

A Graph-to-Sequence Model for AMR-to-Text Generation

Linfeng Song¹, Yue Zhang^{2,4}, Zhiguo Wang³ and Daniel Gildea¹

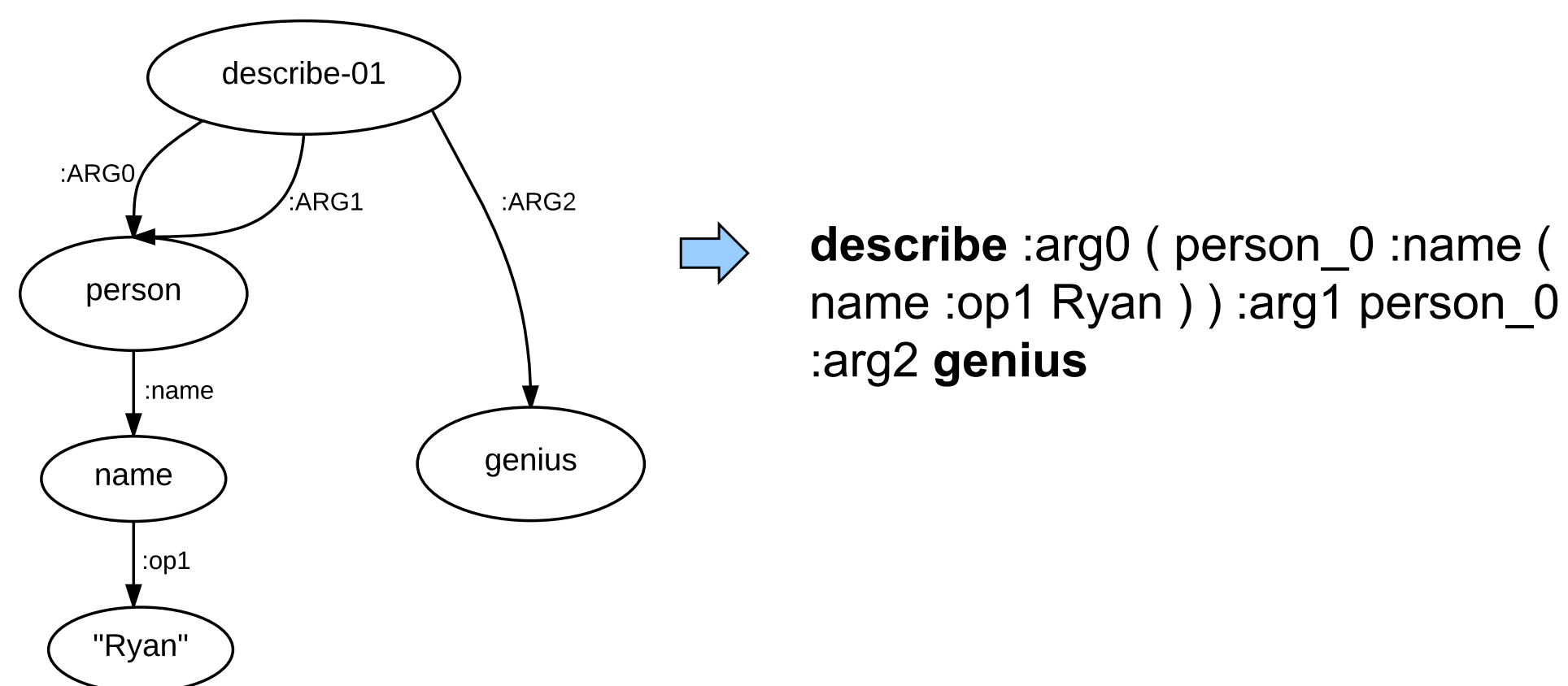
¹University of Rochester, ²Singapore University of Technology and Design, ³IBM Research, ⁴Westlake Institute for Advanced Study

Contribution

- Introduce graph recurrent network (GRN) for modeling AMR graph.
- It shows better performance than a sequential LSTM encoder on linearized AMRs.
- We release our code at <https://github.com/freesunshine0316/neural-graph-to-seq-mp>.

