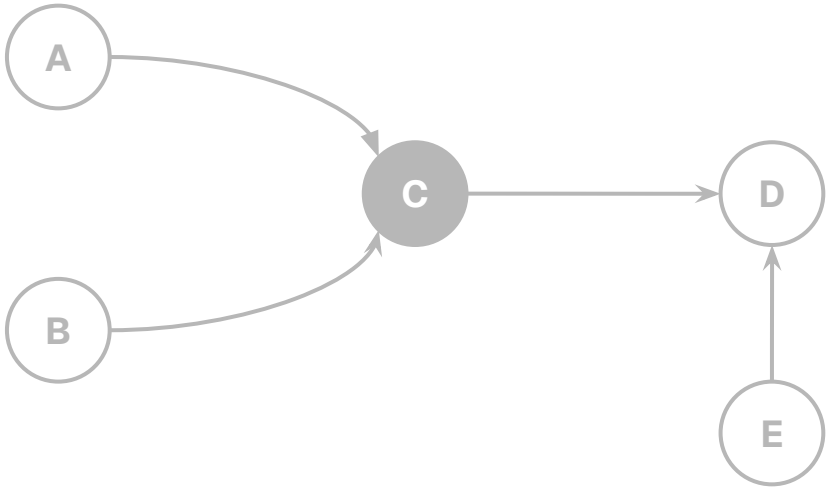
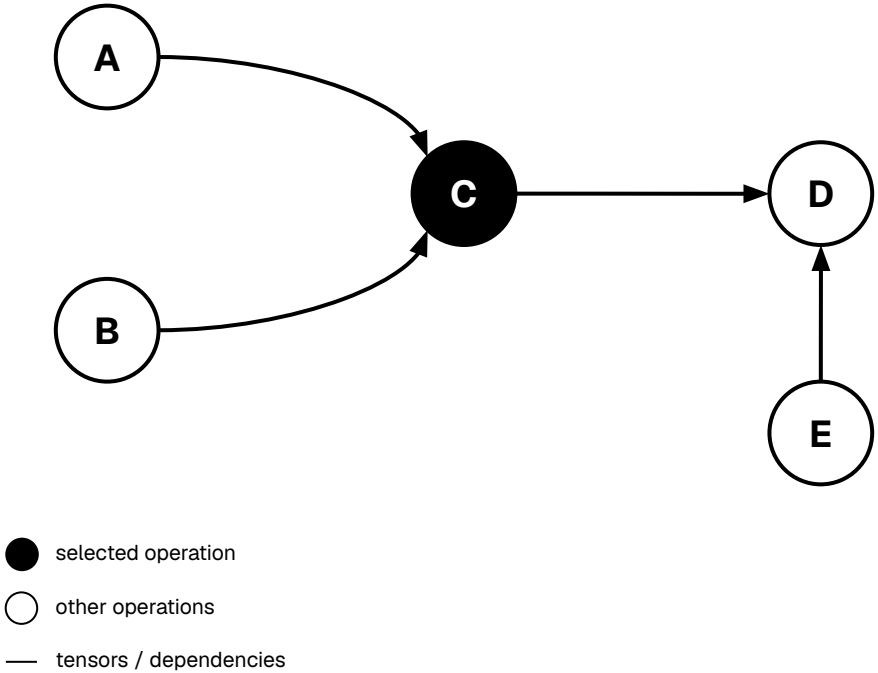
Placeto

Ravichandra Addanki et al.

Presented by Peter



Training as a graph



V

Operation groups

e.g. convolution, matrix multiplication, gradients…

E

Data-flow edges

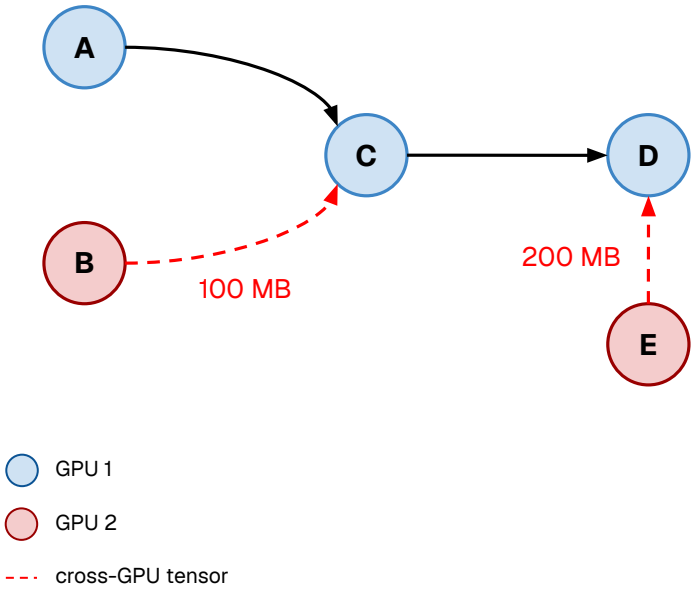
i.e. operation dependencies

Placement is a function

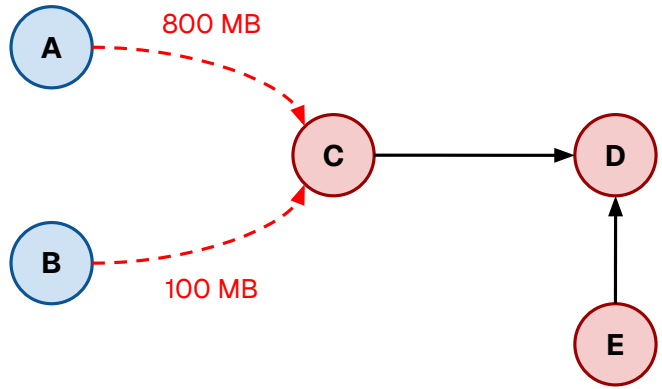
π: V → Devices

Device placement for same graph

OPTION A A, C and D on GPU 1 (300 MB)



