ACL2020

Semantic Graphs for Generating Deep Questions

Liangming Pan[1,2] **Yuxi Xie**[3] **Yansong Feng**[3]
Tat-Seng Chua[2] **Min-Yen Kan**[2]
[1]NUS Graduate School for Integrative Sciences and Engineering
[2]School of Computing, National University of Singapore, Singapore
[3]Wangxuan Institute of Computer Technology, Peking University
e0272310@u.nus.edu, {xieyuxi, fengyansong}@pku.edu.cn
{dcscts@, kanmy@comp.}nus.edu.sg

Deep Question Generation

Input Sentence:
Oxygen is used in cellular respiration and released by photosynthesis, which uses the energy of sunlight to produce oxygen from water.
Question: What life process produces oxygen in the presence of light?
Answer: Photosynthesis

a) Example of **Shallow** Question Generation

Input Paragraph A: **Pago Pago International Airport**
Pago Pago International Airport, also known as Tafuna Airport, is a public airport located 7 miles (11.3 km) southwest of the central business district of Pago Pago, in the village and plains of Tafuna on the island of Tutuila in American Samoa, an unincorporated territory of the United States.
Input Paragraph B: **Hoonah Airport**
Hoonah Airport is a state-owned public-use airport located one nautical mile (2 km) southeast of the central business district of Hoonah, Alaska.
Question: Are Pago Pago International Airport and Hoonah Airport both on American territory?
Answer: Yes

b) Example of **Deep** Question Generation

Figure 1: Examples of shallow/deep QG. The evidence needed to generate the question are highlighted.

Challenges

- First, DQG requires document-level understanding, which may introduce long-range dependencies when the passage is long.

- Second, we must be able to select relevant contexts to ask meaningful questions.

- Third, we need to ensure correct reasoning over multiple pieces of information so that the generated question is answerable by information in the passage.

•

Novel mechanisms

- Proposing a novel graph encoder, which incorporates an attention mechanism into the Gated Graph Neural Network.

- Enhancing the word-level passage embeddings and the node-level semantic graph representations to obtain an unified semantic-aware passage representations for question decoding

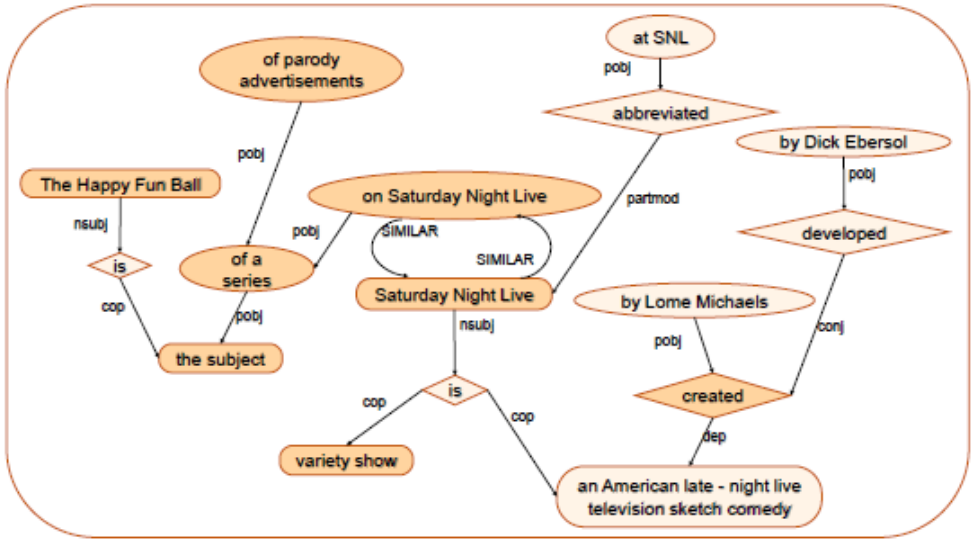
- Introducing an auxiliary content selection task that jointly trains with question decoding, which assists the model in selecting relevant contexts in the semantic graph to form a proper reasoning chain

