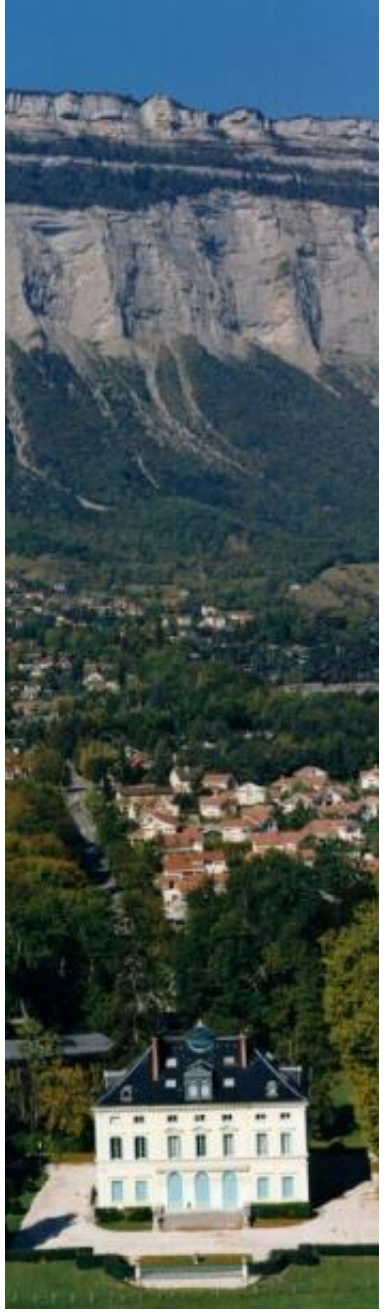
Learning Goal-Oriented Dialog with Gated Memory Networks