OPTION B C, D and E on GPU 2 (900 MB)



Good answer ⇏ reusable algorithm

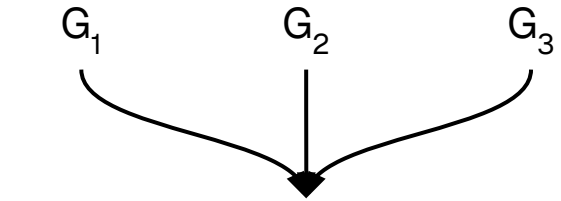
RNN-BASED RL

G_1 → Train agent → π_1

G_2 → Train again → π_2

G_3 → Train again → π_3

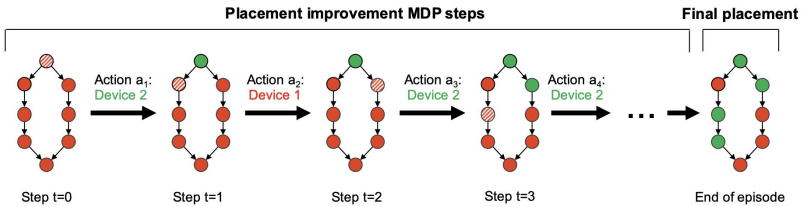
PLACETO



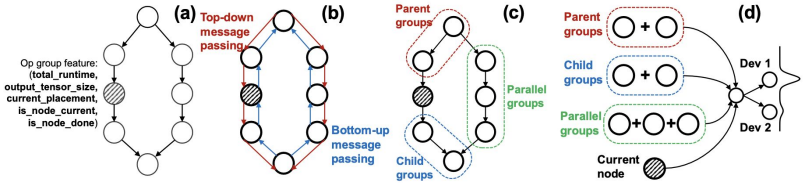
G_4 → Train shared policy θ → π_4

