Grounded language understanding: Speakers: From the world to language

Christopher Potts

Stanford Linguistics

CS224u: Natural language understanding

Color describer: Task formulation and data

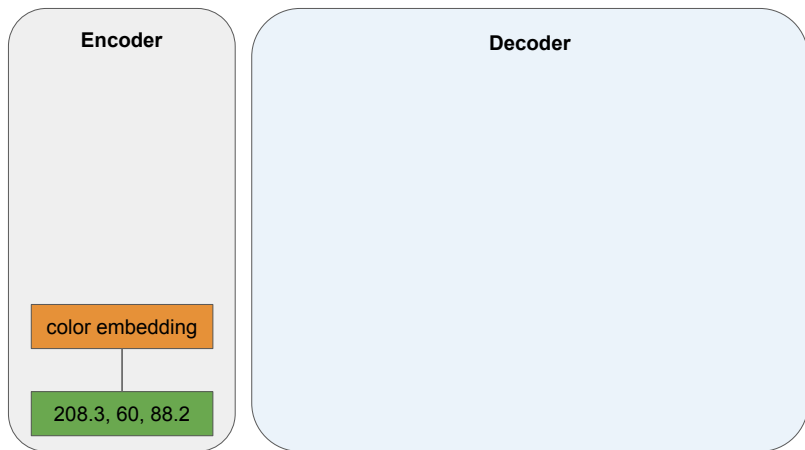| Color	Utterance
	green
	purple
	grape
	turquoise
	moss green
	pinkish purple
	light blue grey
	robin's egg blue
	british racing green
	baby puke green

McMahan and Stone 2015

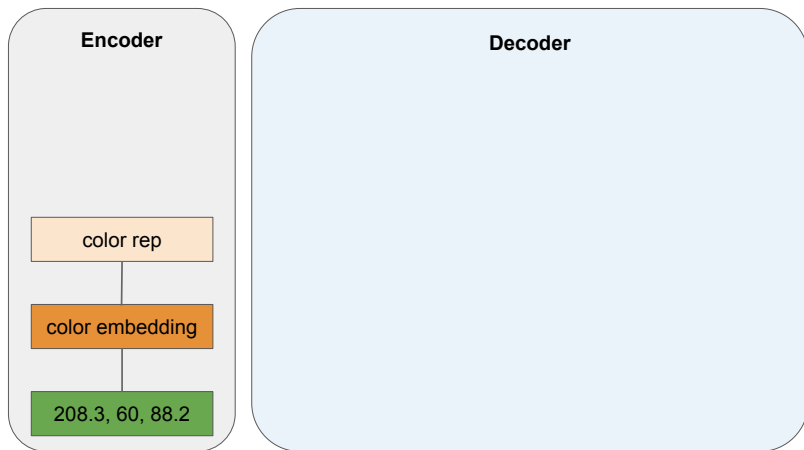# Color describer: Training with *teacher forcing*