Model

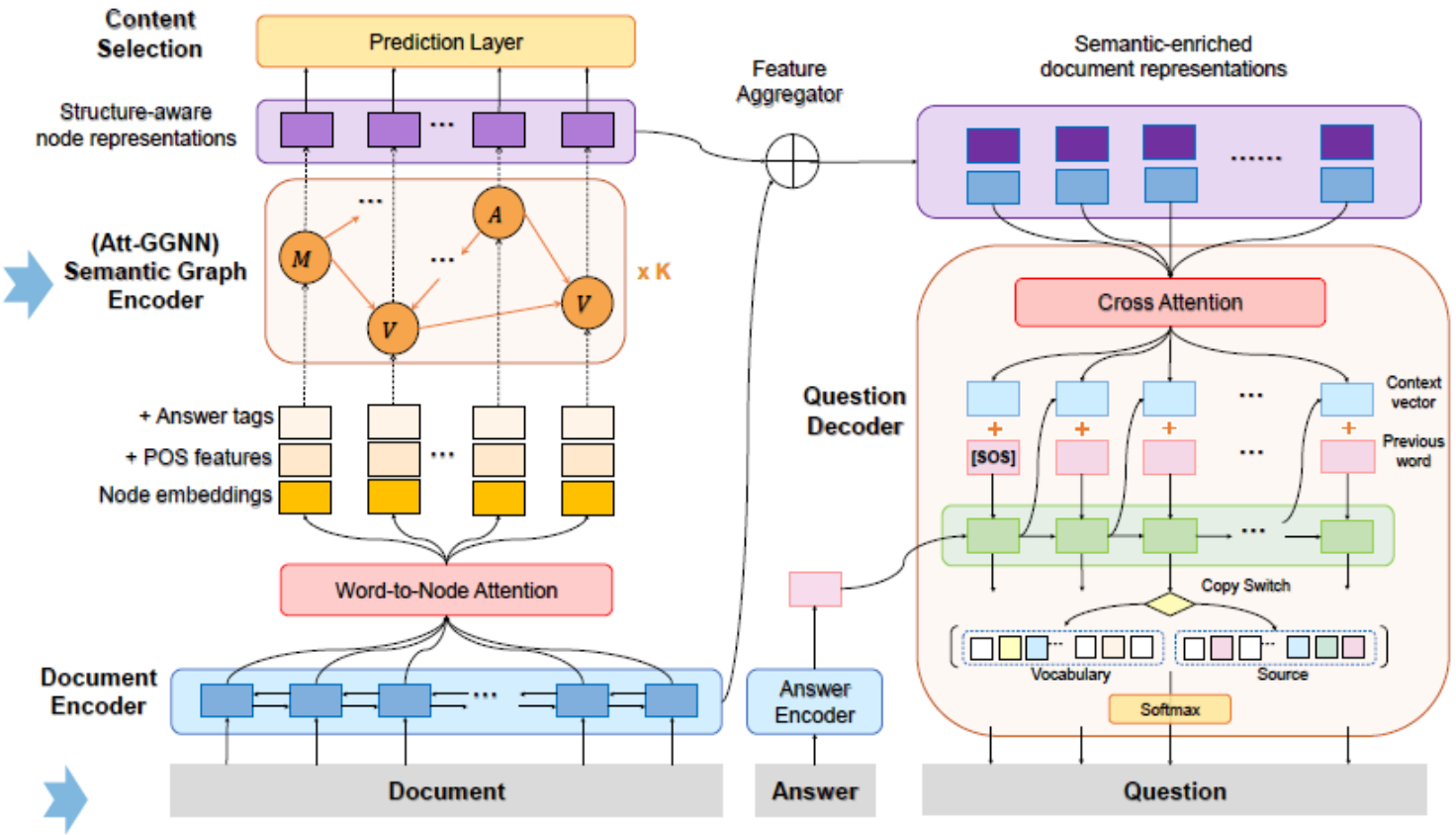
- **A document encoder** to encode the input document
- **A semantic graph encoder** to embed the document-level semantic graph via Att-GGNN
- **A content selector** to select relevant question-worthy contents from the semantic graph
- **A question decoder** to generate question from the semantic-enriched document representation

Framework

Question The "**Happy Fun Ball**" was the **subject** of a **series** of **parody advertisements** on a **show created** by who?
Answer Lorne Michaels



Evidence #1 The "**Happy Fun Ball**" was the **subject** of a **series** of **parody advertisements** on "**Saturday Night Live**" .
Evidence #2 **Saturday Night Live** (abbreviated as SNL) is an American late - night live television sketch comedy and variety **show created** by Lorne Michaels and developed by Dick Ebersol .