Julien Perez and Fei Lui
Xerox Research Centre Europe

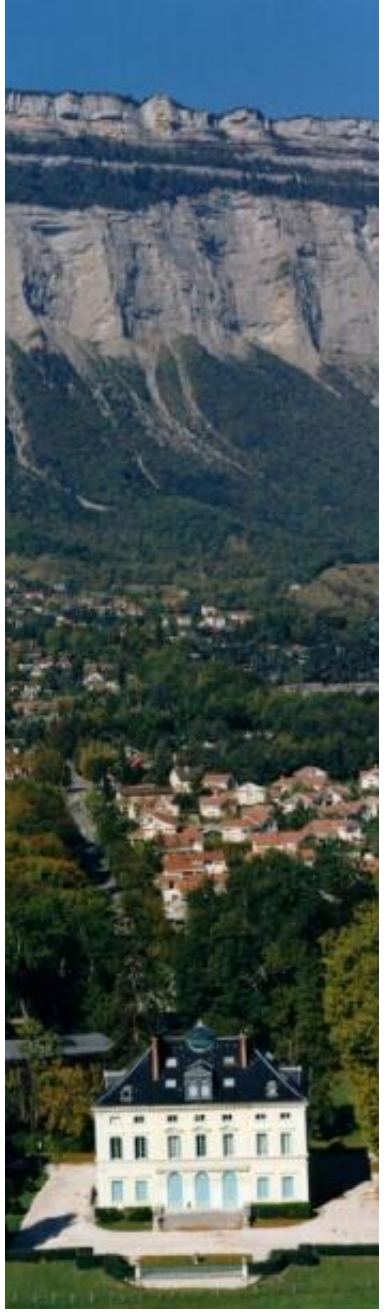


Introduction
Talking the talk

Building Dialog Systems, some current issues

Gated MemN2N for end2end learning

Results and current work



Dialog Systems, some current issues

Modularity is the current solution

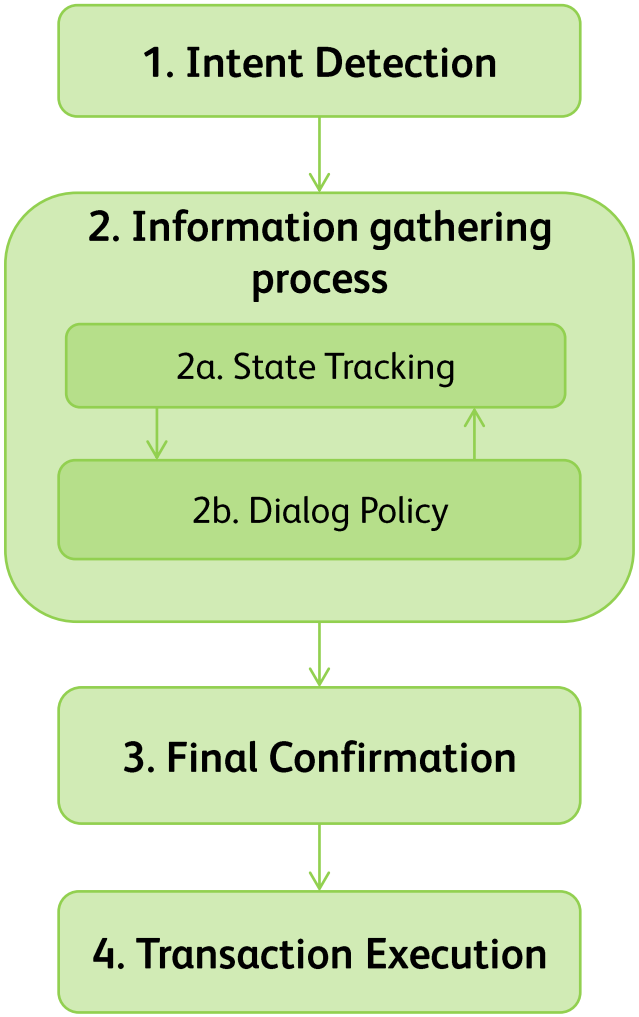
- Divide and Conquer approach
- Annotation processes are required
- Hand-crafted models, case-by-case adaptation

End-to-End opportunities

- Leveraging raw dialogs
- Can be (automatically) enriched with meta-data
- Seamless integration of back-end access

Promising research context

- (Deep) imitation learning
- Non-Markovian control tasks



xerox

End-to-End Memory Network for dialog

FAIR Dialog bAbI tasks

Learning to dialog … from dialogs

- 5 dialog tasks with/without OOV
- 1 DSTC-2 end-to-end dialog task

Goal oriented dialogs

- Learn on synthetic + real dialogs
- Backed with a Knowledge Base

Memory Networks

- End-to-End learnable and flexible
- Non-parametric memory due to attention
- KB-fact and utterance support of the decision

Time	Locutor	Dialog history	Hop #1	Hop #2	Hop #3
3	User	may i have a table in paris	.061	.072	.040
4	Bot	i'm on it	.026	.012	.001
5	User	<silence>	.087	.042	.012
6	Bot	any preference on a type of cuisine	.026	.023	.001
7	User	i love indian food	.081	.070	.055
8	Bot	how many people would be in your party	.025	.006	.001
9	User	we will be six	.059	.051	.018
10	Bot	which price range are looking for	.038	.043	.004
11	User	in a moderate price range please	.080	.095	.096
12	Bot	ok let me look into some options for you	.025	.042	.003
13	User	<silence>	.127	.069	.032
14	Bot	`api_call indian paris six moderate`	.062	.113	.043
15	User	instead could it be with french food	.188	.311	.683
16	Bot	sure is there anything else to update	.016	.007	.001
17	User	no	.028	.013	.007
18	Bot	ok let me look into some options for you	.011	.006	.000
User input		<silence>			
Correct answer		`api_call french paris six moderate`			
Predicted answer		`api_call french paris six moderate` **[Correct]**			

Time	Locutor	Dialog history	Hop #1	Hop #2	Hop #3
14	Bot	`api_call indian paris six moderate`	.012	.000	.000
15	User	instead could it be with french food	.067	.103	.147
20	Bot	`api_call french paris six moderate`	.012	.000	.000
21	User	`resto_1 r_phone rest_1_phone`	.018	.004	.000
23	User	`resto_1 r_cuisine french`	.029	.005	.000
24	User	`resto_1 r_location paris`	.060	.292	.094
25	User	`resto_1 r_number six`	.050	.298	.745
26	User	`resto_1 r_price moderate`	.060	.090	.002
27	User	`resto_1 r_rating 6`	.016	.002	.000
30	User	`resto_2 r_cuisine french`	.031	.007	.000
31	User	`resto_2 r_location paris`	.040	.081	.004
32	User	`resto_2 r_number six`	.020	.012	.000
33	User	`resto_2 r_price moderate`	.029	.009	.000
37	User	`resto_3 r_cuisine french`	.014	.001	.000
38	User	`resto_3 r_location paris`	.028	.016	.001
39	User	`resto_3 r_number six`	.024	.022	.004
40	User	`resto_3 r_price moderate`	.039	.015	.001
User input		<silence>			
Correct answer		what do you think of this option: resto_1			
Predicted answer		what do you think of this option: resto_1 **[Correct]**			