Baseline: sequence-to-sequence

Linearization and Anonymization



Drawbacks:

- Serialization causes loss of structural information. Closely-related nodes, such as parents, children and siblings can be far away after serialization.
- Anonymization requires manual work for defining heuristic rules. The rules have to be adapted for new domains.

Leveraging large-scale automatic data (Konstas et al., 2017)

Data: D_a : automatic training data, D_g : gold training data, D_v : devset

Input: T : total steps of automatic training, E : fine-tune after the number of steps

for $t \in [1 \dots T]$ **do**

 pretrain with D_a for one step;

if $t \% |D_a| == 0$ **or** $t \% E == 0$ **then**

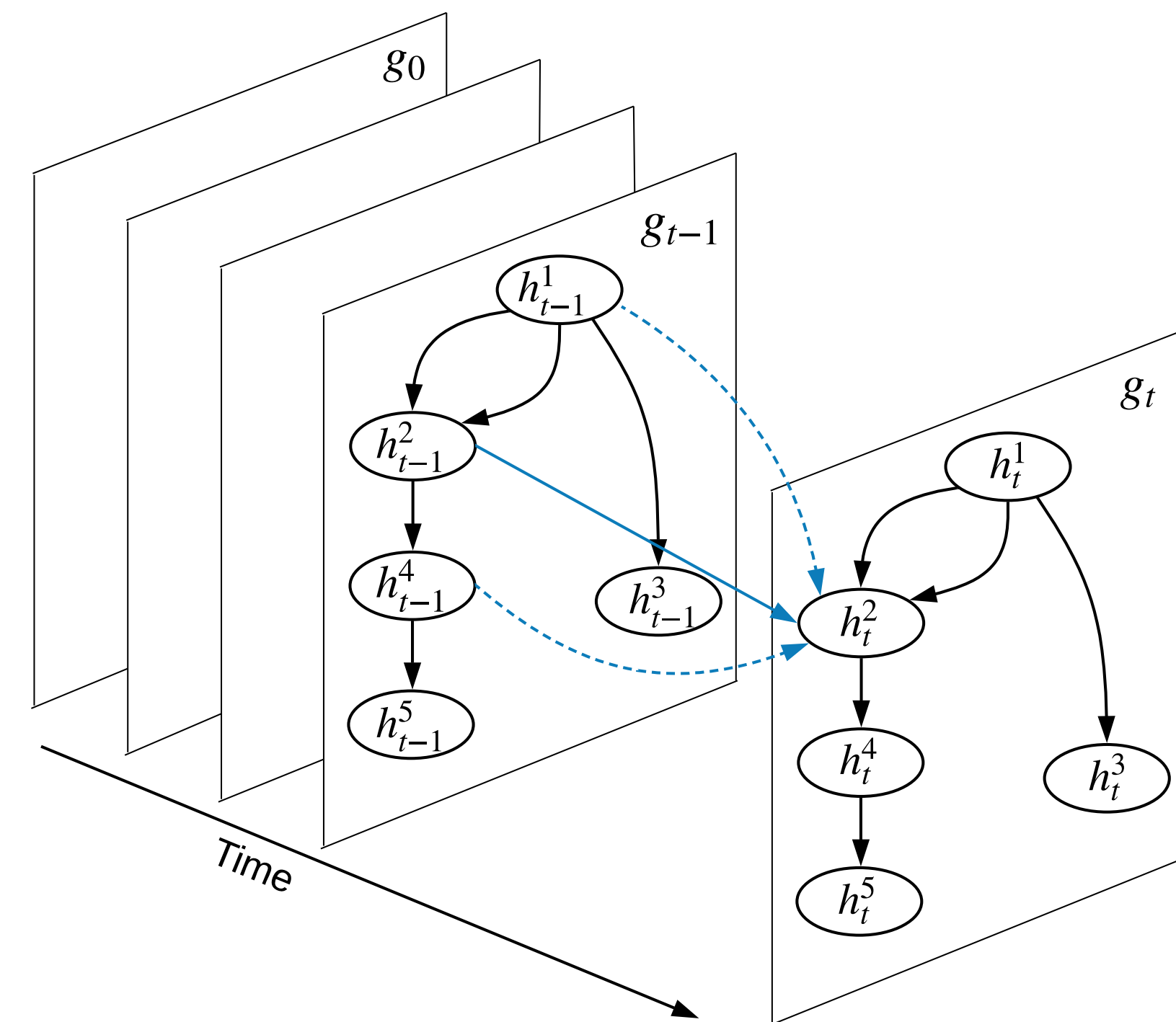
 fine-tune with D_g ;

 evaluate with D_v ;

end

end

Graph recurrent network (GRN)



