1) + N * N + N * K) * M_{write} \\
& (7 * N + 1 + 4 * N * N * T + 4 * N * T + 2 * T + 7 * N * N * (T - 1) \\
& + N * (T - 1) + 2 * T * K * N) \\
& + (2 + 5 * N + 2 * T + N * N * T + 5 * N * T + 2 * N * (T - 1) \\
& + 2 * N * N * (T - 1) + N * N + N * K)
\end{aligned}
$$

If we count both reads and writes as one operation we get a total of

$$
\begin{aligned}
M_{total}(n) =& 7 * N + 1 + 4 * N * N * T + 4 * N * T + 2 * T + 7 * N * N * (T - 1) + N * (T - 1) \\
& + 2 * T * K * N + 2 + 5 * N + 2 * T + N * N * T + 5 * N * T + 2 * N * (T - 1) \\
& + 2 * N * N * (T - 1) + N * N + N * K \\
=& 2KNT + KN + 14N^2T - 8N^2 + 12NT + 9N + 4T + 3
\end{aligned}
$$

Asymptotically:

$$M_{total}(n) = 2KNT + 14N^2T$$

### 2.6 Working set

From cop one iteration of the loop at line 449 (new pi):

$$TN + 4T + N + K$$

## 3 Reordered

### 3.1 Pre-Forward

- add remaining parts

$$M_1(n) = N * 2 * M_r + N * 3 * M_w$$

- new emission matrix

$$M_2(n) = K * N * 2 * M_r + K * N * 2 * M_w$$

$$M_{pre} = M_1 + M_2 = (N * 2 + KN * 2) * M_r + N * (N3 + KN2) * M_w$$

## 3.2 Forward

- **transpose a-new**

$$M_0(n) = N^2 * M_r + N^2 * Mw$$

- alpha(0)

$$M_1(n) = N * 2 * M_r + N * M_w$$

- scale alpha(0)

$$M_2(n) = N * M_r + N * M_w$$

- alpha(1) inner loop

$$M_3(n) = (N - 1) * N * 3 * M_r + (N - 1) * N * 2 * M_w$$

- alpha(1) outer loop

$$M_4(n) = (N - 1) * M_r + (N - 1) * M_w$$

- gamma-sum to zero

$$M_{4.5}(n) = 3N * M_r + 3N * M_w$$

- scale alpha(t)

$$M_5(n) = N * M_r + N * M_w$$

- T, outer loop

$$M_6(n) = (T - 2) * M_r + (T - 2) * M_w$$

- T, scale loop

$$M_7(n) = (T - 2) * N * M_r + (T - 2) * N * M_w$$

- T, most inner loop

$$M_8(n) = (T - 2) * N * N * 2 * M_r$$

- T, write alpha loop

$$M_9(n) = (T - 2) * N * M_r + (T - 2) * N * M_w$$

- alpha(T-1), outer loop

$$M_{10}(n) = N * M_r + N * M_w$$

- alpha(T-1), inner loop

$$M_{11}(n) = 2N^2 * M_r$$

- scale alpha(t)

$$M_{12}(n) = N * M_r + N * 2 * M_w$$

$$M_{forward} = \Sigma_i M_i = (2N^2T+2N^2+2NT+3N+T-3)*M_r+(3N^2+2NT+4N+T-3)*M_w$$

## 3.3 Fused Backward and Update

- transpose transitionMatrix

$$M_0(n) = N^2 * M_r + N^2 * M_w$$

- pre-compute ab

$$M_{0.5}(n) = 2KN^2 * M_r + KN^2 * M_w$$

- Init Beta

$$M_1(n) = N * M_w$$

- T, outer loop

$$M_2(n) = 2(T-1) * M_r$$

- T, middle loop

$$M_3(n) = 3(T-1)N * M_r + 4(T-1) * N * M_w$$

- T, inner loop

$$M_4(n) = 3(T-1)N^2 * M_r + (T-1)N^2 * M_w$$

$$M_{fused} = \Sigma_i M_i = (2KN^2+3N^2T-2N^2+3NT-3N+2T-2)*M_r+(KN^2+N(NT+4T-3))*M_w$$

## 3.4 Finishing

$$M_{finishing} = T * M_r$$

## 3.5 Total

$$M_{total} = M_{pre} + M_{forward} + M_{fused} + M_{finishing}$$
$$M_{total_r} = (2KN^2 + 2KN + 5N^2T + 5NT + 2N + 4T - 5) * M_r$$

$$M_{total_w} = (KN^2 + 2KN + N^2T + 3N^2 + 6NT + 4N + T - 3) * M_w$$

Asymptotically:

$$M_{total_r} = (2KN^2 + 5N^2T) * M_r$$
$$M_{total_w} = (KN^2 + N^2T) * M_w$$

## 3.6 Working set

From reo one iteration of the loop at line 577 (fused update and backward):

$$2N^2 + 6N$$