Semantic Graph Construction

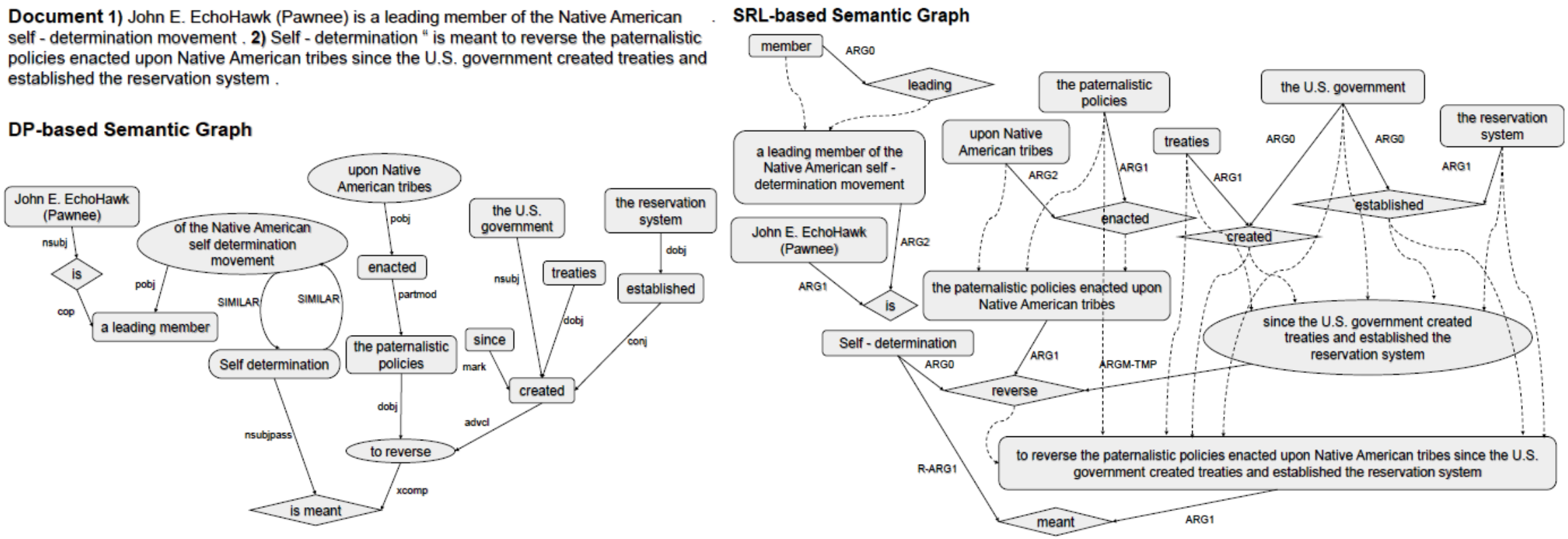


Figure 4: An example of constructed DP- and SRL- based semantic graphs, where ⇢ indicates *CHILD* relation, and rectangular, rhombic and circular nodes represent arguments, verbs and modifiers respectively.

Baseline

• **Seq2Seq + Attn** (Bahdanau et al., 2014): the basic Seq2Seq model with attention, which takes the document as input to decode the question.

• **NQG++** (Zhou et al., 2017): which enhances the Seq2Seq model with a feature-rich encoder containing answer position, POS and NER information.

• **ASs2s** (Kim et al., 2019): learns to decode questions from an answer-separated passage encoder together with a keyword-net based answer encoder.