We model the graph state ($g = \{h^i\}_{v_i \in V}$) via state transition, h^i incorporates larger context through the gated (LSTM-based) state transition.

$$c_t^i, h_t^i = \text{LSTM}([x_j^i; x_j^o], [h_j^i; h_j^o], c_{t-1}^i)$$

where x_j^i and x_j^o are the sum of all incoming and outgoing edge representations of v_j , respectively. Similarly, h_j^i and h_j^o are the sum of the hidden states of all incoming and outgoing neighbors:

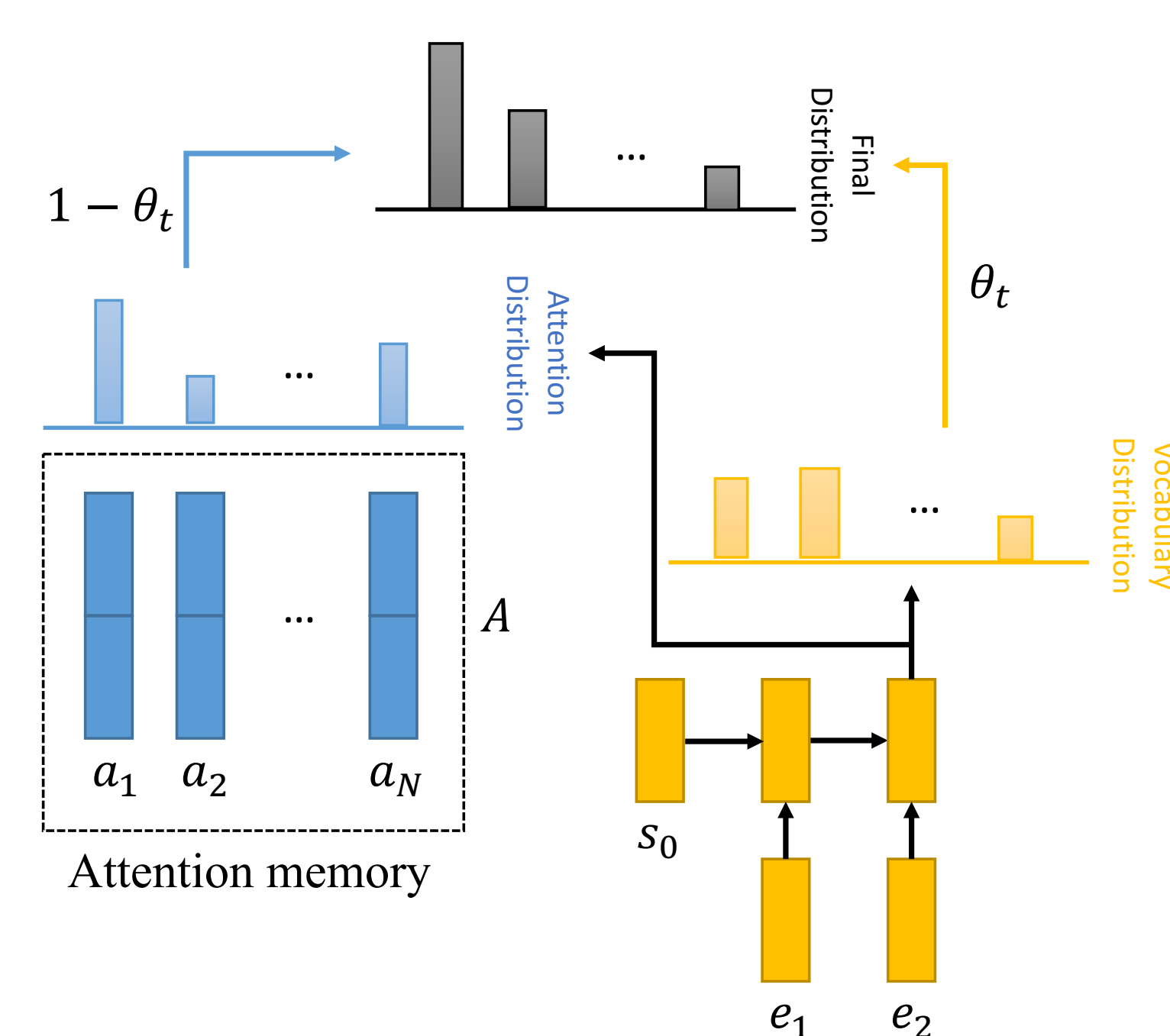
$$x_j^i = \sum_{(i,j,l) \in E_{in}(j)} x_{i,j}^l; \quad x_j^o = \sum_{(j,k,l) \in E_{out}(j)} x_{j,k}^l$$