Encoder

Decoder

208.3, 60, 88.2

Color describer: Training with *teacher forcing*

Encoder

Decoder

color embedding

208.3, 60, 88.2

Color describer: Training with *teacher forcing*

Encoder

Decoder

color rep

color embedding

208.3, 60, 88.2

Color describer: Training with *teacher forcing*

Encoder

Decoder

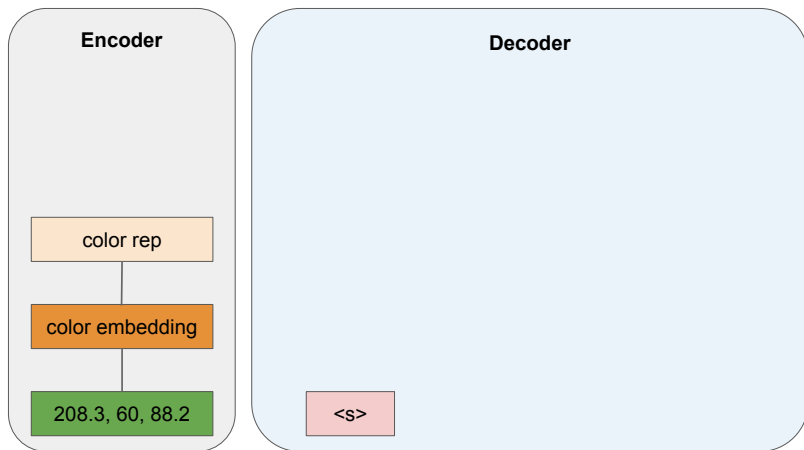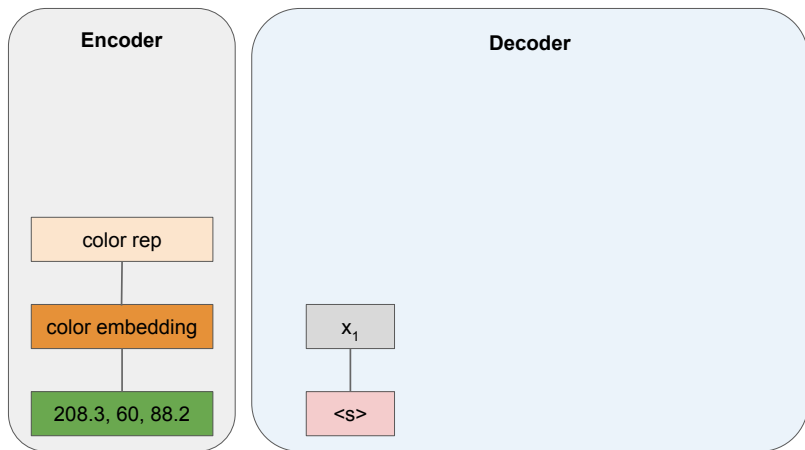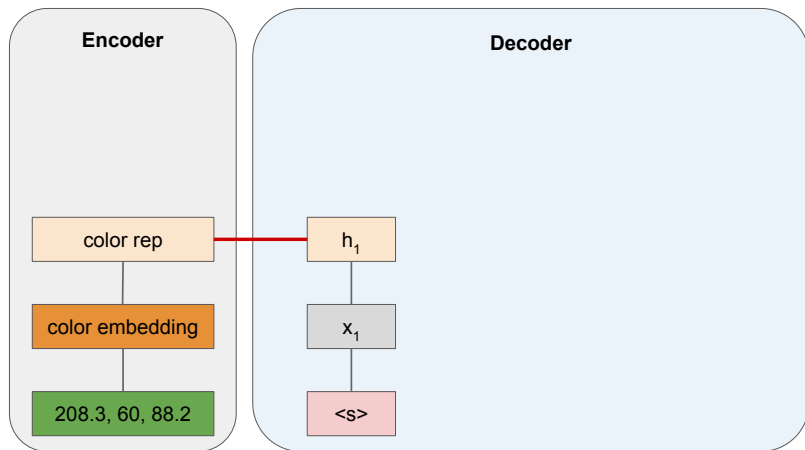color rep