Iteration and embedding

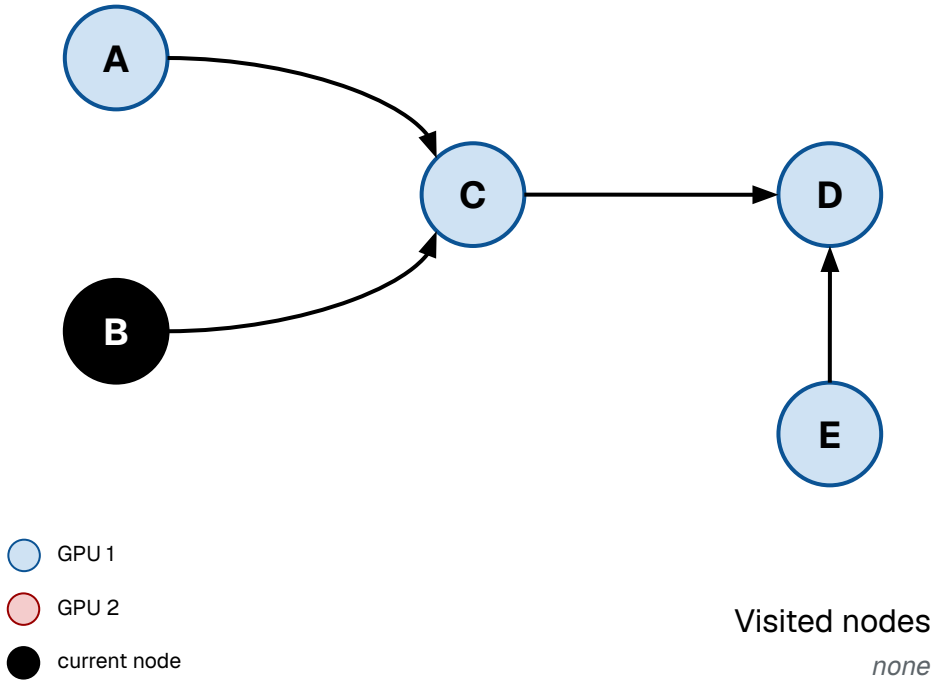
ITERATION



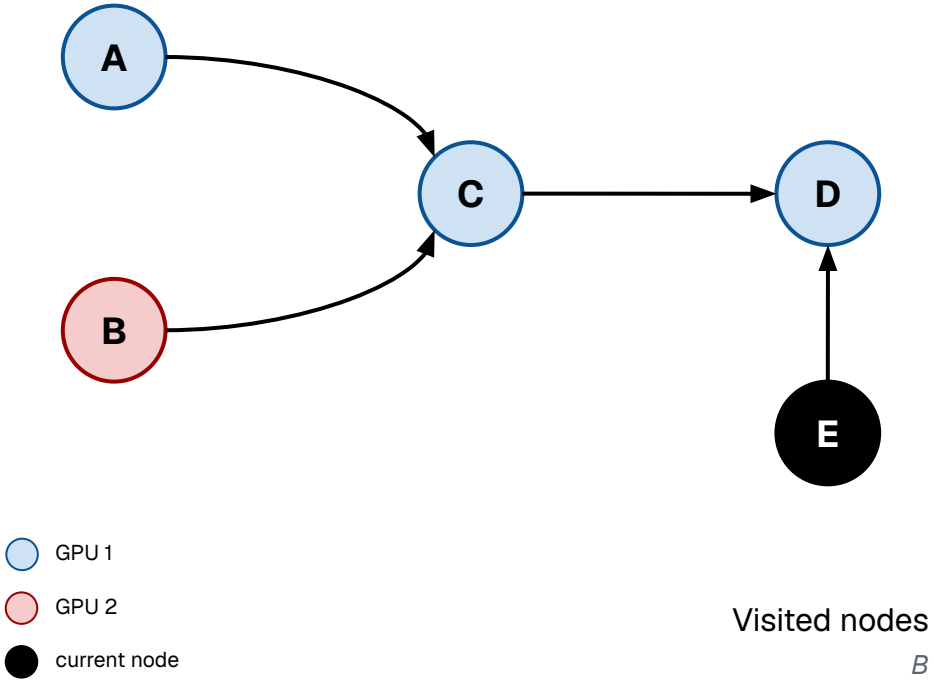
EMBEDDING



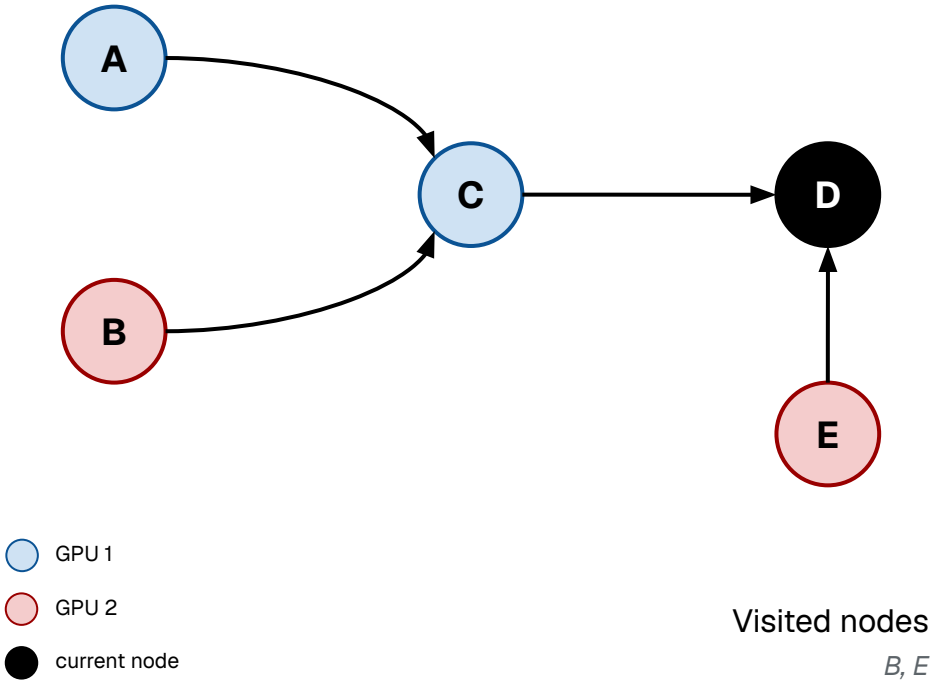
Iteration



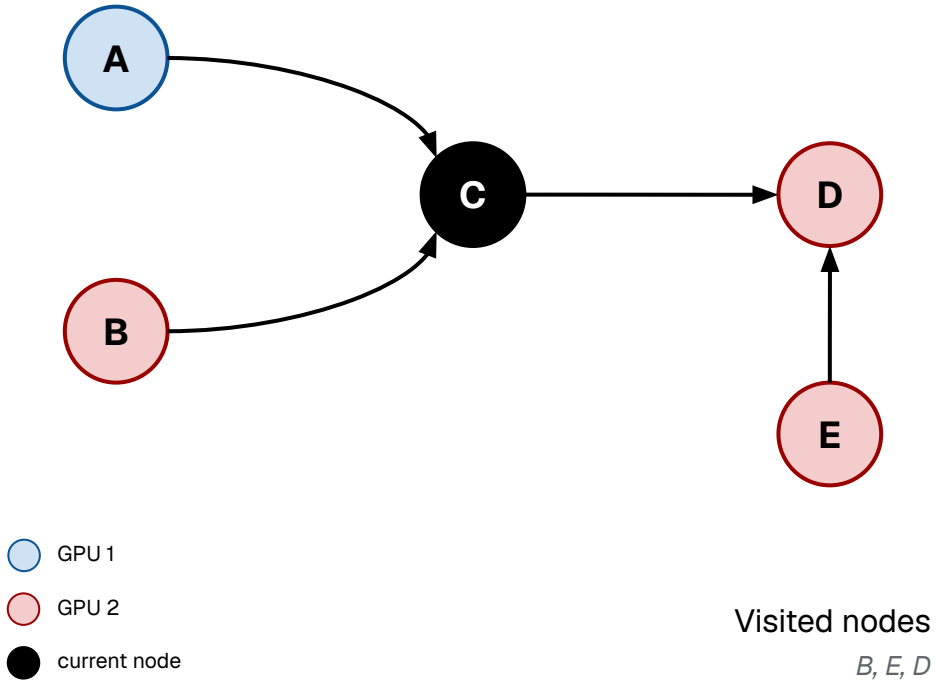
Iteration



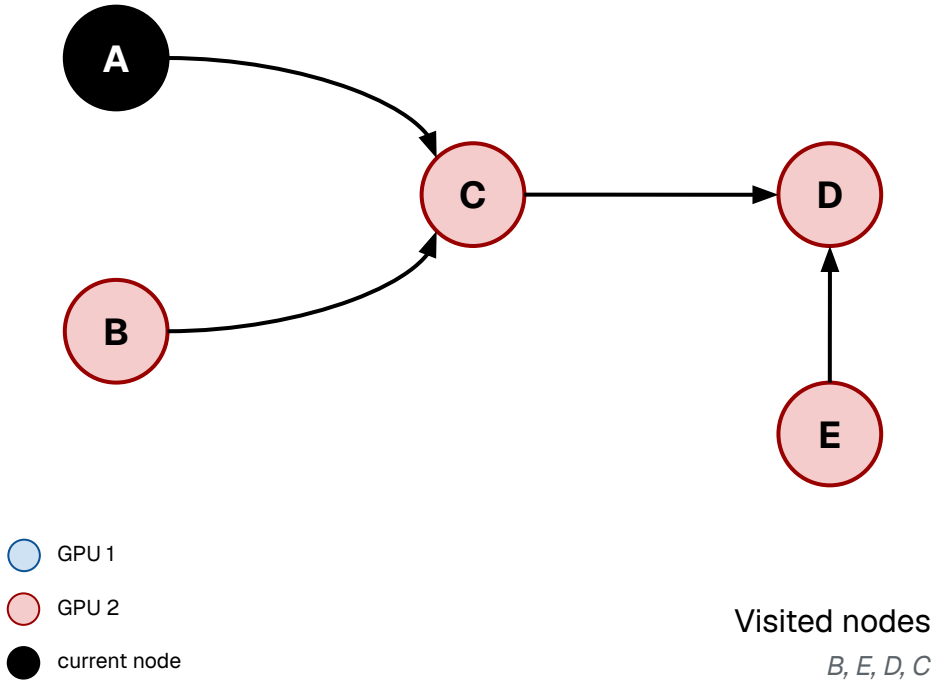
Iteration



