

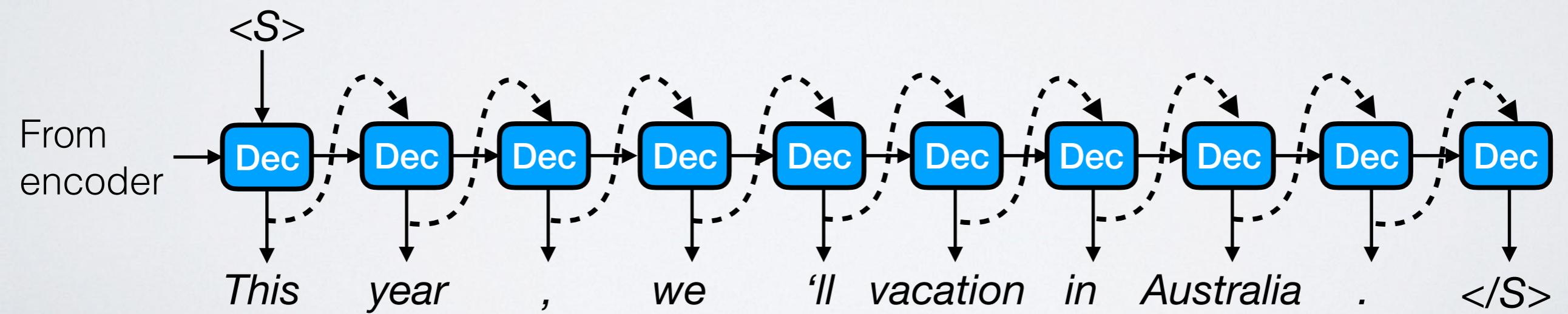
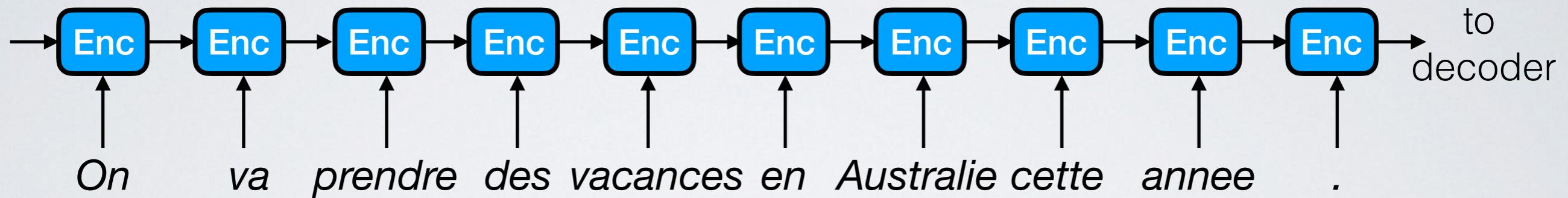
Discourse-Aware Neural Rewards for Coherent Recipe Generation

Antoine Bosselut, Asli Celikyilmaz, Xiaodong He,
Jianfeng Gao, Po-sen Huang, Yejin Choi

Accepted to NAACL 2018



Encoder-decoder approach



Assumptions

- Strong correspondence between input and output
- Most disambiguating information provided at the token level

What happens when these assumptions don't hold true?

Recipe Generation

Title:

Fried

Chicken

Ingredients:

Chicken

Flour

Spices

Recipe Generation

Title:

Fried

Chicken

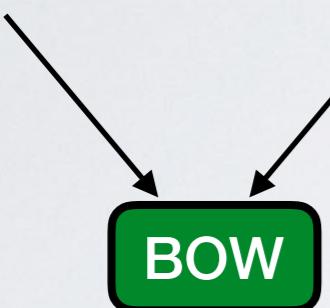
Ingredients:

Chicken

Flour

Spices

BOW



Recipe Generation

Title:

Fried

Chicken

BOW

Ingredients:

Chicken

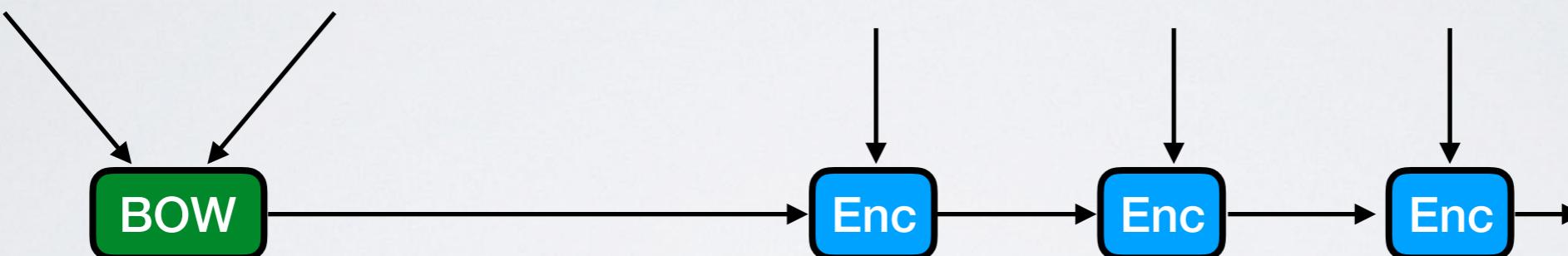
Flour

Spices

Enc

Enc

Enc



Recipe Generation

Title:

Fried

Chicken

BOW

Ingredients:

Chicken

Flour

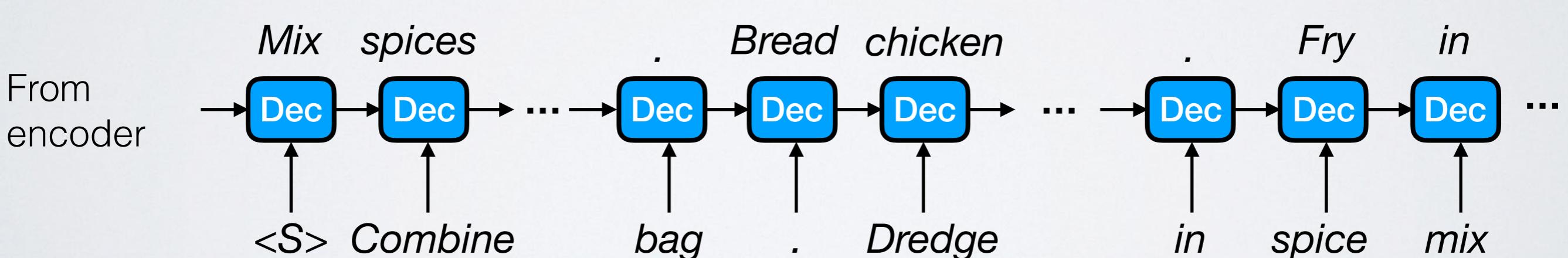
Spices

Enc

Enc

Enc

to decoder



$$L_{mle} = - \sum_{t=1}^T \log P(x_t | x_0, \dots, x_{t-1}, \mathbf{g}, \mathbf{e})$$

No Coherence

sausage sandwiches

- Cut each sandwich in halves.
- Sandwiches with sandwiches.
- Sandwiches, sandwiches, Sandwiches,
sandwiches, sandwiches
- sandwiches, sandwiches, sandwiches,
sandwiches, sandwiches, sandwiches, or
sandwiches or triangles, a griddle, each
sandwich.
- Top each with a slice of cheese, tomato,
and cheese.
- Top with remaining cheese mixture.
- Top with remaining cheese.
- Broil until tops are bubbly and cheese is
melted, about 5 minutes.

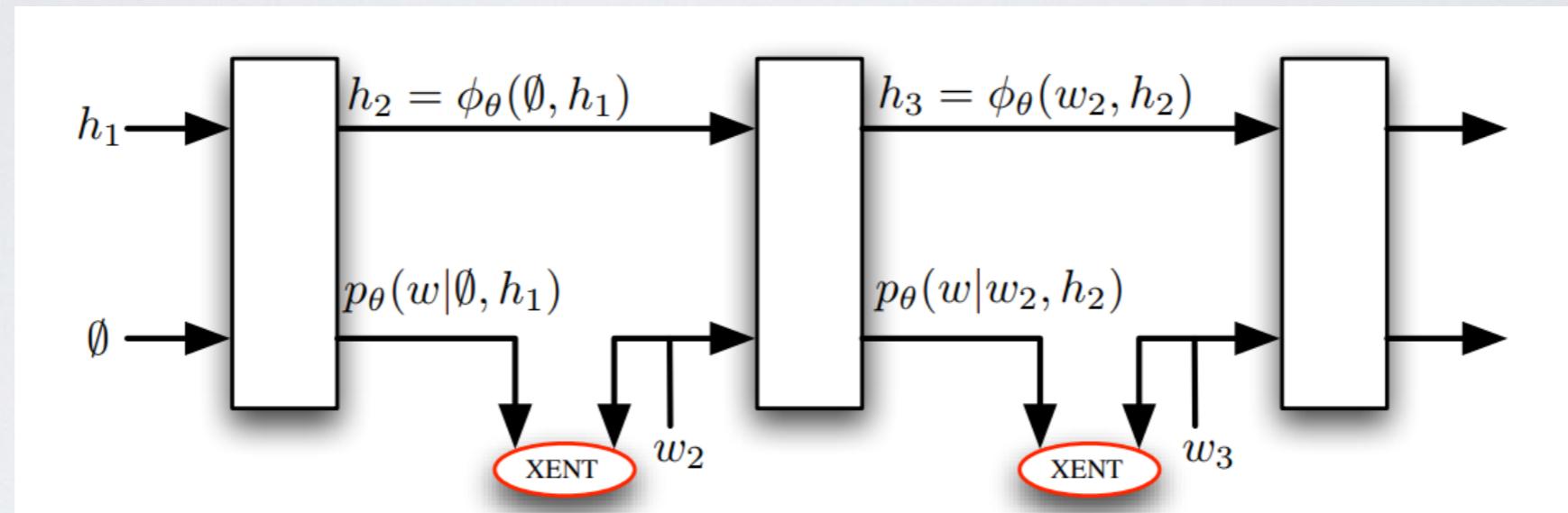
No Specificity

Mix all the ingredients in a bowl.
Stir until thick.
Add flavoring.
Pour into baking pan.
Bake at 350 F for 1 hour.
Let cool.

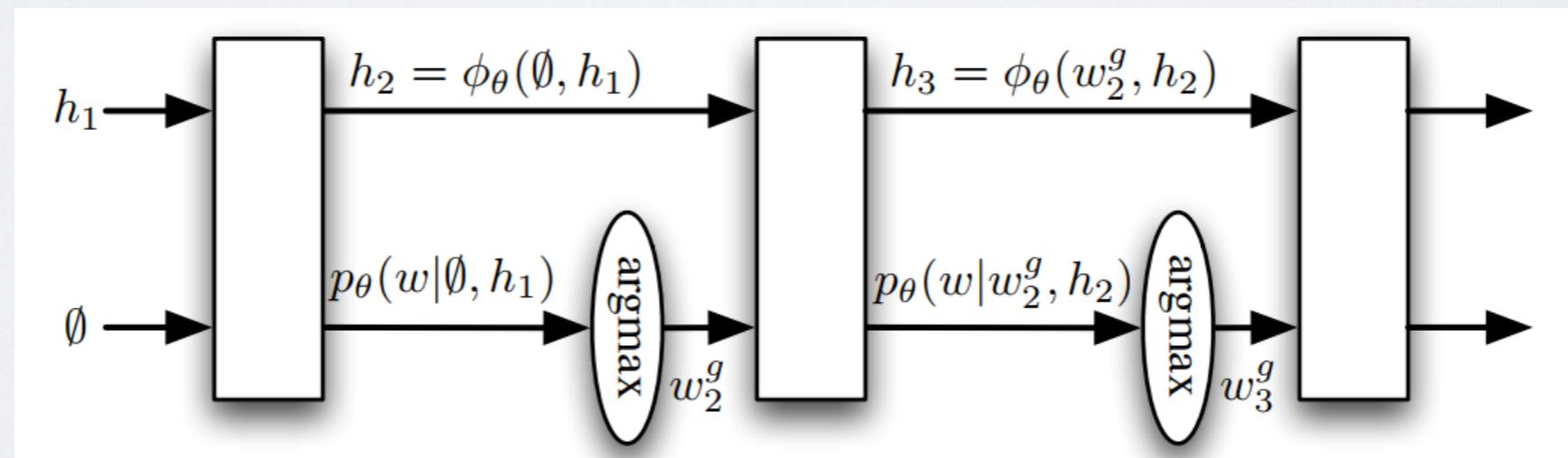
What are we making?