[1] Learning End-to-End Goal-Oriented Dialog, Bordes and Weston, 2016

xerox

End-to-End Memory Network for dialog

Description

Model

$$\boldsymbol{m}_i = \boldsymbol{A}\Phi(x_i) \qquad \boldsymbol{u} = \boldsymbol{B}\Phi(q)$$

$$\boldsymbol{c}_i = \boldsymbol{C}\Phi(x_i)$$

$$p_i = \text{softmax}(\boldsymbol{u}^\top \boldsymbol{m}_i)$$

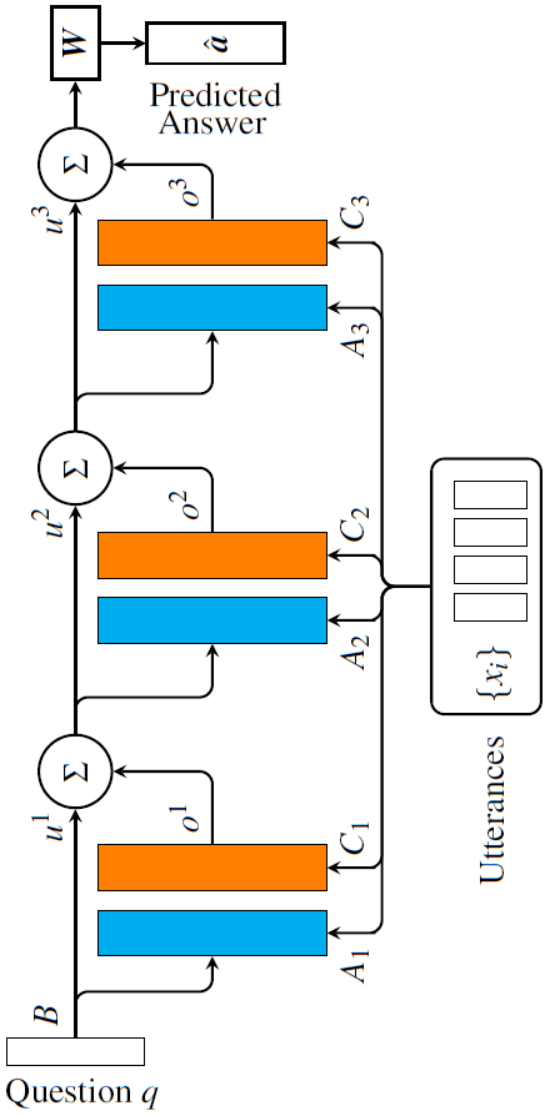
$$\boldsymbol{o} = \sum_i p_i \boldsymbol{c}_i$$

$$\boldsymbol{u}^{k+1} = \boldsymbol{o}^k + \boldsymbol{u}^k$$

$$\hat{\boldsymbol{a}} = \text{softmax}(\boldsymbol{u}^\top \boldsymbol{W}' \Phi(\boldsymbol{y}_1), \ldots, \boldsymbol{u}^\top \boldsymbol{W}' \Phi(\boldsymbol{y}_{|C|}))$$

Optimization task

- Categorical cross-entropy
- Stochastic Gradient Descent with clipping
- Grid-searched Hyper Parameters