color embedding

208.3, 60, 88.2

\<s\>

Color describer: Training with *teacher forcing*

Color describer: Training with *teacher forcing*
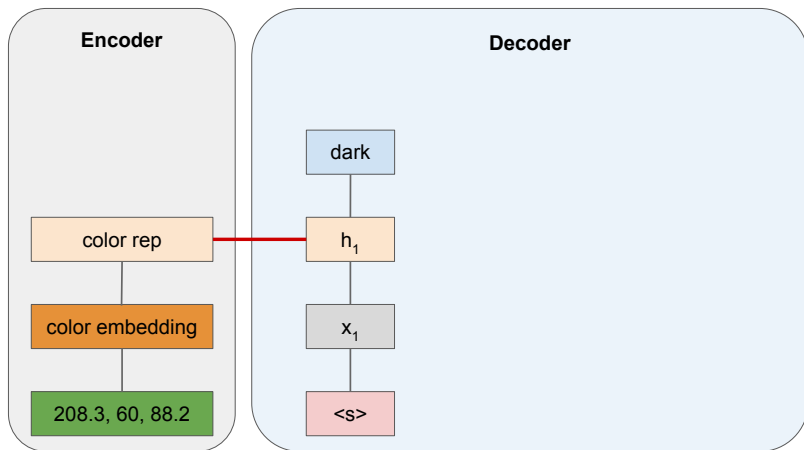
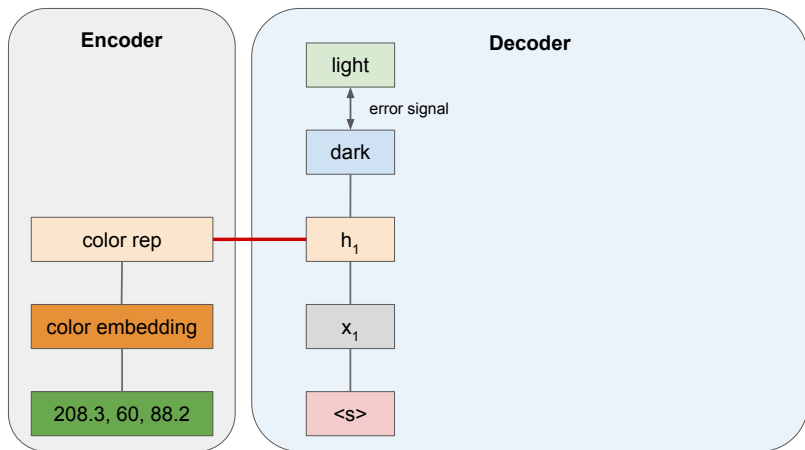# Color describer: Training with *teacher forcing*

Color describer: Training with *teacher forcing*

Color describer: Training with *teacher forcing*

Color describer: Training with *teacher forcing*
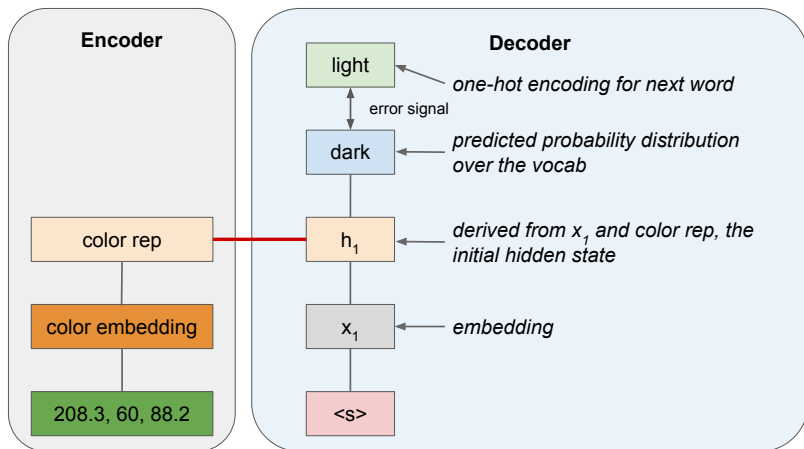
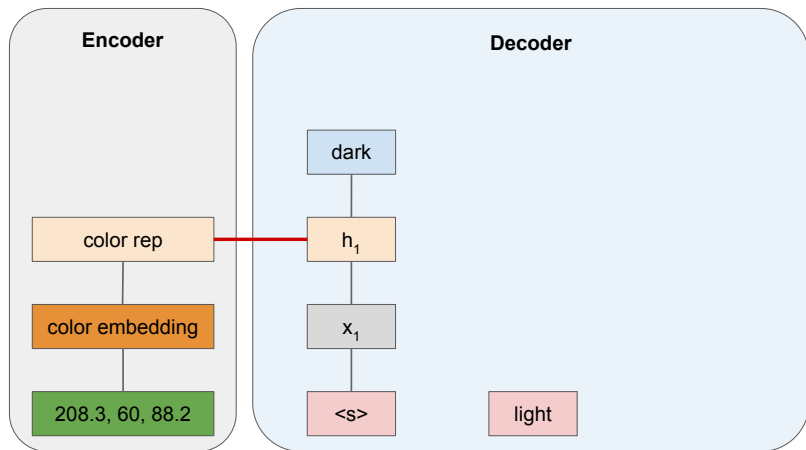# Color describer: Training with *teacher forcing*