• **S2sa-at-mp-gsa** (Zhao et al., 2018): an enhanced Seq2Seq model incorporating gated self-attention and maxout-pointers to encode richer passage-level contexts (B4 in Table 1). We also implement a ver-

• **CGC-QG** (Liu et al., 2019a): another enhanced Seq2Seq model that performs word-level content selection before generation; *i.e.*, making decisions on which words to generate and to copy using rich syntactic features, such as NER, POS, and DEP.

Results

Model		BLEU1	BLEU2	BLEU3	BLEU4	METEOR	ROUGE-L
Baselines	B1. Seq2Seq + Attn	32.97	21.11	15.41	11.81	18.19	33.48
	B2. NQG++	35.31	22.12	15.53	11.50	16.96	32.01
	B3. ASs2s	34.60	22.77	15.21	11.29	16.78	32.88
	B4. S2s-at-mp-gsa	35.36	22.38	15.88	11.85	17.63	33.02
	B5. S2s-at-mp-gsa (+cov, +ans)	38.74	24.89	17.88	13.48	18.39	34.51
	B6. CGC-QG	31.18	22.55	17.69	14.36	**25.20**	**40.94**
Proposed	P1. SRL-Graph	40.40	26.83	19.66	15.03	19.73	36.24
	P2. DP-Graph	**40.55**	**27.21**	**20.13**	**15.53**	20.15	36.94
Ablation	A1. *-w/o* Contexts	36.48	20.56	12.89	8.46	15.43	30.86
	A2. *-w/o* Semantic Graph	37.63	24.81	18.14	13.85	19.24	34.93
	A3. *-w/o* Multi-Relation & Attention	38.50	25.37	18.54	14.15	19.15	35.12
	A4. *-w/o* Multi-Task	39.43	26.10	19.14	14.66	19.25	35.76

Human Evaluation

Model	Short Contexts			Medium Contexts			Long Contexts			Average		
	Flu.	Rel.	Cpx.	Flu.	Rel.	Cpx.	Flu.	Rel.	Cpx.	Flu.	Rel.	Cpx.
B4. S2sa-at-mp-gsa	3.76	4.25	3.98	3.43	4.35	4.13	3.17	3.86	3.57	3.45	4.15	3.89
B6. CGC-QG	3.91	4.43	3.60	3.63	4.17	4.10	**3.69**	3.85	**4.13**	3.75	4.15	3.94
A2. *-w/o* Semantic Graph	4.01	4.43	4.15	3.65	4.41	4.12	3.54	3.88	3.55	3.73	4.24	3.94
A4. *-w/o* Multi-Task	4.11	4.58	4.28	3.81	4.27	**4.38**	3.44	3.91	3.84	3.79	4.25	4.17
P2. DP-Graph	**4.34**	**4.64**	**4.33**	**3.83**	**4.51**	4.28	3.55	**4.08**	4.04	**3.91**	**4.41**	**4.22**
G1. Ground Truth	4.75	4.87	4.74	4.65	4.73	4.73	4.46	4.61	4.55	4.62	4.74	4.67