xerox

Gated End-to-End Memory Network

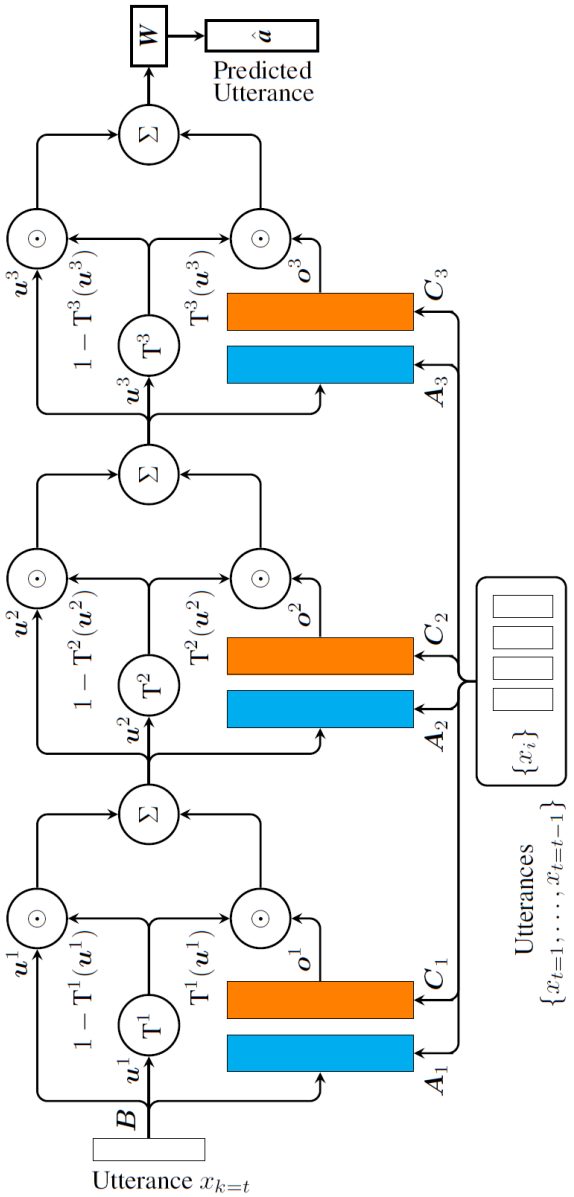
Proposition

Model

$$\boldsymbol{m}_i = \boldsymbol{A}\Phi(x_i) \qquad \boldsymbol{u} = \boldsymbol{B}\Phi(q)$$
$$\boldsymbol{c}_i = \boldsymbol{C}\Phi(x_i)$$
$$p_i = \text{softmax}(\boldsymbol{u}^\top \boldsymbol{m}_i)$$
$$\boldsymbol{o} = \sum_i p_i \boldsymbol{c}_i$$

$$\mathrm{T}^k(\boldsymbol{u}^k) = \sigma(\boldsymbol{W}_T^k \boldsymbol{u}^k + \boldsymbol{b}_T^k)$$
$$\boldsymbol{u}^{k+1} = \boldsymbol{o}^k \odot \mathrm{T}^k(\boldsymbol{u}^k) + \boldsymbol{u}^k \odot (1 - \mathrm{T}^k(\boldsymbol{u}^k))$$

$$\hat{\boldsymbol{a}} = \text{softmax}(\boldsymbol{u}^\top \boldsymbol{W}' \Phi(\boldsymbol{y}_1), \ldots, \boldsymbol{u}^\top \boldsymbol{W}' \Phi(\boldsymbol{y}_{|C|}))$$



Properties

- End-to-End memory access regulation
- Close to Highway Network and Residual Network

[2] Gated End-to-End Memory Network, Fei Liu and Julien Perez, EACL 2017

xerox

End-to-End Memory Network for dialog

Some technical details

Sentence Representation (PE)

$$c_i = \sum_j Ax_{ij} \Rightarrow m_i = \sum_j l_j \cdot Ax_{ij}$$

$$l_{kj} = (1 - j/J) - (k/d)(1 - 2j/J)$$

Temporal Encoding

$$m_i = \sum_j Ax_{ij} + T_A(i)$$

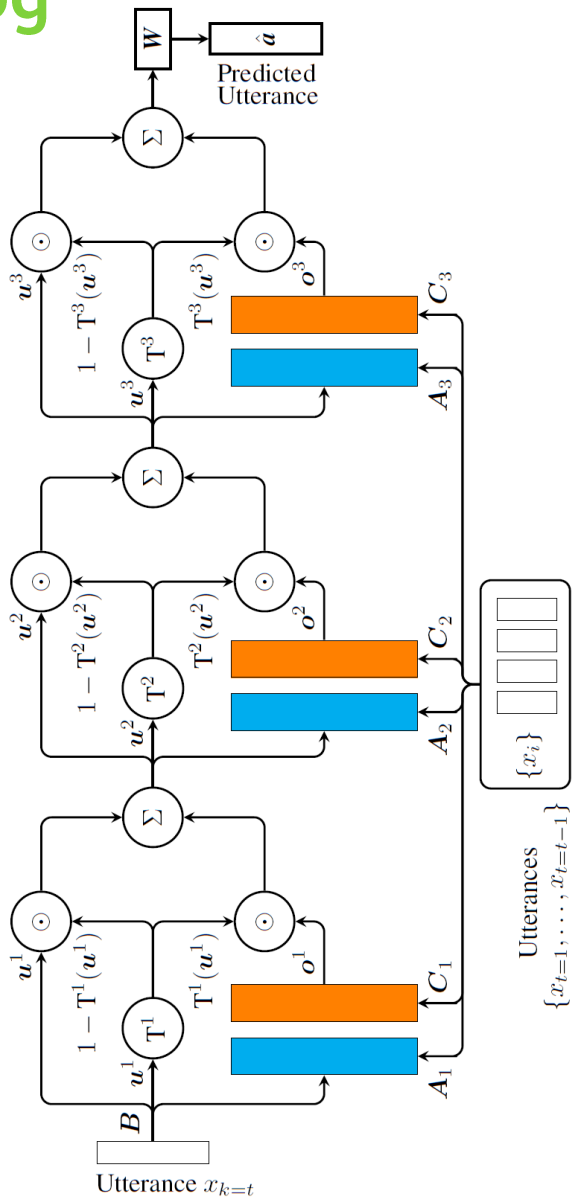
$T_A(i)$: Special matrix encode temporal info