Exposure Bias

Training



Testing



Fixes

- SEARN, DAGGER (Daume et al., 2009, Ross et al., 2011)
- Scheduled Sampling (Bengio et al, 2016)

Fixes

- SEARN, DAGGER (Daume et al., 2009, Ross et al., 2011)
- Scheduled Sampling (Bengio et al, 2016)
- Curriculum Learning (Ranzato et al., 2016)
- Reward Sampling (Pasunuru et al., 2017, Liu et al., 2017)

Fixes

- SEARN, DAGGER (Daume et al., 2009, Ross et al., 2011)
- Scheduled Sampling (Bengio et al, 2016)
- Curriculum Learning (Ranzato et al., 2016)
- Reward Sampling (Pasunuru et al., 2017, Liu et al., 2017)
- Self-critical Learning (Rennie et al., 2017, Paulus et al., 2018, Celikyilmaz et al., 2018)

Reward Matters

	CIDEr	BLEU-4	ROUGE-L	METEOR
XENT	94.0	29.6	52.6	25.2
CIDEr	106.3	31.9	54.3	25.5
BLEU-4	94.4	33.2	53.9	24.6
ROUGE-L	97.7	31.6	55.4	24.5
METEOR	80.4	25.3	51.3	25.9

Reward Matters

	CIDEr	BLEU-4	ROUGE-L	METEOR
XENT	94.0	29.6	52.6	25.2
CIDEr	106.3	31.9	54.3	25.5
BLEU-4	94.4	33.2	53.9	24.6
ROUGE-L	97.7	31.6	55.4	24.5
METEOR	80.4	25.3	51.3	25.9

Reward Matters

- No correlated metric?
 - Paraphrase Generation - entailment model for rewards (Li et al., 2017)
 - Dialogue Response Generation - mixture of similarity and perplexity rewards (Li et al., 2016)
 - Sentence Simplification - mixture of similarity and perplexity rewards (Zhang and Lapata, 2017)
 - Image Captioning - image and caption representation similarity (Ren et al., 2017)
 - Actor-critic Methods (Bahdanau et al., 2016)

Temporal Patterns



- 1) Wash the lettuce and place in a bowl
- 2) Dry the lettuce with a paper towel or a salad spinner
- 3) Mix oil, mustard and vinegar in a small bowl with a spoon.
- 4) Toss the salad with the vinaigrette before serving.

Temporal Patterns

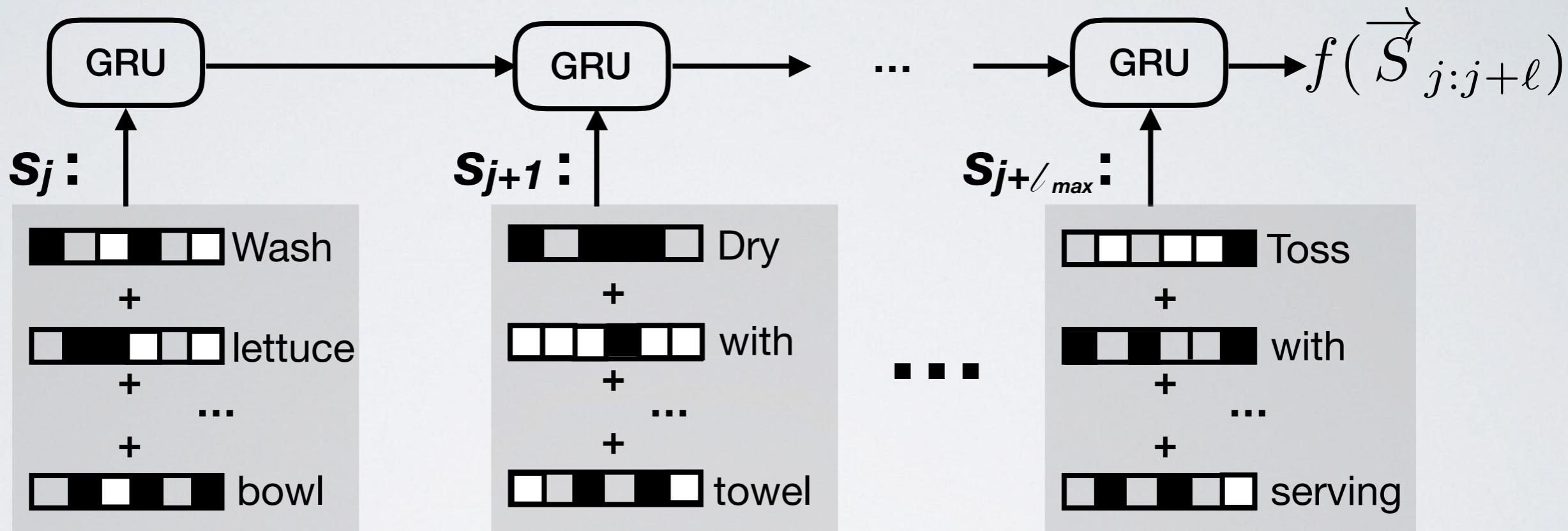


- 1) Wash the lettuce and place in a bowl
- 2) Dry the lettuce with a paper towel or a salad spinner
- 3) Toss the salad with the vinaigrette before serving.**
- 4) Mix oil, mustard and vinegar in a small bowl with a spoon.**