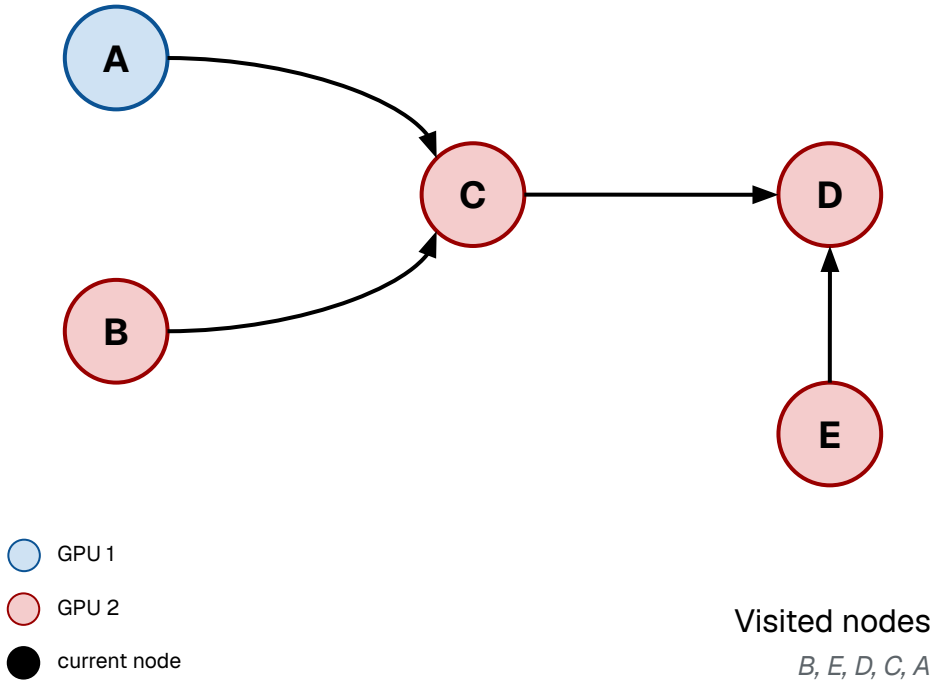
Iteration



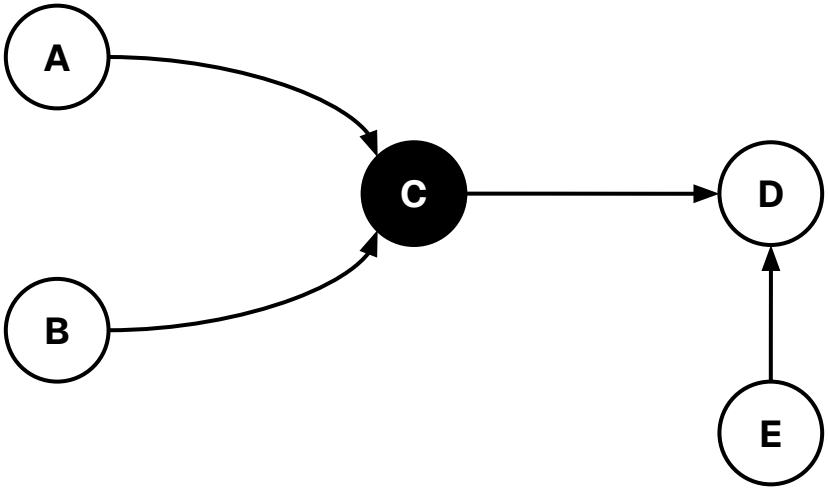
Iteration



Iteration



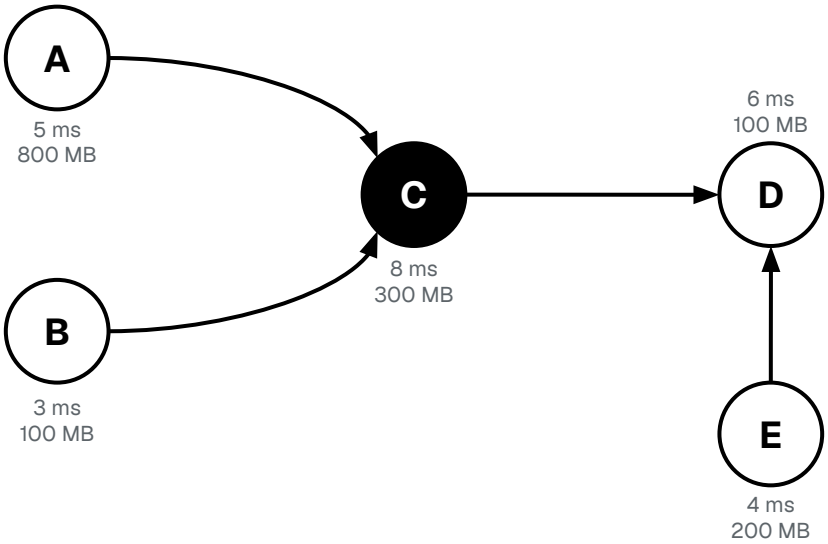
Embedding: message passing



$x_C^{(0)}$ ← message passing round

Estimated runtime	8 ms
Output tensor size	300 MB
Current device	GPU-1
Already visited?	No
Current node?	Yes