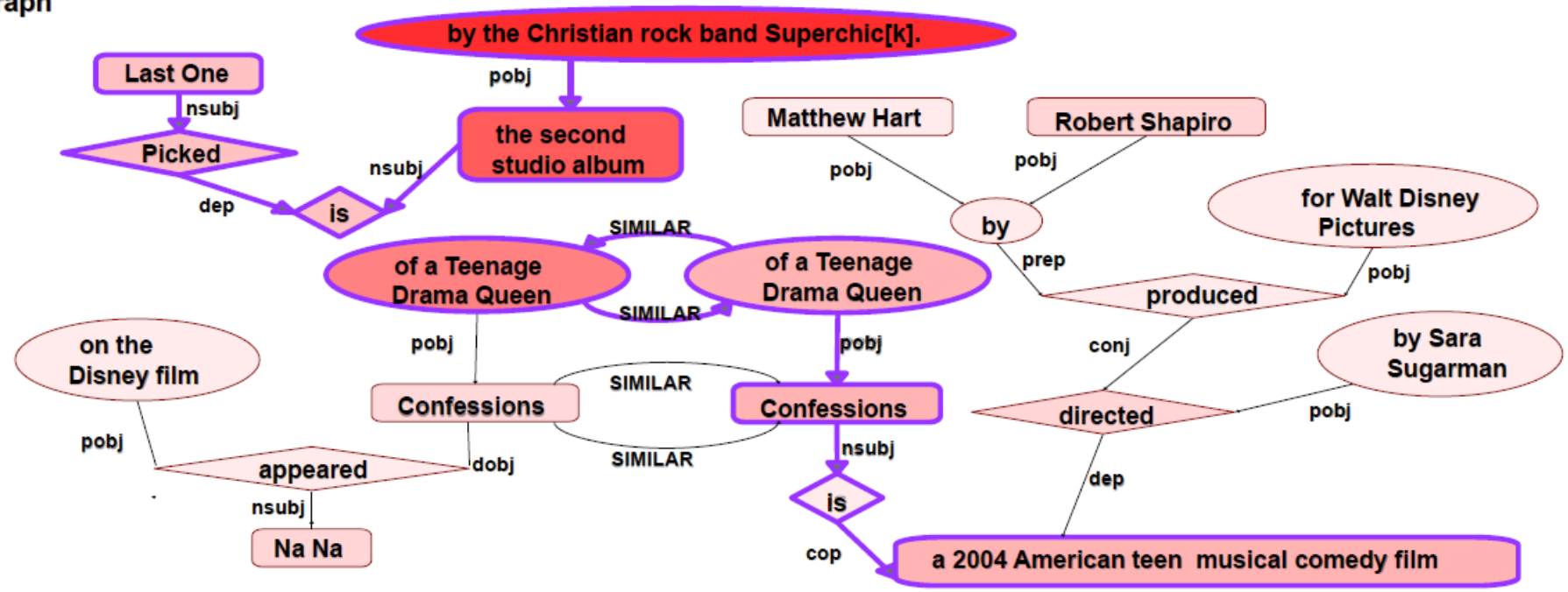
Error Analysis

Types	Examples		S2sa-at-mp-gsa	CGC-QG	DP-Graph
Correct	(Pred.) (G.T.)	Between Kemess Mine and Colomac Mine, which mine was operated earlier? What mine was operated at an earlier date, Kemess Mine or Colomac Mine?	56.5%	52.9%	67.4%
Semantic Error	(Pred.) (G.T.)	Lawrence Ferlinghetti is an American poet, **he is** a short story **written by who**? Lawrence Ferlinghetti is an American poet, he wrote a short story named what ?	17.7%	26.4%	8.3%
Answer Revealing	(Pred.) (G.T.)	What is the release date of this game **released on 17 October 2006**? What is the release date of this game named Hurricane?	2.1%	5.7%	1.4%
Ghost Entity	(Pred.) (G.T.)	When was the video game on which **Michael Gelling** plays Dr. Promoter? When was the video game on which Drew Gelling plays Dr. Promoter?	6.8%	0.7%	4.9%
Redundant	(Pred.) (G.T.)	What town did **Walcha and Walcha** belong to? What town did Walcha belong to?	16.3%	14.3%	13.9%
Unanswerable	(Pred.) (G.T.)	What is **the population of** the city Barack Obama was born? What was the ranking of the population of the city Barack Obama was born in 1999?	8.2%	18.6%	8.3%

Case

Passage 1) Last One Picked is the second studio album by the Christian rock band Superchic[k].
2) " Na Na " appeared on the Disney film , " Confessions of a Teenage Drama Queen " .
3) Confessions of a Teenage Drama Queen is a 2004 American teen musical comedy film directed by Sara Sugarman and produced by Robert Shapiro and Matthew Hart for Walt Disney Pictures .

Semantic Graph



Question(Ours) What is the name of the American teen musical comedy in which the second studio album by the Christian rock band Superchic[k]. " Na Na appeared ?
Question(Humans) Which song by Last One Picked appeared in a 2004 American teen musical comedy film directed by Sara Sugarman ?
Question(Baseline) Who directed the 2004 American musical comedy Na in the film confessions " Na " ?
Question (CGC) Last One Picked is the second studio album by which 2004 American teen musical comedy film directed by Sara Sugarman and produced by Robert Shapiro and Matthew Hart for Walt Disney Pictures ?