$$h_j^i = \sum_{(i,j,l) \in E_{in}(j)} h_{t-1}^i; \quad h_j^o = \sum_{(j,k,l) \in E_{out}(j)} h_{t-1}^k$$

The representation vector $x_{i,j}^l$ for edge (i, j, l) is calculated by the edge label embedding e_l and the concept of the other node e_i .

$$x_{i,j}^l = W_4([e_l; e_i]) + b_4 \quad x_{i,j}^c = W_4([e_l; e_i; h_i^c]) + b_4$$

Copy mechanism



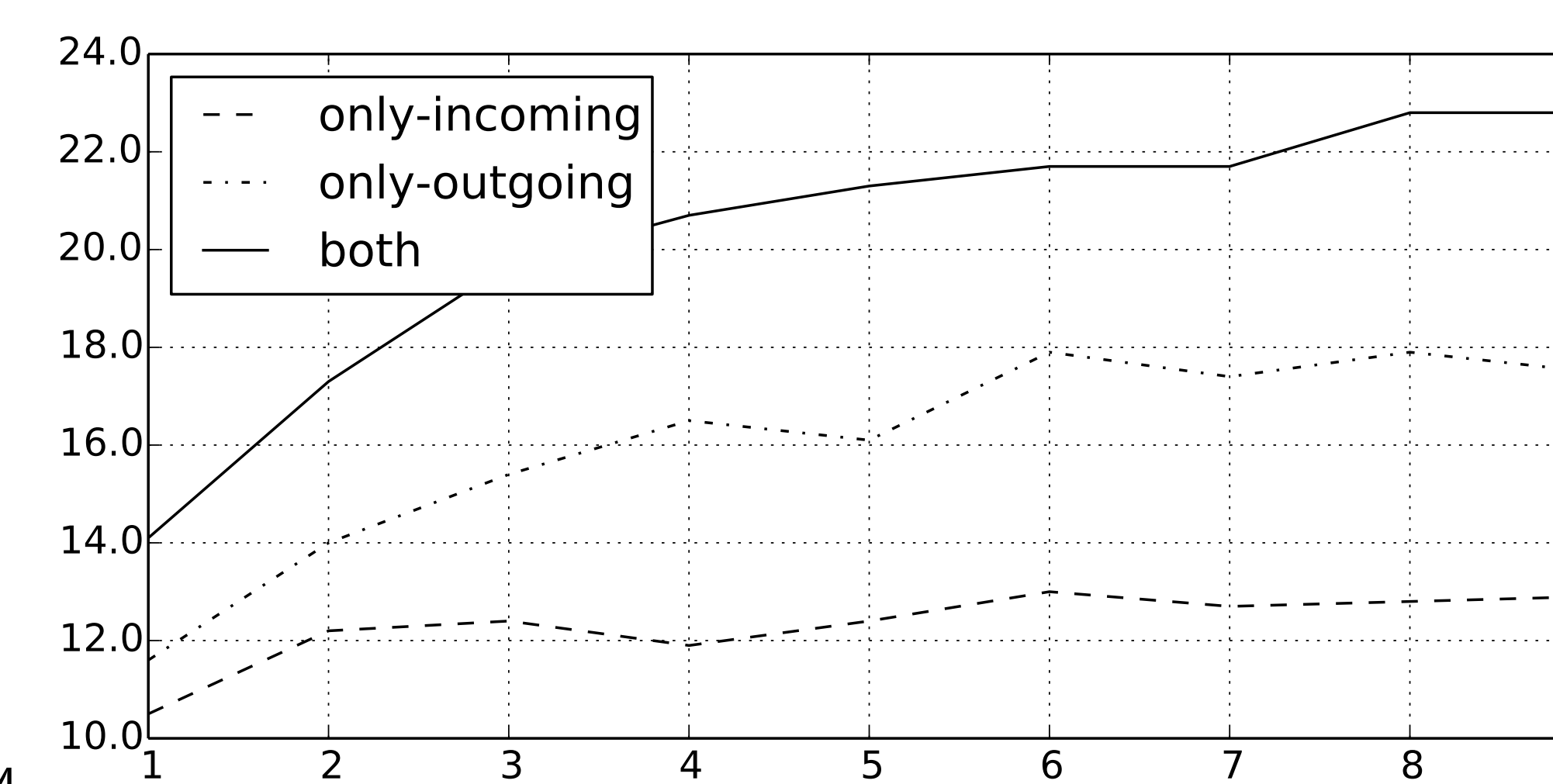
Main results

Data:

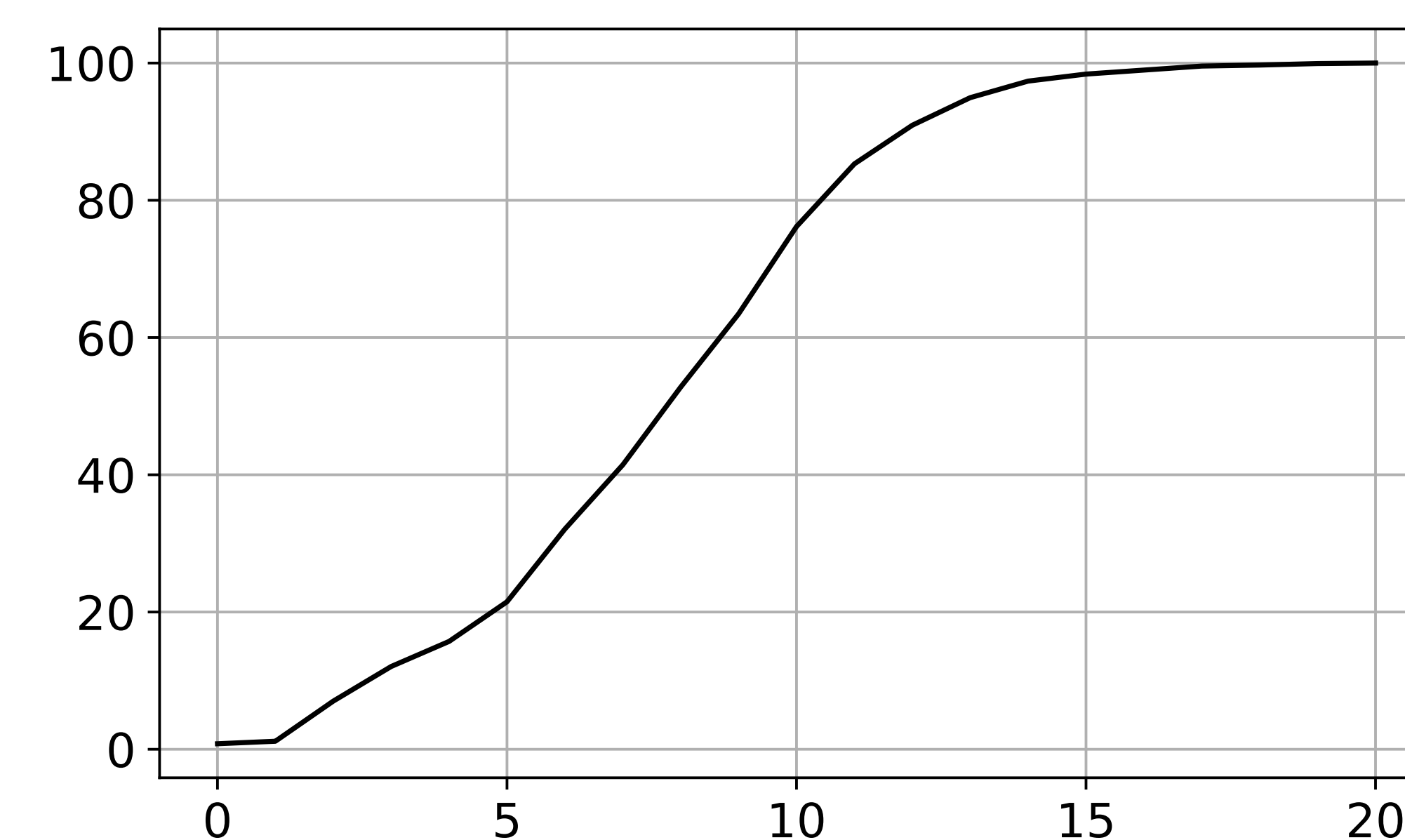
- LDC2015E86, Train/Dev/Test: 16833/1368/1371
- Up to 2M raw data (Gigaword) parsed by JAMR