Embedding: message passing



$$x_v \leftarrow g(\Sigma_{u \in \xi(v)} f(x_u))$$

1 Transform

$f(x_u)$ creates a message

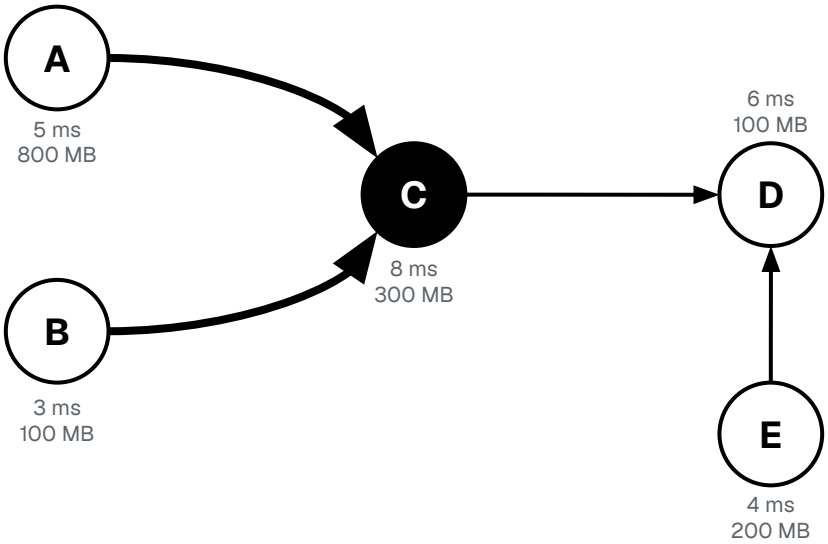
2 Aggregate

Sum the neighbours messages

3 Update

$g(\cdot)$ returns a new node embedding

Embedding: message passing



$$x_v \leftarrow g(\Sigma_{u \in \xi(v)} f(x_u))$$

1 Transform

$f(x_u)$ creates a message

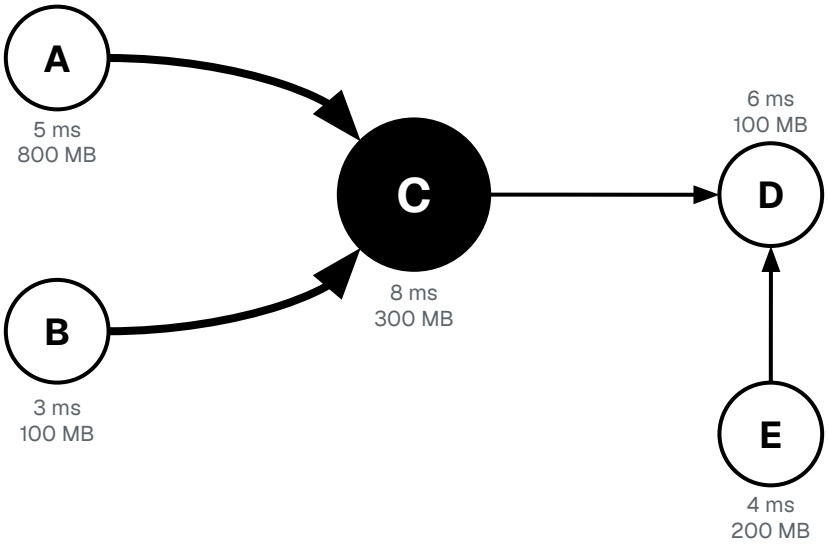
2 Aggregate

Sum the neighbours messages

3 Update

$g(\cdot)$ returns a new node embedding

Embedding: message passing



$$x_v \leftarrow g(\Sigma_{u \in \xi(v)} f(x_u))$$

1 Transform

$f(x_u)$ creates a message

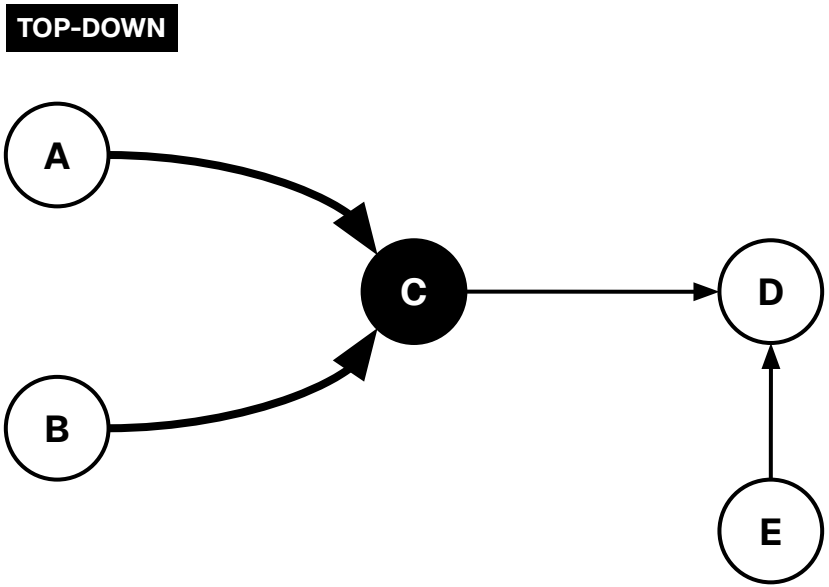
2 Aggregate

Sum the neighbours messages

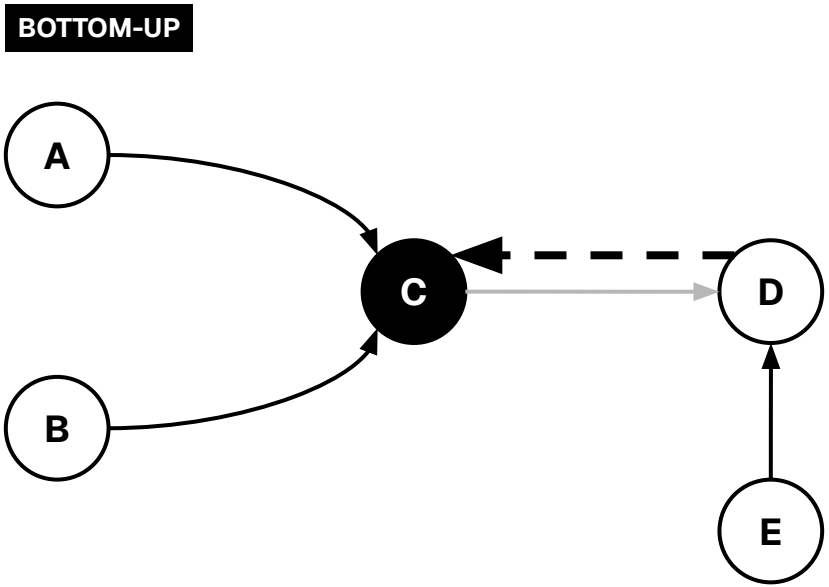
3 Update

$g(\cdot)$ returns a new node embedding

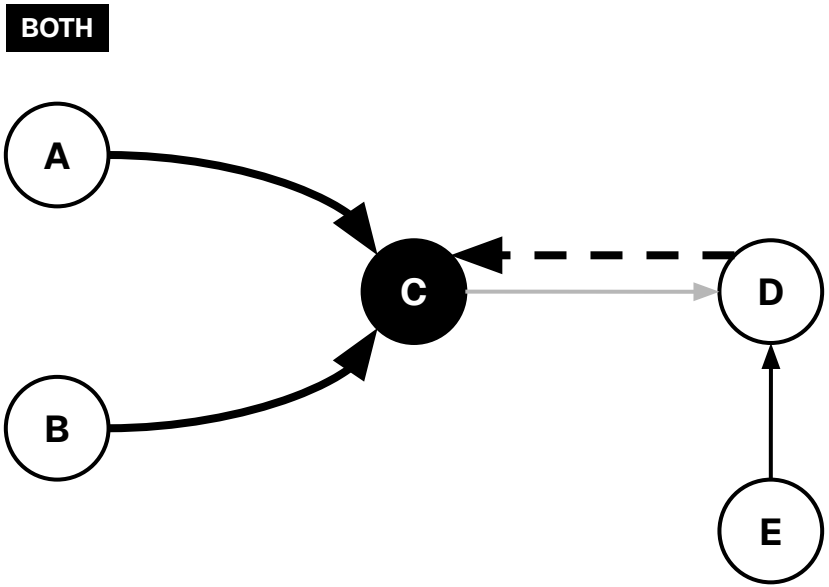
Embedding: message passing



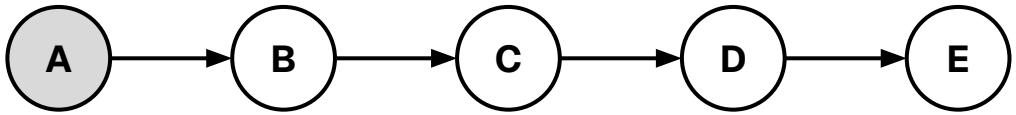
Embedding: message passing