Color describer: Training with *teacher forcing*

Color describer: Training with *teacher forcing*

Color describer: Training with *teacher forcing*

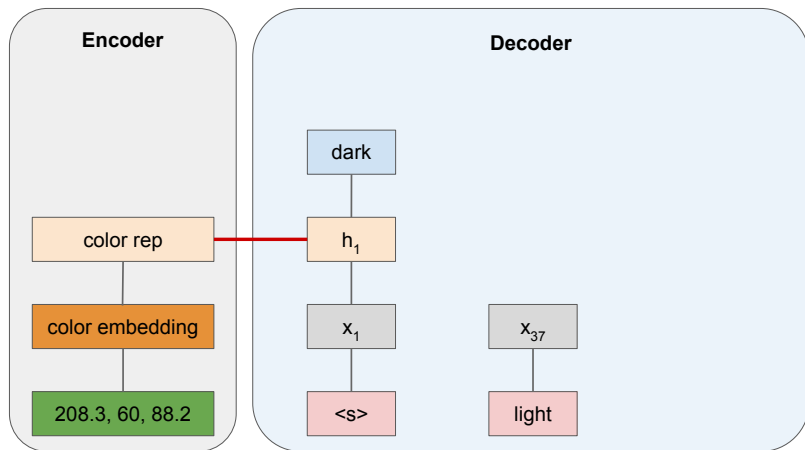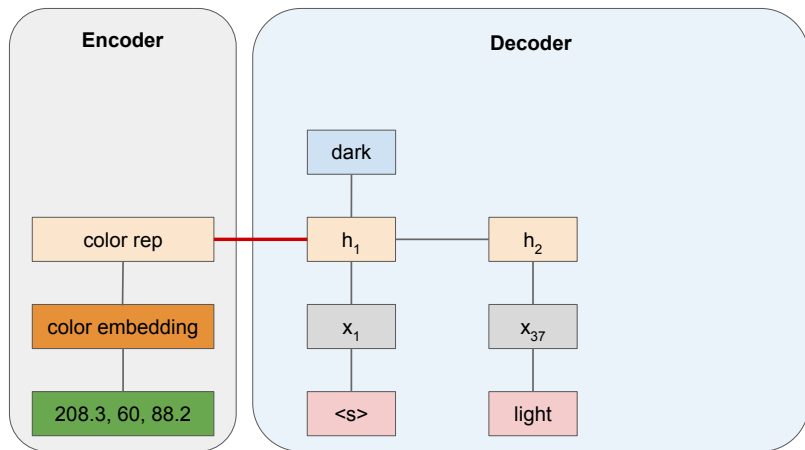# Color describer: Training with *teacher forcing*

Color describer: Training with *teacher forcing*

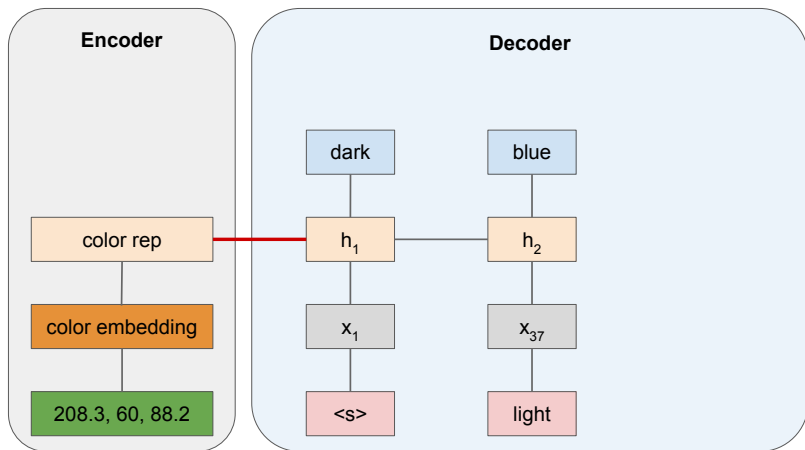# Color describer: Training with *teacher forcing*