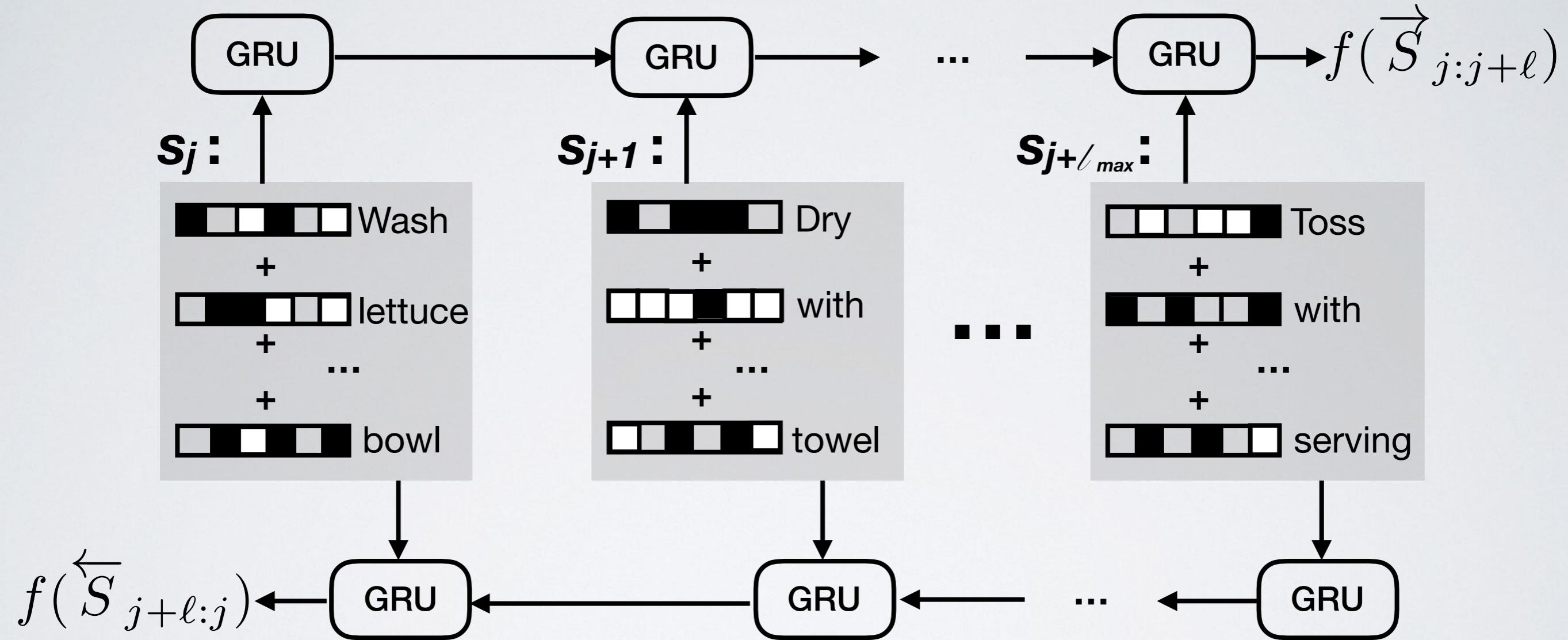
Outline

- Recipe Generation
- Neural Teachers
- Self-critical Sequence Training
- Evaluations

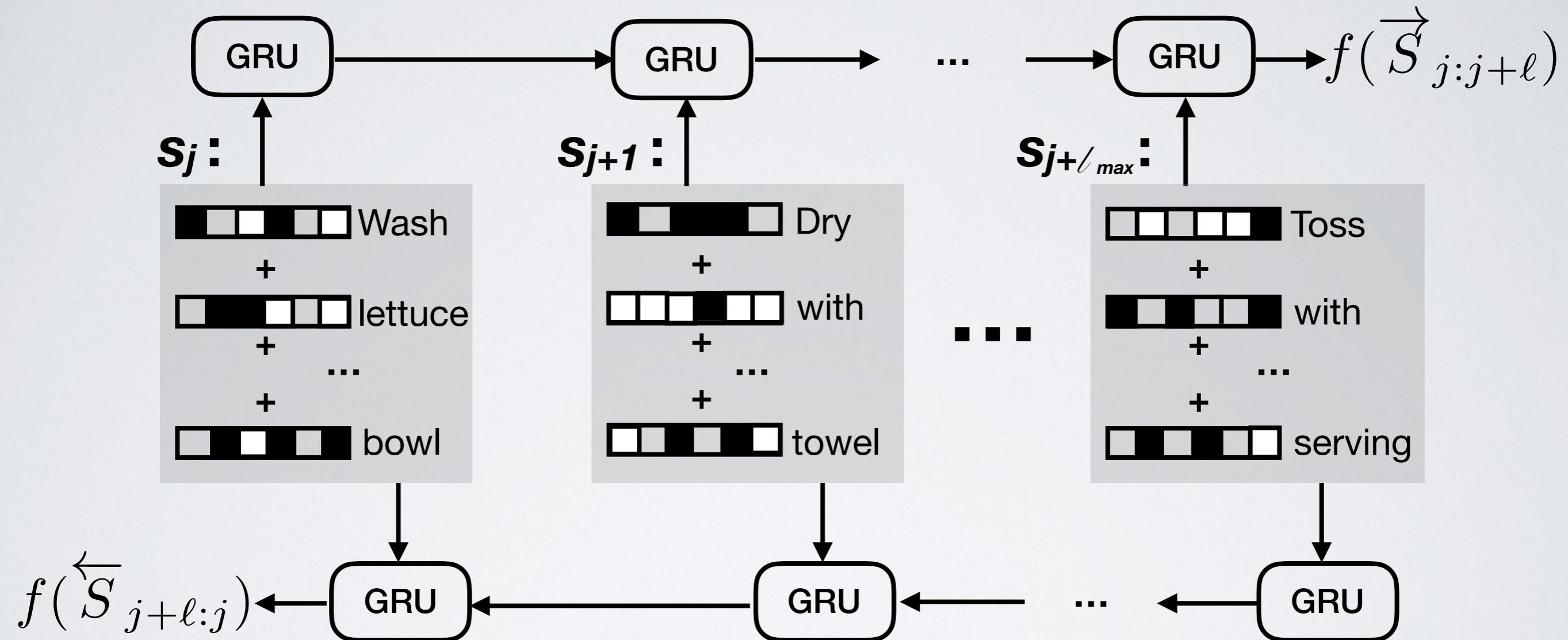
Neural Teacher



Neural Teacher



Neural Teacher



$$\mathcal{L}_{rel} = \frac{\langle f(\vec{S}_{j:j+\ell}), f(\overleftarrow{S}_{j+\ell:j}) \rangle}{\|f(\vec{S}_{j:j+\ell})\| \|f(\overleftarrow{S}_{j+\ell:j})\|}$$