Embedding: message passing



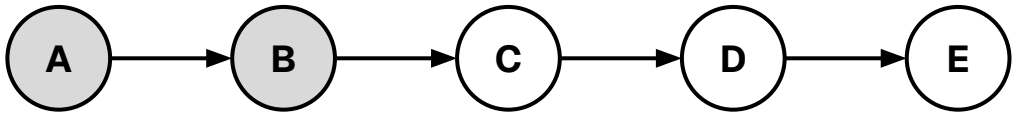
Embedding: message passing



k = 0	k = 1	k = 2	k = 3

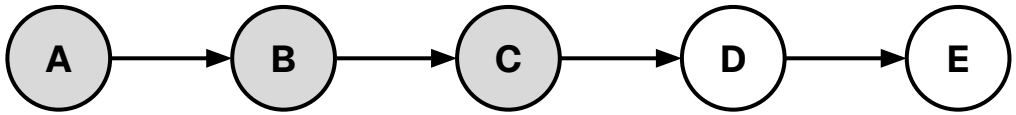
Information is local to A

Embedding: message passing



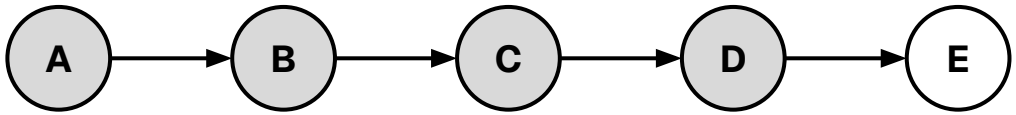
k = 0 | **k = 1** | k = 2 | k = 3

Embedding: message passing



k = 0	k = 1	**k = 2**	k = 3

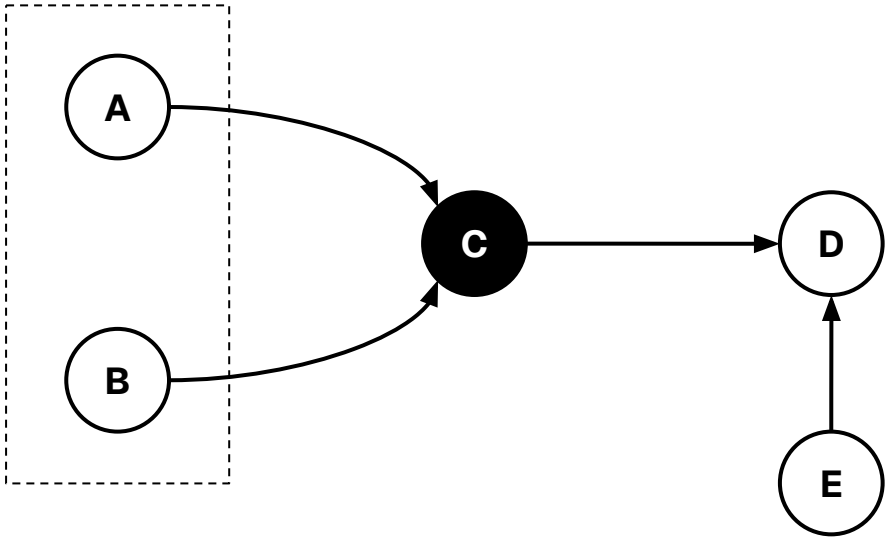
Embedding: message passing



k = 0	k = 1	k = 2	**k = 3**

Information has reached D

Embedding: pooling



$S_{parents}(C)$

Nodes that can reach C

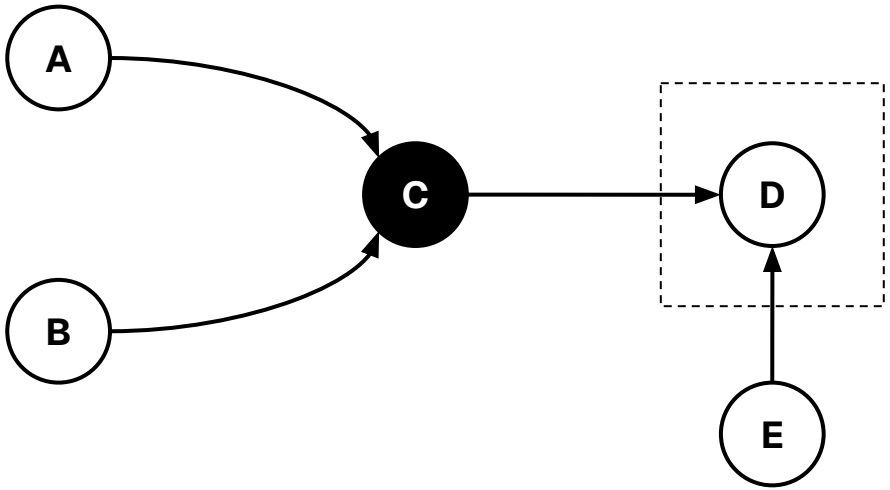
$S_{children}(C)$

Nodes that C can reach

$S_{parallel}(C)$

Neither direction has a path

Embedding: pooling



$S_{parents}(C)$

Nodes that can reach C

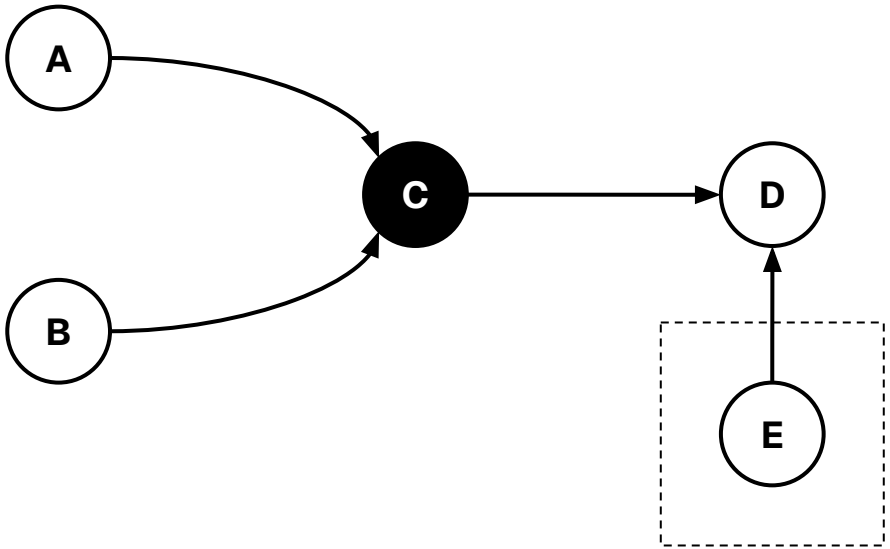
$S_{children}(C)$

Nodes that C can reach

$S_{parallel}(C)$

Neither direction has a path

Embedding: pooling



$S_{parents}(C)$

Nodes that can reach C

$S_{children}(C)$