Color describer: Training with *teacher forcing*

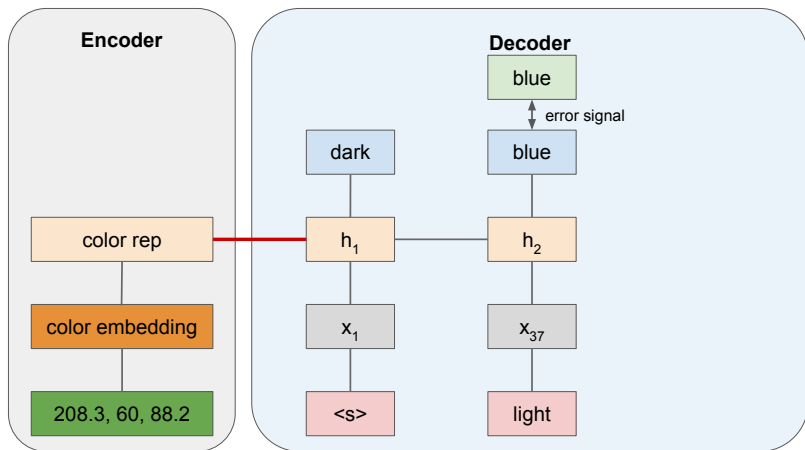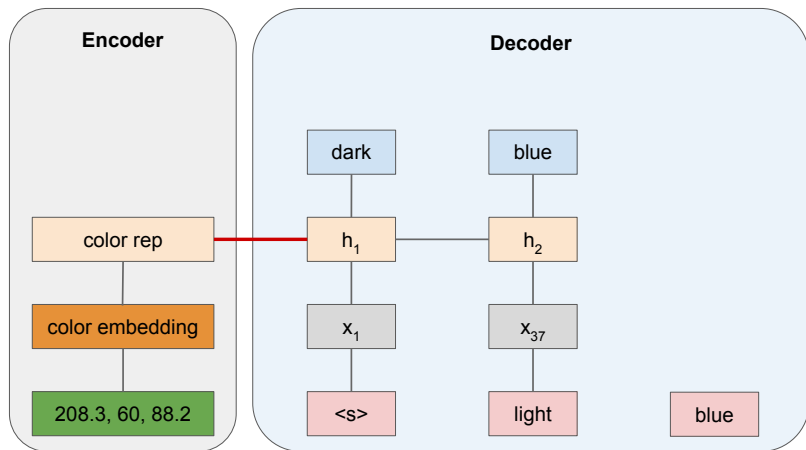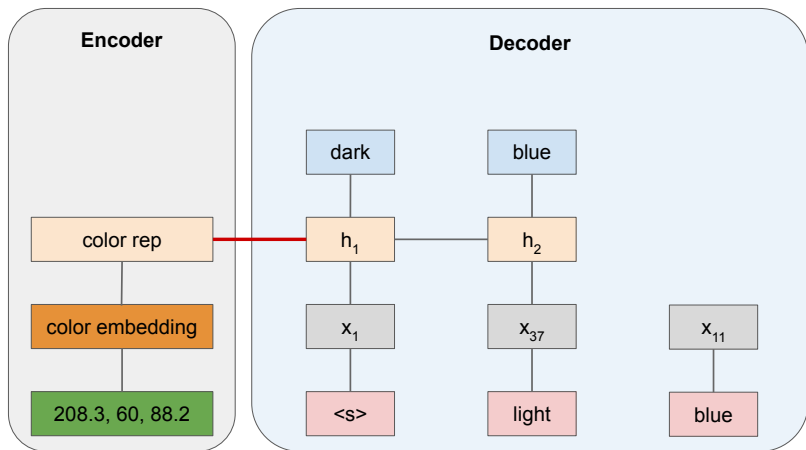# Color describer: Training with *teacher forcing*