Model	BLEU
PBMT	26.9
SNRG	25.6
Tree2Str	23.0
MSeq2seq+Anon	22.0
Graph2seq+copy	22.7
Graph2seq+charLSTM+copy	23.3
MSeq2seq+Anon (200K)	27.4
MSeq2seq+Anon (2M)	32.3
MSeq2seq+Anon (20M)	33.8
Seq2seq+charLSTM+copy (200K)	27.4
Seq2seq+charLSTM+copy (2M)	31.7
Graph2seq+charLSTM+copy (200K)	28.2
Graph2seq+charLSTM+copy (2M)	33.0 33.6

Analysis



DEV BLEU scores against transition steps for the graph encoder.



Percentage of DEV AMRs with different diameters.

Development experiments

Model	BLEU	Time
Seq2seq	18.8	35.4s
Seq2seq+copy	19.9	37.4s
Seq2seq+charLSTM+copy	20.6	39.7s
Graph2seq	20.4	11.2s
Graph2seq+copy	22.2	11.1s
Graph2seq+Anon	22.1	9.2s
Graph2seq+charLSTM+copy	22.8	16.3s

Example Outputs

(p / possible-01 :polarity -
:ARG1 (l / look-over-06
:ARG0 (w / we)
:ARG1 (a / account-01
:ARG1 (w2 / war-01
:ARG1 (c2 / country :wiki "Japan"
:name (n2 / name :op1 "Japan"))
:time (p2 / previous)
:ARG1-of (c / call-01 :mod (s / so)))
:mod (o / old)))

Lin: possible :polarity - :arg1 (look-over :arg0 we :arg1 (account :arg1 (war :arg1 (country :wiki japan :name (name :op1 japan)) :time previous :arg1-of (call :mod so)) :mod old))

Ref: we can n't look over the old accounts of the previous so-called anti-japanese war .

S2S: we can n't be able to account the past drawn out of japan 's entire war .

G2S: we can n't be able to do old accounts of the previous and so called japan war .

G2S+CP: we can n't look-over the old accounts of the previous so called war on japan .

(p / provide-01
:ARG0 (a / agree-01
:ARG1 (a2 / and
:op1 (s / staff
:prep-for (c / center :mod (r / research-01)))
:op2 (f / fund-01
:prep-for c)))

Lin: provide :arg0 agree :arg1 (and :op1 (staff :prep-for (center :mod research)) :op2 (fund :prep-for center))

Ref: the agreement will provide staff and funding for the research center .

S2S: agreed to provide research and institutes in the center .

G2S: the agreement provides the staff of research centers and funding .

G2S+CP: the agreement provides the staff of the research center and the funding .