Linear Start(LS)

Initial train with remove all the non-linear activation functions

Weight types

Adjacent or Layer-Wise



xerox

Gated Memory Network

Experiments

Task	MemN2N	GMemN2N	MemN2N +match	GMemN2N +match
T1: Issuing API calls	99.9 (99.6)	**100.0** (100.0)	**100.0** (100.0)	**100.0** (100.0)
T2: Updating API calls	**100.0** (100.0)	**100.0** (100.0)	98.3 (83.9)	**100.0** (100.0)
T3: Displaying options	**74.9** (2.0)	**74.9** (0.0)	**74.9** (0.0)	**74.9** (0.0)
T4: Providing information	59.5 (3.0)	57.2 (0.0)	**100.0** (100.0)	**100.0** (100.0)
T5: Full dialogs	96.1 (49.4)	96.3 (52.5)	93.4 (19.7)	**98.0** (72.5)
Average	86.1 (50.8)	85.7 (50.5)	93.3 (60.7)	**94.6** (74.5)
T1 (OOV): Issuing API calls	72.3 (0.0)	82.4 (0.0)	96.5 (82.7)	**100.0** (100.0)
T2 (OOV): Updating API calls	78.9 (0.0)	78.9 (0.0)	**94.5** (48.4)	94.2 (47.1)
T3 (OOV): Displaying options	74.4 (0.0)	**75.3** (0.0)	75.2 (0.0)	75.1 (0.0)
T4 (OOV): Providing information	57.6 (0.0)	57.0 (0.0)	**100.0** (100.0)	**100.0** (100.0)
T5 (OOV): Full dialogs	65.5 (0.0)	66.7 (0.0)	77.7 (0.0)	**79.4** (0.0)
Average	69.7 (0.0)	72.1 (0.0)	88.8 (46.2)	**89.7** (49.4)
T6: Dialog state tracking 2	41.1 (0.0)	47.4 (1.4)	41.0 (0.0)	**48.7** (1.4)

Performances on Facebook bAbI End-to-End dialog tasks

Gated Memory Network

End-to-End Transactional dialog management

Dialog	MemN2N			GMemN2N		
	Hop 1	Hop 2	Hop 3	Hop 1	Hop 2	Hop 3
. . . (omitted due to space constraints)	0.00	0.00	0.00	0.00	0.00	0.00
api_call r_cuisine south cheap.	0.01	0.00	0.00	0.00	0.00	0.00
api_call_return (the_lucky_star).	0.17	0.00	0.00	0.00	0.00	0.00
<silence>.	0.00	0.00	0.00	0.00	0.00	0.00
nandos is a nice place in the south of town and the prices are cheap.	0.00	0.00	0.36	0.00	0.00	0.08
<silence>.	0.00	0.00	0.00	0.00	0.00	0.00
the_lucky_star is a nice place in the south of town and the prices are cheap.	0.00	0.00	0.47	0.00	0.28	0.91
what else is there.	0.00	0.00	0.00	0.00	0.00	0.00
the_lucky_star is a great restaurant.	0.01	0.00	0.00	0.00	0.50	0.00
anything else.	0.00	0.00	0.00	0.00	0.00	0.00
the post code of the_lucky_star is the_lucky_star_post_code.	0.03	0.00	0.00	1.00	0.23	0.00
is there anything else.	0.00	0.00	0.00	0.00	0.00	0.00
sorry but there is no other restaurant in the cheap price range and the south of town.	0.77	1.00	0.17	0.00	0.00	0.00
Avg. transform gate cell value, $\sum_i \mathrm{T}(u)_i$	N/A	N/A	N/A	0.12	0.30	0.76
Question: can i have the phone number?						
Answer: the phone number of the_lucky_star is the_lucky_star_phone						
MemN2N prediction: the phone number of nandos is nandos_phone						
GMemN2N prediction: the phone number of the_lucky_star is the_lucky_star_phone						