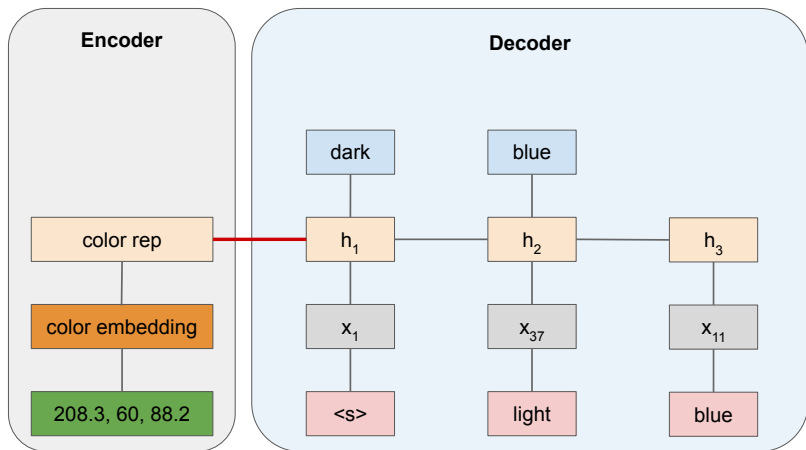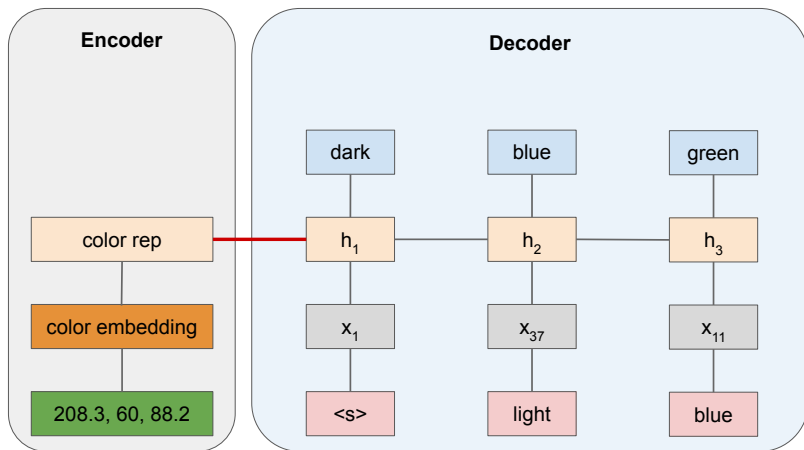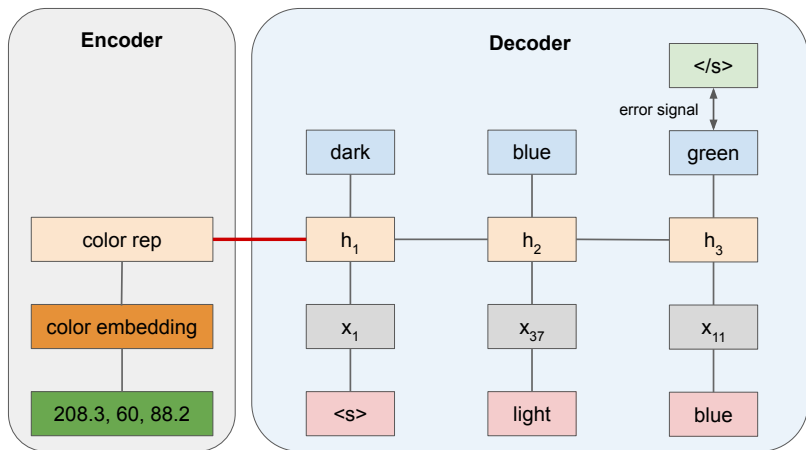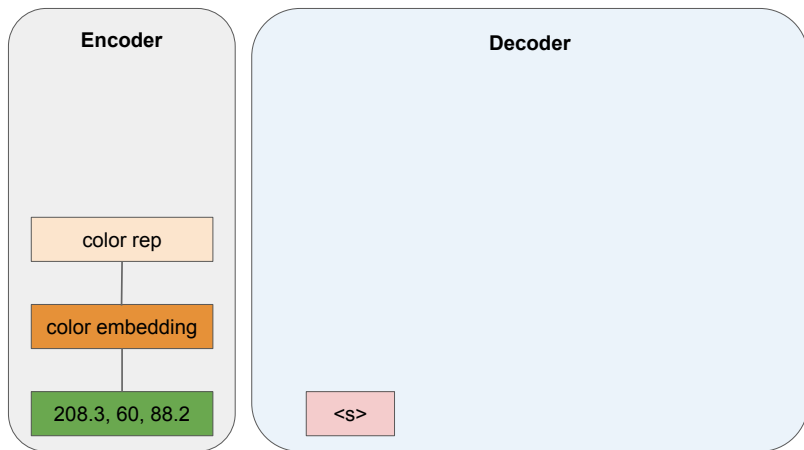
Color describer: Prediction