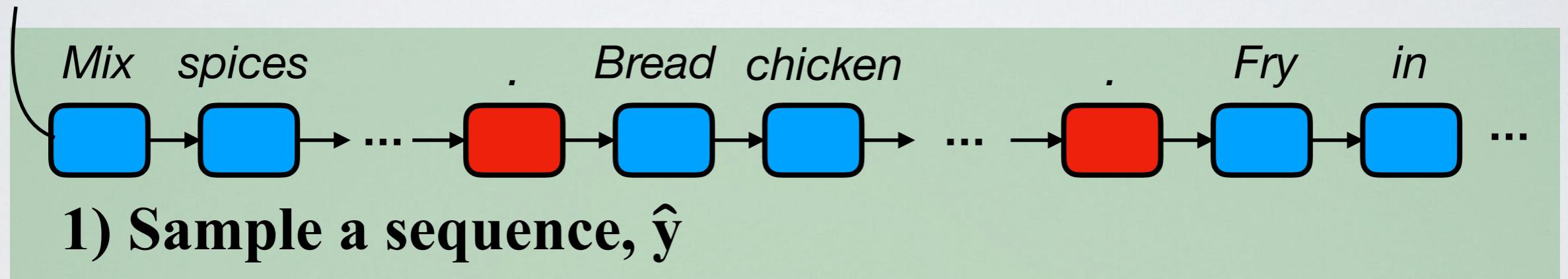
Outline

- Recipe Generation
- Neural Teachers
- **Self-critical Sequence Training**
- Evaluations

Self-critical Training

Fried Chicken

- **Chicken**
- **Flour**
- **Spices**

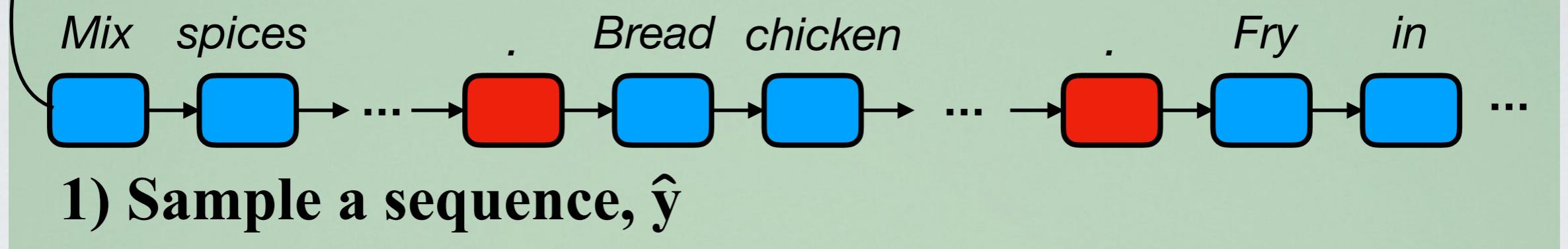


Self-critical Training

Fried Chicken

- Chicken
- Flour
- Spices

Teacher



$$r(s_j) = \sum_{\ell=\ell_{min}}^{\ell_{max}} \left(\frac{\langle f(S_{j-\ell:j}), f(\vec{S}_{j-\ell:j}) \rangle}{\|f(S_{j-\ell:j})\| \|f(\vec{S}_{j-\ell:j})\|} - \frac{\langle f(S_{j-\ell:j}), f(\overleftarrow{S}_{j-\ell:j}) \rangle}{\|f(S_{j-\ell:j})\| \|f(\overleftarrow{S}_{j-\ell:j})\|} \right)$$