Table 2: MemN2N vs. GMemN2N- Dialog dataset - Task 6 - DSTC2

Gated End-to-End Memory Network
visualizations

Gated Memory Access Regulation

$$\mathrm{T}^k(\boldsymbol{u}^k) = \sigma(\boldsymbol{W}_T^k \boldsymbol{u}^k + \boldsymbol{b}_T^k)$$
$$\boldsymbol{u}^{k+1} = \boldsymbol{o}^k \odot \mathrm{T}^k(\boldsymbol{u}^k) + \boldsymbol{u}^k \odot (1 - \mathrm{T}^k(\boldsymbol{u}^k))$$

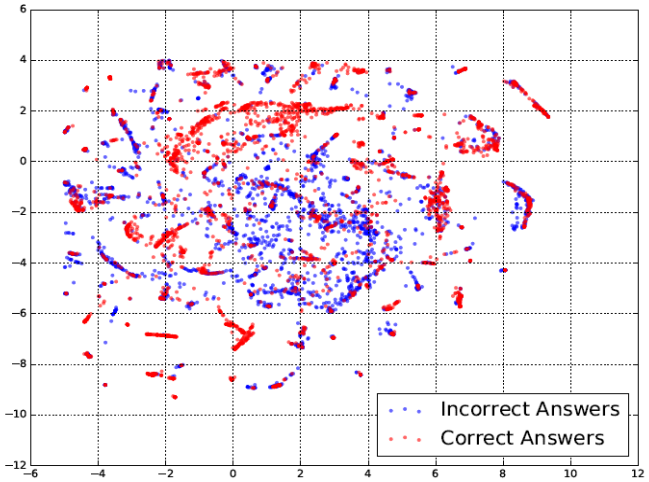


Figure 3: t-SNE scatter plot of the flattened gate values

Remarks

- End-to-end training of the reasoning capability
- Prototypical memory access patterns can be aggregated and visualized
- Open for improvement
- Variance issues (regularization ? curriculum ?)

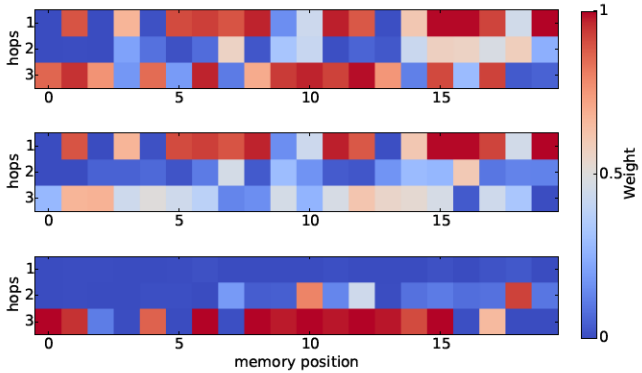
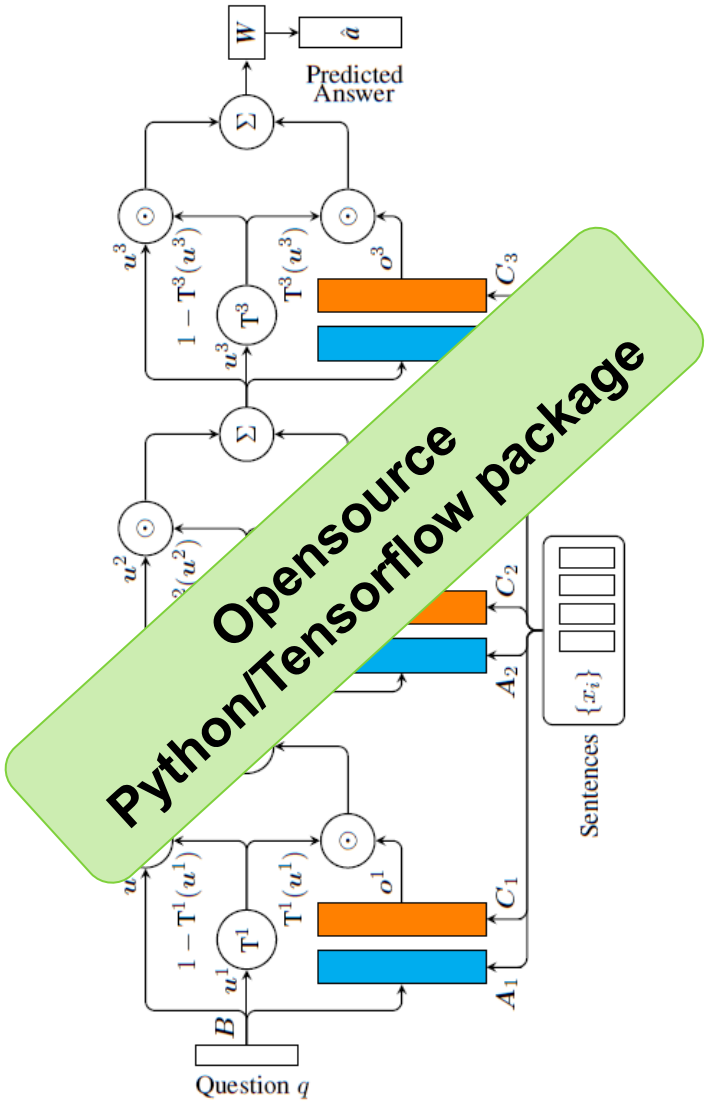