Color describer: Prediction

Encoder

color rep

color embedding

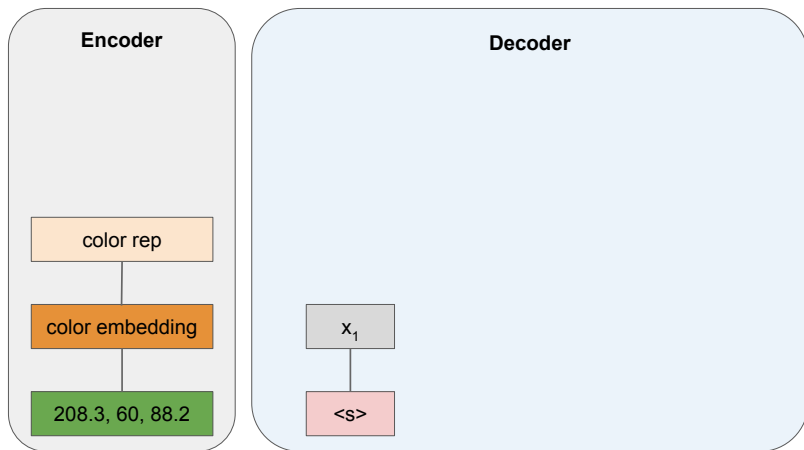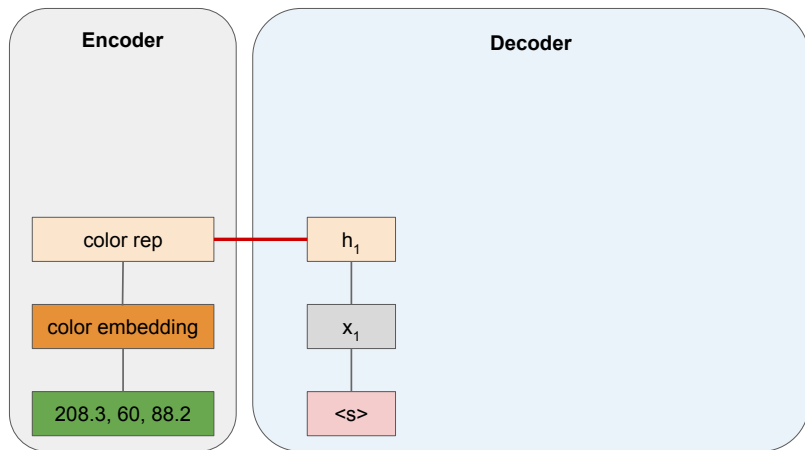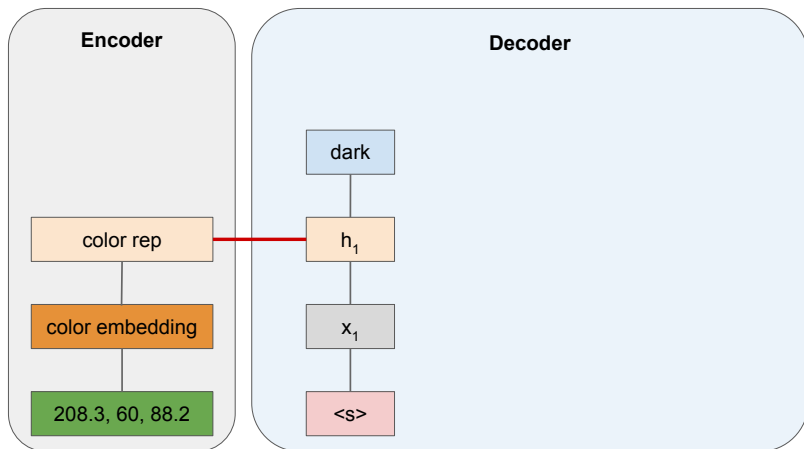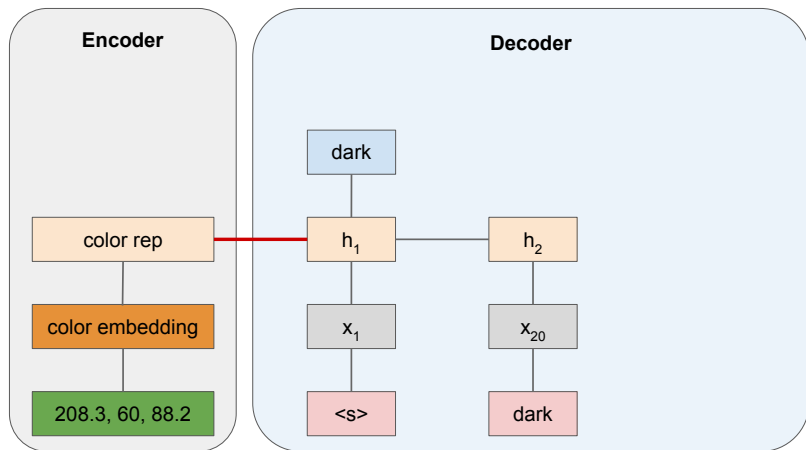208.3, 60, 88.2