Self-critical Training

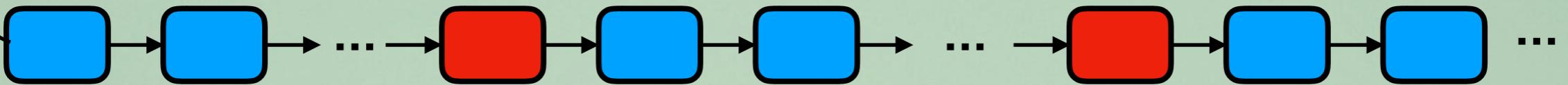
Fried Chicken

- Chicken
- Flour
- Spices

Teacher

$r(\hat{s}_1), r(\hat{s}_2), \dots, r(\hat{s}_n)$

Mix spices



1) Sample a sequence, \hat{y}

Self-critical Training

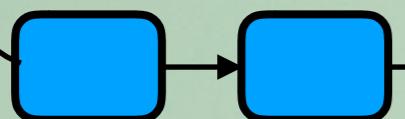
Fried Chicken

- Chicken
- Flour
- Spices

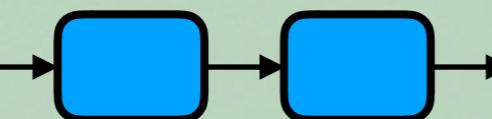
Teacher

$r(\hat{s}_1), r(\hat{s}_2), \dots, r(\hat{s}_n)$

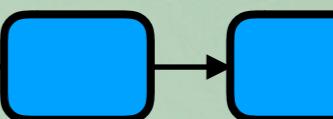
Mix spices



Bread chicken

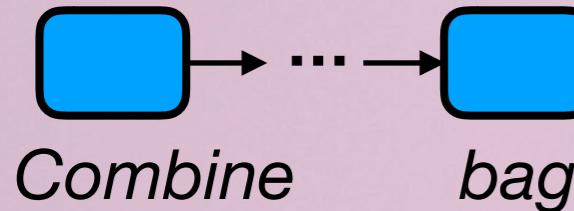


Fry in



1) Sample a sequence, \hat{y}

2) Greedily decode a sequence, y^*



Teacher

Self-critical Training

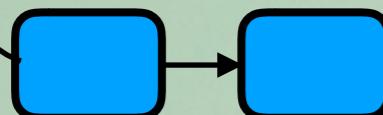
Fried Chicken

- Chicken
- Flour
- Spices

Teacher

$r(\hat{s}_1), r(\hat{s}_2), \dots, r(\hat{s}_n)$

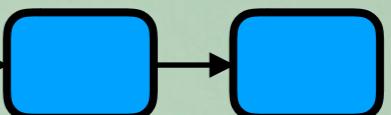
Mix spices



Bread chicken

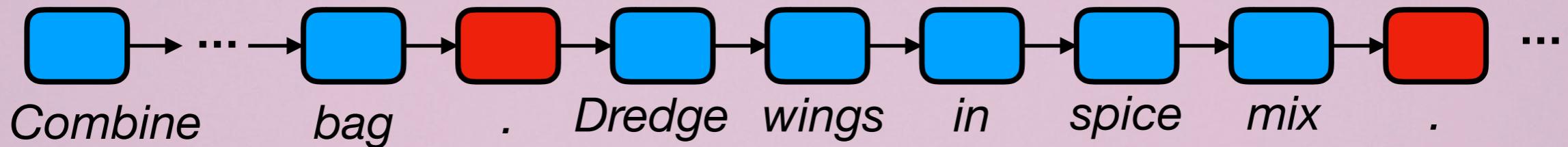


Fry in



1) Sample a sequence, \hat{y}

2) Greedily decode a sequence, y^*



Teacher

$r(s_1^*), r(s_2^*), \dots, r(s_n^*)$

3) Compute rewards

Sentence-level Rewards

$$r(s_j) = \sum_{\ell=\ell_{min}}^{\ell_{max}} \left(\frac{\langle f(S_{j-\ell:j}), f(\vec{S}_{j-\ell:j}) \rangle}{\|f(S_{j-\ell:j})\| \|f(\vec{S}_{j-\ell:j})\|} - \frac{\langle f(S_{j-\ell:j}), f(\overleftarrow{S}_{j-\ell:j}) \rangle}{\|f(S_{j-\ell:j})\| \|f(\overleftarrow{S}_{j-\ell:j})\|} \right)$$

$$r_t = \sum_{j=1}^{|S|} \mathbb{1}(y_t \in \hat{s}_j) (r(\hat{s}_j) - r(s_j^*))$$

sampled

decoded

Sentence-level Rewards

$$r(s_j) = \sum_{\ell=\ell_{min}}^{\ell_{max}} \left(\frac{\langle f(S_{j-\ell:j}), f(\vec{S}_{j-\ell:j}) \rangle}{\|f(S_{j-\ell:j})\| \|f(\vec{S}_{j-\ell:j})\|} - \frac{\langle f(S_{j-\ell:j}), f(\overleftarrow{S}_{j-\ell:j}) \rangle}{\|f(S_{j-\ell:j})\| \|f(\overleftarrow{S}_{j-\ell:j})\|} \right)$$

$$r_t = \sum_{j=1}^{|S|} \mathbb{1}(y_t \in \hat{s}_j) (r(\hat{s}_j) - r(s_j^*))$$

sampled

decoded

$$\mathcal{L}_{rl} = - \sum_{t=1}^T r_t \log P(\hat{y}_t | \hat{y}_0, \dots, y_{t-1}, \mathbf{g}, \mathbf{e})$$

Mixed Training

Title: Jellied Horseradish

Ings: horseradish, sugar, vinegar, fruit pectin

Generated Recipe:

Add sugar and sugar. **Add sugar and cook. Add sugar and cook. Add sugar and cook.** Remove from heat and **add sugar**. Fold in whipped cream. Chill. **Add sugar** and lemon juice. Add sugar.

Mixed Training

$$\mathcal{L}_{mix} = \gamma \mathcal{L}_{rl} + (1 - \gamma) \mathcal{L}_{mle}$$

The diagram illustrates the formula for mixed training loss. At the top, the equation $\mathcal{L}_{mix} = \gamma \mathcal{L}_{rl} + (1 - \gamma) \mathcal{L}_{mle}$ is shown. Below it, two upward-pointing arrows originate from the terms \mathcal{L}_{rl} and \mathcal{L}_{mle} . The arrow from \mathcal{L}_{rl} is labeled "RL Loss" and the arrow from \mathcal{L}_{mle} is labeled "MLE Loss".

Datasets

- Now You're Cooking Dataset
 - Kiddon et al., 2016
- 140000 training recipes
- 1000 development recipes
- 1000 test recipes

Experimental Setup

- Pretrain recipe generator
- Use mixing weight of $\gamma = 0.97$ during RL
- Reward is average of teacher scores for encoding 3 to 6 sentences
- Train for 600000 episodes use checkpoint that achieves highest reward

Outline

- Recipe Generation
- Neural Rewards
- Self-critical Sequence Training
- **Evaluations**

Metrics

- Word-level
 - BLEU-1, BLEU-4, ROUGE-L
- Action-level
 - BLEU-1, BLEU-4
- State-change level
 - BLEU-1, BLEU-4