Nodes that C can reach

$S_{parallel}(C)$

Neither direction has a path

Embedding: pooling

$$P_i(v) = h_i(\Sigma_{u \in S_i(v)} l_i(x_u))$$

TRANSFORM **SUM** **INTERPRET**

$l(A) = [2, 1]$
$l(B) = [1, 3]$



$[2, 1] + [1, 3] = [3, 4]$



$h_{parents}([3, 4])$

Embedding: pooling

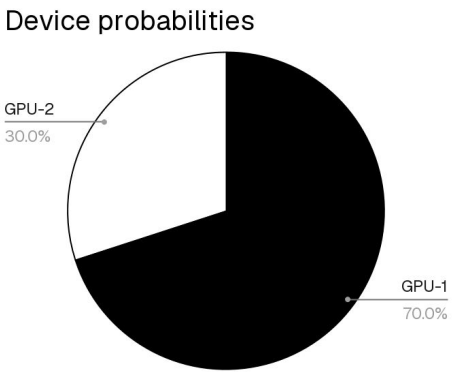
x_c C's embedding

$P_{parents}$ Upstream

$P_{children}$ Downstream

$P_{parallel}$ Independent

$$\rightarrow [\, x_c \,||\, P_{parents} \,||\, P_{children} \,||\, P_{parallel} \,] \rightarrow$$



Performance

	Placement runtime (sec)							Training time (# placements sampled)		Improvement	
Model	CPU only	Single GPU	#GPUs	Expert	Scotch	Placeto	RNN-based	Placeto	RNN-based	Runtime Reduction	Speedup factor
Inception-V3	12.54	1.56	2	1.28	1.54	1.18	1.17	1.6 K	7.8 K	- 0.85%	**4.8** ×
			4	1.15	1.74	1.13	1.19	5.8 K	35.8 K	5%	**6.1** ×
NMT	33.5	OOM	2	OOM	OOM	2.32	2.35	20.4 K	73 K	1.3 %	**3.5** ×
			4	OOM	OOM	2.63	3.15	94 K	51.7 K	**16.5 %**	0.55 ×