Figure 2: 3 most frequently observed gate value $\mathrm{T}^k(\boldsymbol{u}^k)$ patterns on T6 of the `Dialog bAbI` dataset

Gated Memory Network on reasoning tasks

Performance on bAbI 20 Tasks

Task	10k			
	MemN2N	Our MemN2N	GMemN2N +global	GMemN2N +hop
1: 1 supporting fact	**100.0**	**100.0**	**100.0**	**100.0**
2: 2 supporting facts	99.7	99.7	**100.0**	**100.0**
3: 3 supporting facts	90.7	89.1	94.7	**95.5**
4: 2 argument relations	**100.0**	**100.0**	**100.0**	**100.0**
5: 3 argument relations	99.4	99.4	**99.9**	99.8
6: yes/no questions	**100.0**	**100.0**	96.7	**100.0**
7: counting	96.3	96.8	96.7	**98.2**
8: lists/sets	99.2	98.1	**99.9**	99.7
9: simple negation	99.2	99.1	**100.0**	**100.0**
10: indefinite knowledge	97.6	98.0	**99.9**	99.8
11: basic coreference	**100.0**	**100.0**	**100.0**	**100.0**
12: conjunction	**100.0**	**100.0**	**100.0**	**100.0**
13: compound coreferenc	**100.0**	**100.0**	**100.0**	**100.0**
14: time reasoning	**100.0**	**100.0**	**100.0**	**100.0**
15: basic deduction	**100.0**	**100.0**	**100.0**	**100.0**
16: basic induction	99.6	**100.0**	**100.0**	**100.0**
17: positional reasoning	59.3	62.1	68.8	**72.2**
18: size reasoning	93.3	**93.4**	92.0	91.5
19: path finding	33.5	47.2	54.8	**69.0**
20: agent's motivation	**100.0**	**100.0**	**100.0**	**100.0**
Average	93.4	94.1	95.2	**96.3**



xerox

Gated Memory Network fitted-Q Iteration

Non-Markovian decision process

Current issues of policy learning

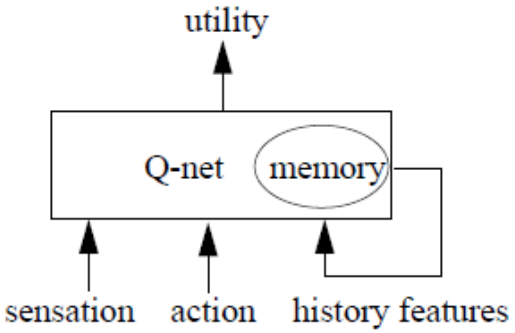
- Long range dependencies
- State Feature selection
- Selective memory through attention

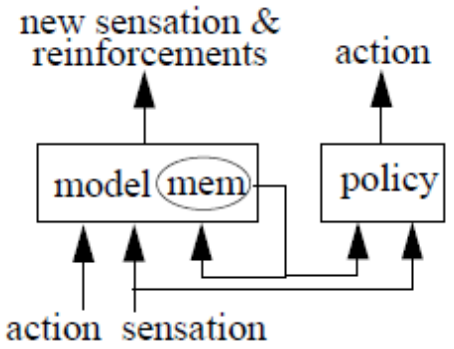
E2E Attention mechanism

- Trainable via back propagation
- Unbounded memory
- Data-type adaptable (encoding …)

Applications

- E2E Reward driven Transactional agent
- Negotiation agent
- Recommendation systems