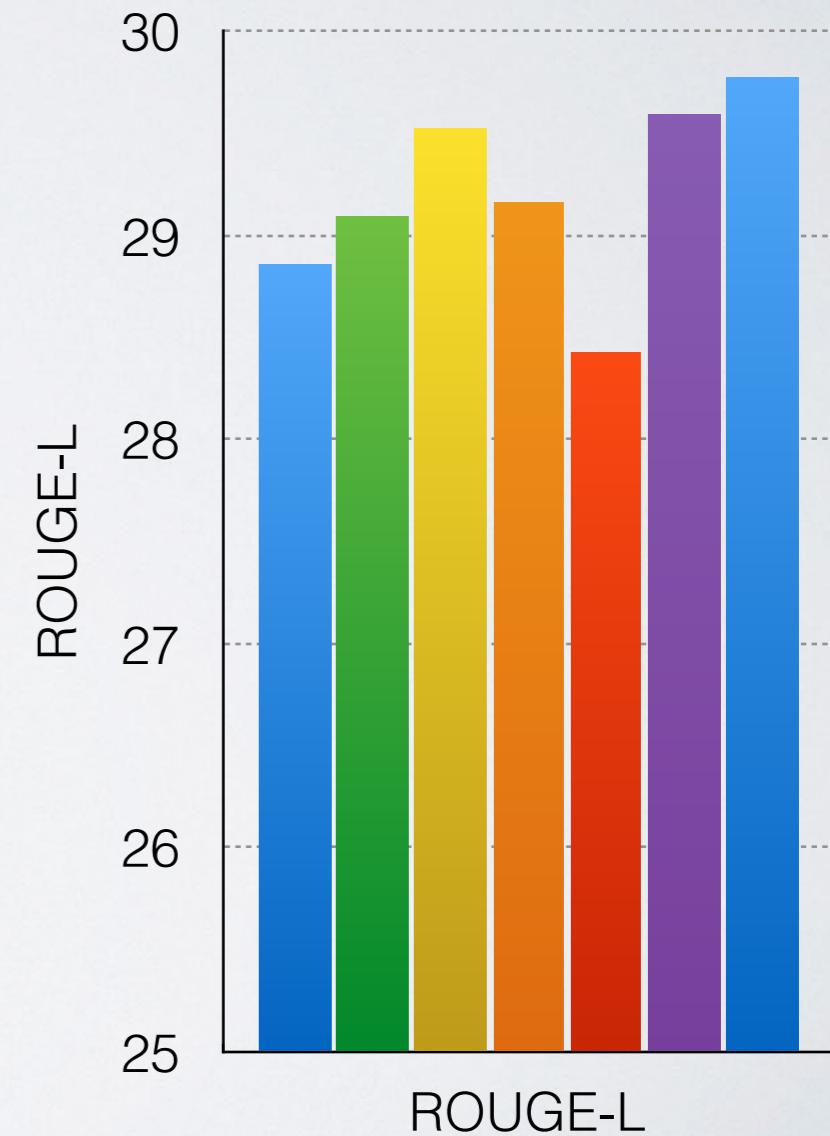
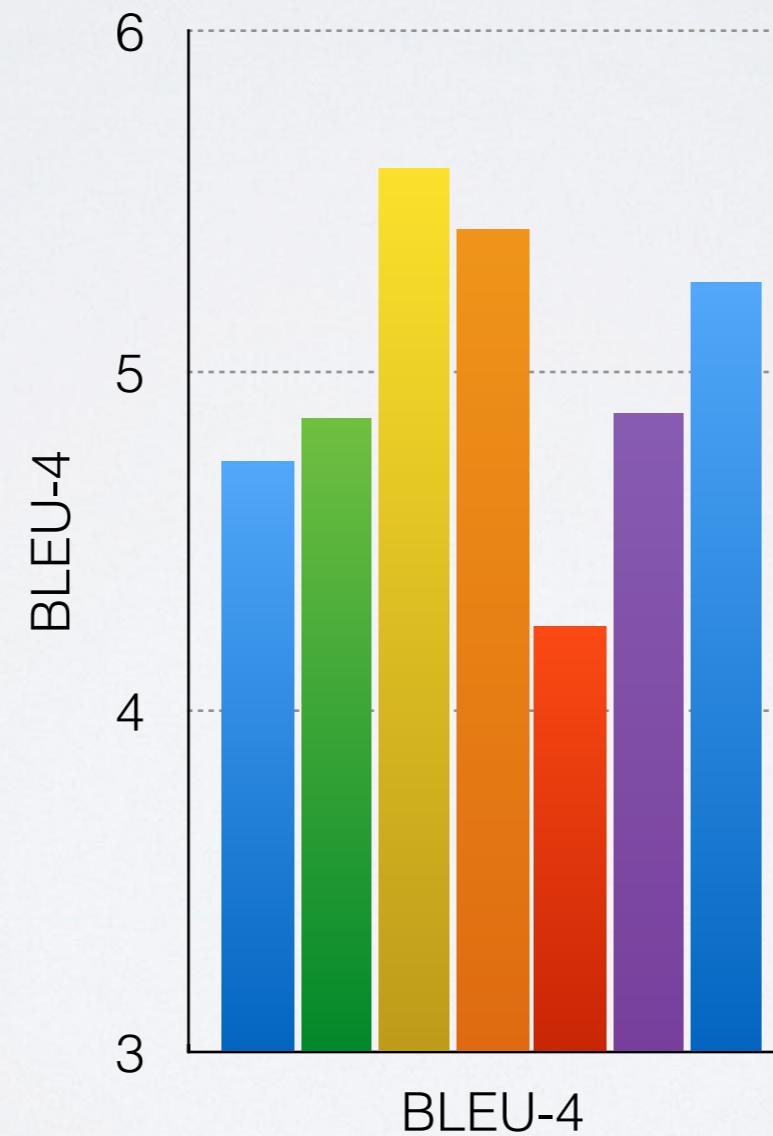
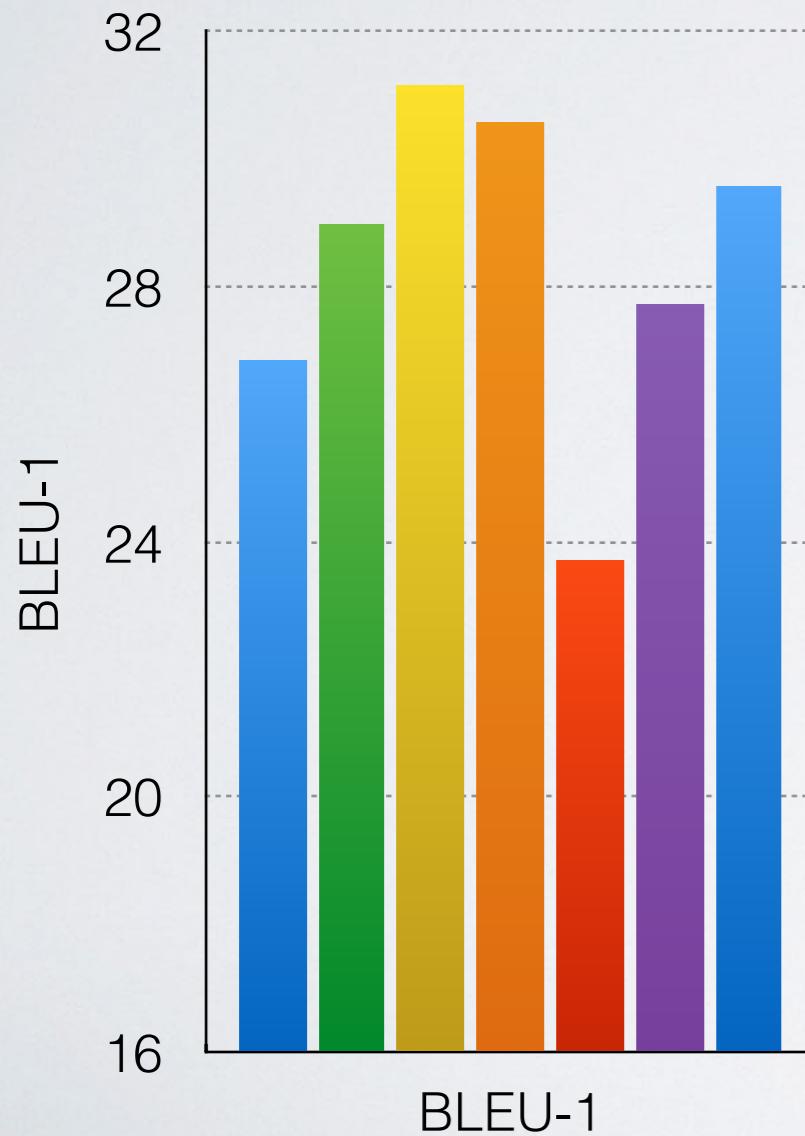
Baselines