NASNet	37.5	1.28	2	0.86	1.28	0.86	0.89	3.5 K	16.3 K	3.4%	**4.7** ×
			4	0.84	1.22	0.74	0.76	29 K	37 K	2.6%	**1.3** ×

INCEPTION-V3

6.1□

fewer placement evaluations

(5.8K versus 35.8K)

NMT

16.5%

lower execution runtime

(2.63 s versus 3.15 s)

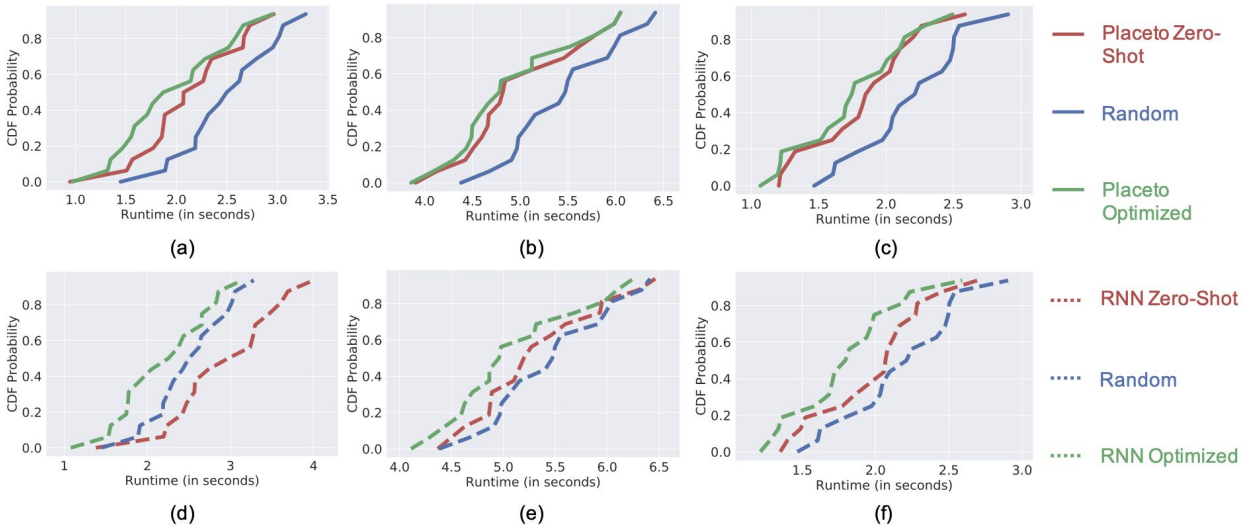
NMT

1.82□

more placement evaluations

(94K versus 51.7K)

Performance



Placeto zero-shot

w/ graph-specific optimisation

RNN zero-shot

w/o graph-specific optimisation

Critical analysis

CONTRIBUTION

Learning a policy for related graphs

Clear graph-specific bottleneck

Functional structural design

LIMITATION

Generalisation robustness unclear

Only 16 unseen graphs per family

No test of increasing structural variation

Predetermined operation grouping (?)

DEVELOPMENTS

Graph-aware placement

GiPH (2023):
dynamic heterogeneous clusters

NeurIPS (2024):
structure-aware placement

① Local messages

② Global context

③ Reusable decisions