Decoder

<s>

Color describer: Prediction

Color describer: Prediction

Color describer: Prediction

Color describer: Prediction

Color describer: Prediction

Color describer: Prediction

Color describer: Prediction

Miscellaneous design choices

- The Encoder and Decoder could have more hidden layers. We would expect the layer counts to match to facilitate the hand-off between Encoder and Decoder, though pooling or copying might work too.

- It seems very common at present for researchers to tie the embedding and classifier parameters (Press and Wolf 2017)

- During training, one might drop teacher forcing a small percentage of the time to encourage the model to explore.

Color describer of Monroe et al. (2016)
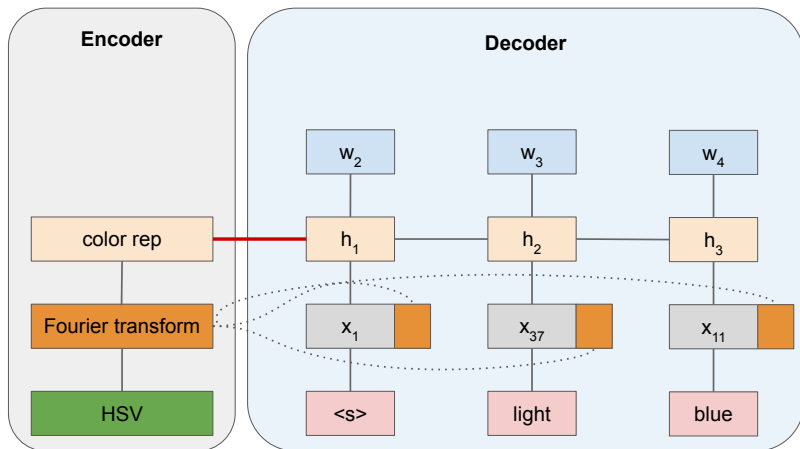
Related tasks

Non-linguistic representation ⇒ Language

- Image captioning

- Scene description

- Visual Question Answering
                    (Image + Question-text ⇒ Answer-text)

- Instruction giving (State ⇒ Language)

- . . .

References I

Brian McMahan and Matthew Stone. 2015. A Bayesian model of grounded color semantics. *Transactions of the Association for Computational Linguistics*, 3:103–115.

Will Monroe, Noah D. Goodman, and Christopher Potts. 2016. Learning to generate compositional color descriptions. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 2243–2248, Stroudsburg, PA. Association for Computational Linguistics.

Ofir Press and Lior Wolf. 2017. Using the output embedding to improve language models. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, pages 157–163, Valencia, Spain. Association for Computational Linguistics.