

A Mixed Hierarchical Attention based Encoder-Decoder Approach for Standard Table Summarization

NAACL

2018

Parag Jain¹, Anirban Laha¹, Karthik Sankaranarayanan¹, Preksha Nema², Mitesh M. Khapra², Shreyas Shetty²

¹IBM Research, India

²Indian Institute Of Technology, Madras

{pajain34,anirlaha,kartsank}@in.ibm.com

{preksha,miteshk,shshett}@cse.iitm.ac.in

Overview

- **Background:** In most practical applications such as finance, healthcare or weather report, data is stored in a form of a table with known schema. Summarizing such data instances into natural language is an important problem.
- **Objective:** Generate abstractive summaries of tables conforming to a predefined fixed schema.
- **Results:** Our experiments on the publicly available WEATHERGOV dataset (Liang et al., 2009) show around 18 BLEU (~ 30%) improvement over the current state-of-the-art.

Background and Motivation

Example:

Input table 1

TYPE	TIME	MIN	MAX	MEAN	MODE	MB100-4	MB20-2
temperature	17-30	33	60	44	-	-	-
windChill	17-30	0	47	31	-	-	-
windSpeed	17-30	5	9	7	-	-	0-10
windDir	17-30	-	-	-	NNW	-	-
gust	17-30	0	0	0	-	-	-
skyCover	17-30	-	-	-	-	0-25	-
skyCover	17-21	-	-	-	-	0-25	-
skyCover	17-26	-	-	-	-	0-25	-
skyCover	21-30	-	-	-	-	0-25	-
skyCover	26-30	-	-	-	-	0-25	-
precipPotential	17-30	0	0	0	-	-	-

Reference summary 1

Clear, with a low around 33 . North wind between 5 and 9 mph becoming calm .

Input table 2

TYPE	TIME	MIN	MAX	MEAN	MODE	MB100-4	MB20-2
temperature	6-21	35	53	46	-	-	-
windChill	6-21	0	47	27	-	-	-
windSpeed	6-21	6	8	7	-	-	0-10
windDir	6-21	-	-	-	WSW	-	-
gust	6-21	0	0	0	-	-	-
skyCover	6-21	-	-	-	-	75-100	-
skyCover	6-9	-	-	-	-	75-100	-
skyCover	6-13	-	-	-	-	50-75	-
skyCover	9-21	-	-	-	-	75-100	-
skyCover	13-21	-	-	-	-	75-100	-
precipPotential	6-21	4	48	17	-	-	-
rainChance	13-21	-	-	-	Che	-	-

Reference summary 2

A slight chance of showers after 3pm . Mostly cloudy , with a high near 54 .Southwest wind between 6 and 8 mph . Chances of precipitation is 20 % .

Summary:

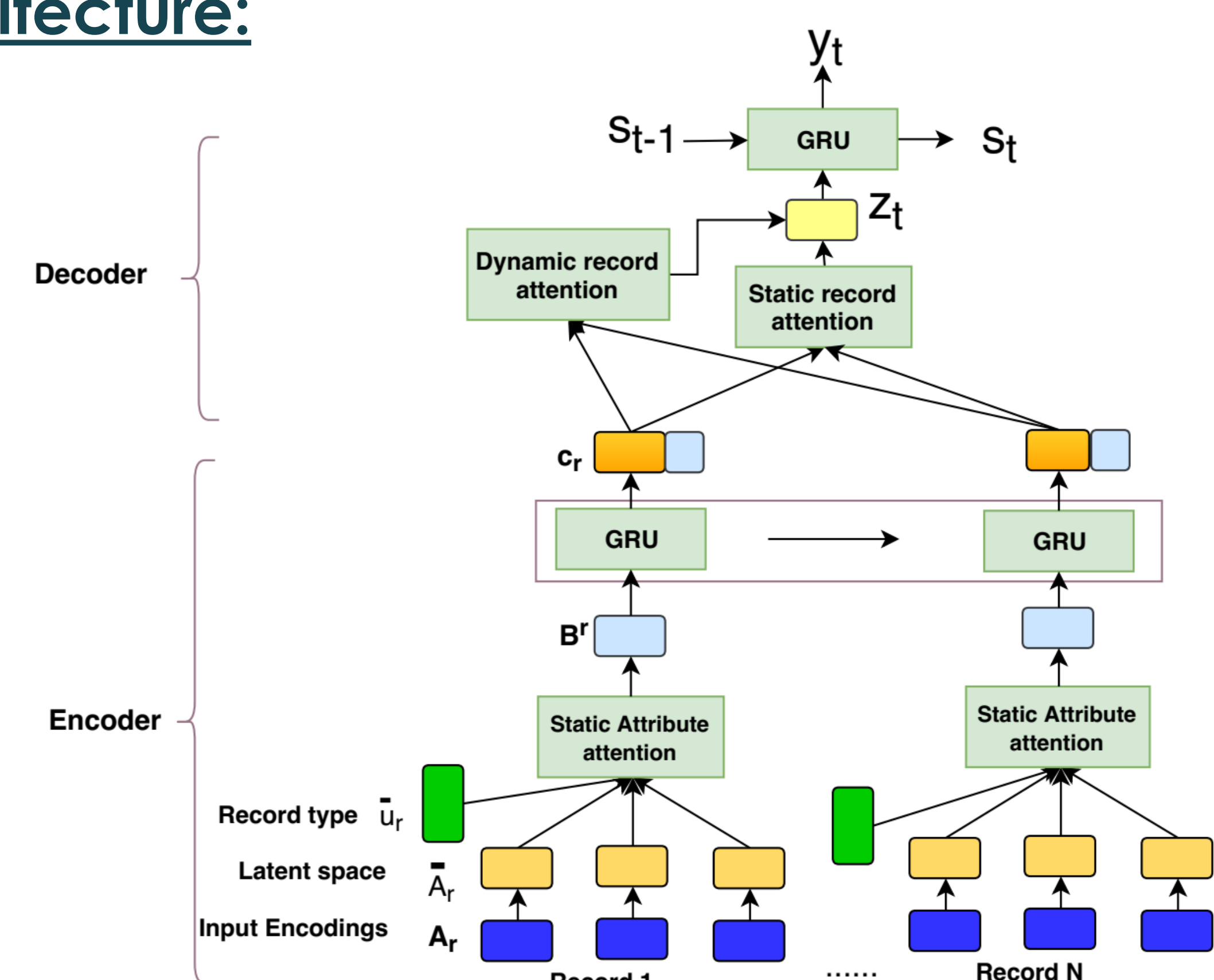
- Structured data summarization involves generation of natural language summaries from structured input data.
- Tabular data summarization is challenging as the schema or structure can vary across tables.
- We formulate the standard table summarization problem, where the tables conform to a fixed schema.
- We propose a novel mixed hierarchical attention based approach suitable for tabular inputs arising from a fixed schema.