[4] **Memory Approaches to Reinforcement Learning in non-Markovian Domains**, Tom M. Mitchell and al, 1992

xerox

Gated Memory Network fitted-Q Iteration

Non-Markovian decision process

Current issues of policy learning

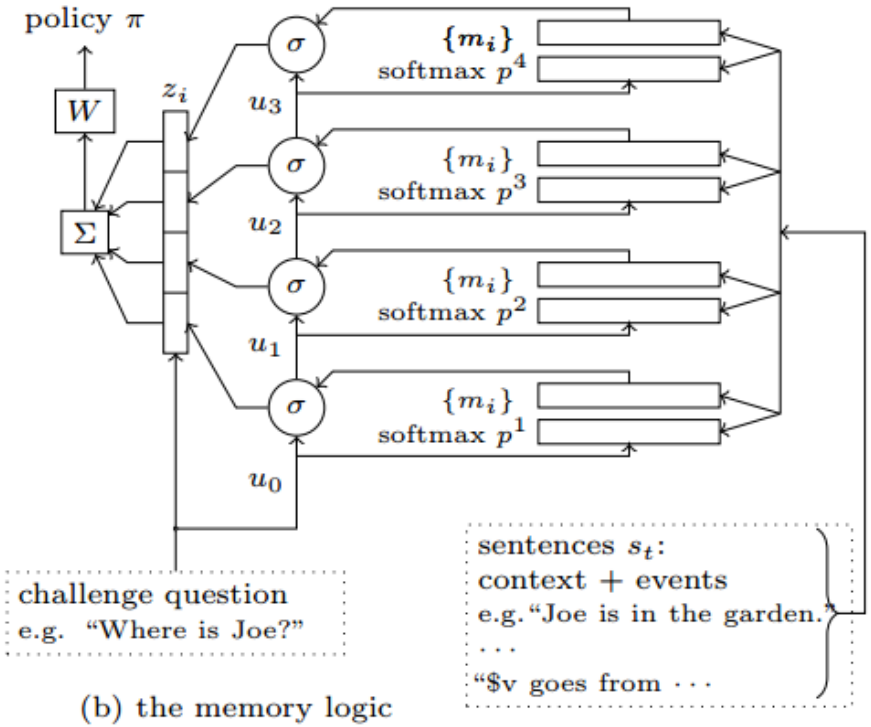
- Long range dependencies
- State Feature selection
- Selective memory through attention

E2E Attention mechanism

- Trainable via back propagation
- Unbounded memory
- Data-type adaptable (encoding …)

Applications

- E2E Reward driven Transactional agent
- Negotiation agent
- Recommendation systems



(b) the memory logic

[3] G. Tesauro, S. Singh and al , **Learning to Query, Reason, and Answer Questions On Ambiguous Texts,** ICLR submission 2017

xerox

Gated Memory Network fitted-Q Iteration

Non-Markovian decision process

Current issues of policy learning

- Long range dependencies
- State Feature selection
- Selective memory through attention

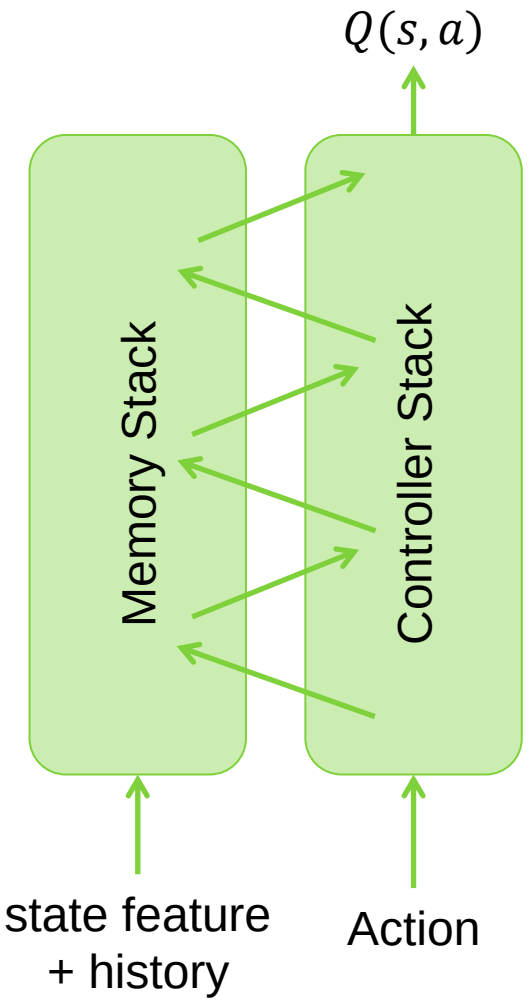
E2E Attention mechanism

- Trainable via back propagation
- Unbounded memory
- Data-type adaptable (encoding …)

Applications

- E2E Reward driven Transactional agent
- Negotiation agent
- Recommendation systems

{Work in progress} – Students wanted…



xerox

Thanks !