Some relevant structured data summarization task:

- Summarizing Wikipedia infobox, restricted to biography domain – Lebre et al., 2016
- Summarizing source code using a neural attention model – Iyer et al., 2016
- Generating natural language descriptions from Ontologies – Androutsopoulos et al., 2014; Colin et al., 2016
- Challenges in data-to-document generation – Wiseman et al., 2017

Mixed Hierarchical Attention Model

Architecture:



Incorporates static attention and dynamic attention at two different levels of the encoder.

Record level: The attention over record representations is dynamic and changes with each decoder time step.

Attribute level: Since the schema is fixed, static attention is used.

Complexity: $O(TM + T')$ compared to fully dynamic attention at both level $O(TMT')$. T is the number of records, M is the number of attributes and T' is the number of decoder steps.

Dataset & Results

- **WeatherGov dataset:** 25000/3528/1000 (train/dev/test) instances
- **Note:** We haven't exploited the soft alignments present in the dataset

Model	sBleu	cBleu	Rouge-L
KL	36.54	-	-
AKL	38.40	51.50	-
MBW	61.0	70.4	-
NHM	76.2	85.0	86.4
MHAM	79.3	87.0	88.5

KL : Konstas et al., 2013

AKL : Angeli et al., 2010

MBW: Mei et al., 2016

NHM: Non hierarchical version

MHAM: Mixed hierarchical attention model

Example summaries

TYPE	TIME	MIN	MAX	MEAN	MODE	MB100-4	MB20-2
temperature	17-30	27	36	29	-	-	-
windChill	17-30	14	26	18	-	-	-
windSpeed	17-30	14	20	16	-	-	10-20
windDir	17-30	-	-	-	SSW	-	-
gust	17-30	0	0	0	-	-	-
skyCover	17-30	-	-	-	-	75-100	-
skyCover	17-21	-	-	-	-	75-100	-
skyCover	17-26	-	-	-	-	75-100	-
skyCover	21-30	-	-	-	-	75-100	-
skyCover	26-30	-	-	-	-	75-100	-
precipPotential	17-30	26	58	43	-	-	-
rainChance	17-21	-	-	-	Lkly	-	-
snowChance	17-30	-	-	-	Che	-	-
snowChance	17-26	-	-	-	Lkly	-	-

Reference: Rain and snow likely , becoming all snow after 8pm . Cloudy , with a low around 22 . South southwest wind around 15 mph . Chance of precipitation is 60% . New snow accumulation of less than one inch possible .

NHM: Rain or freezing rain likely before 8pm , then snow after 11pm , snow showers and sleet likely before 8pm , then a chance of rain or freezing rain after 3am . Mostly cloudy , with a low around 27 . South southeast wind between 15 and 17 mph . Chance of precipitation is 80% . Little or no ice accumulation expected . Little or no snow accumulation expected .

MHAM: Snow , and freezing rain , snow after 9pm . Cloudy , with a steady temperature around 23 . Breezy , with a south wind between 15 and 20 mph . Chance of precipitation is 60% . New snow accumulation of around an inch possible .

References:

- Gabor Angeli, Percy Liang, and Dan Klein. 2010. A simple domain-independent probabilistic approach to generation. EMNLP 2010
- Hongyuan Mei, Mohit Bansal, and Matthew R. Walter. 2016. What to talk about and how? selective generation using lstms with coarse-to-fine alignment. NAACL 2016
- Remi Lebre, David Grangier, and Michael Auli. 2016. Neural text generation from structured data with application to the biography domain. EMNLP 2016
- Srinivasan Iyer, Ioannis Konstas, Alvin Cheung, and Luke Zettlemoyer. 2016. Summarizing source code using a neural attention model. ACL 2016
- Ion Androutsopoulos, Gerasimos Lampouras, and Dimitrios Galanis. 2014. Generating natural language descriptions from OWL ontologies: the naturalowl system. CoRR, abs/1405.6164
- Sam Wiseman, Stuart M. Shieber, and Alexander M. Rush. 2017. Challenges in data-to-document generation. EMNLP 2017
- Ioannis Konstas and Mirella Lapata. 2013. Inducing document plans for concept-to-text generation. EMNLP 2013
- Percy Liang, Michael I. Jordan, and Dan Klein. 2009. ACL 2009

IBM Research