- MLE-trained recipe generator
- Socher et al., 2018 model
- Rennie et al., 2017 model
- BLEU-1 RL trained model
- BLEU-4 RL trained model

Word Results

Word Scores

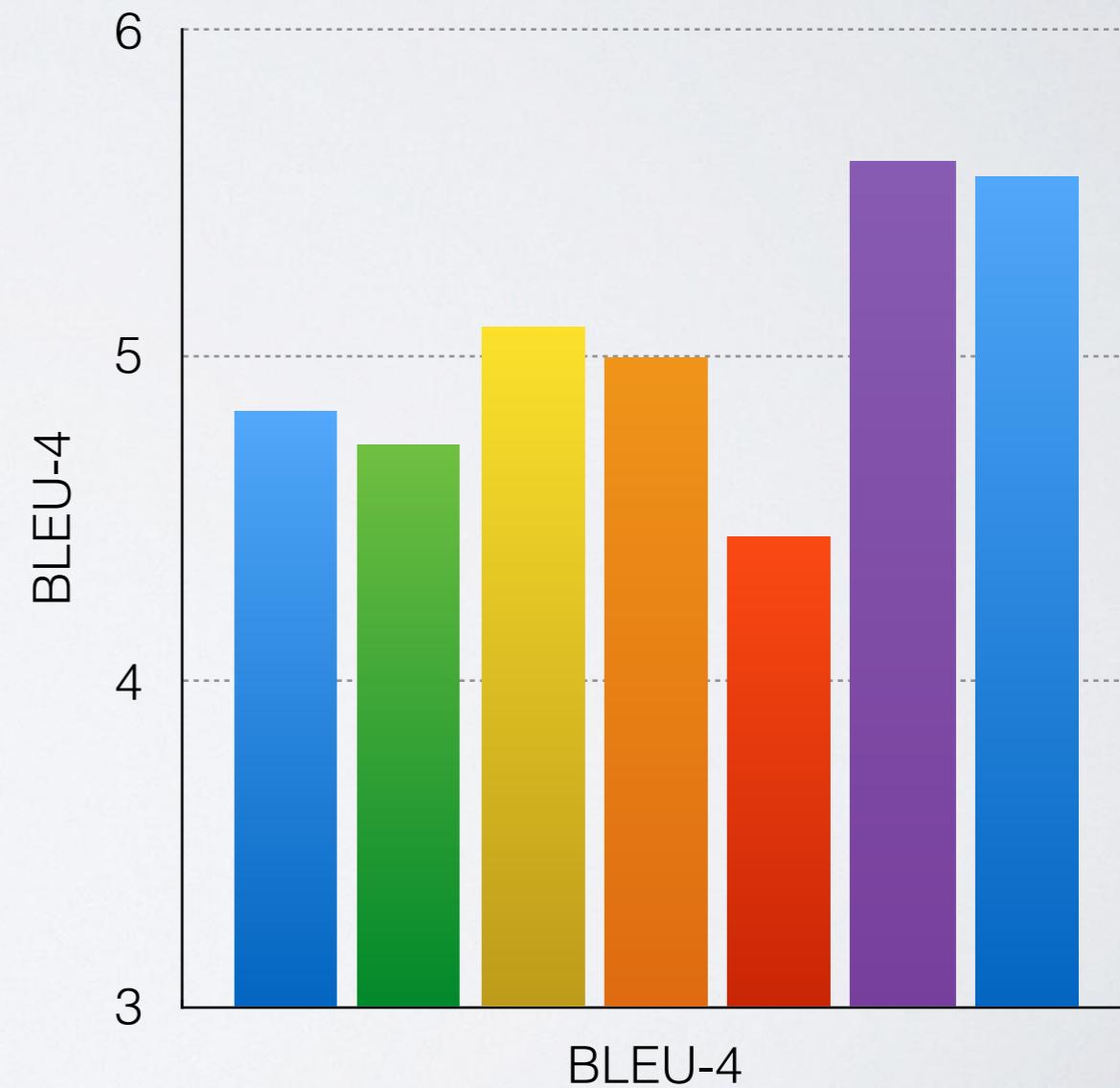
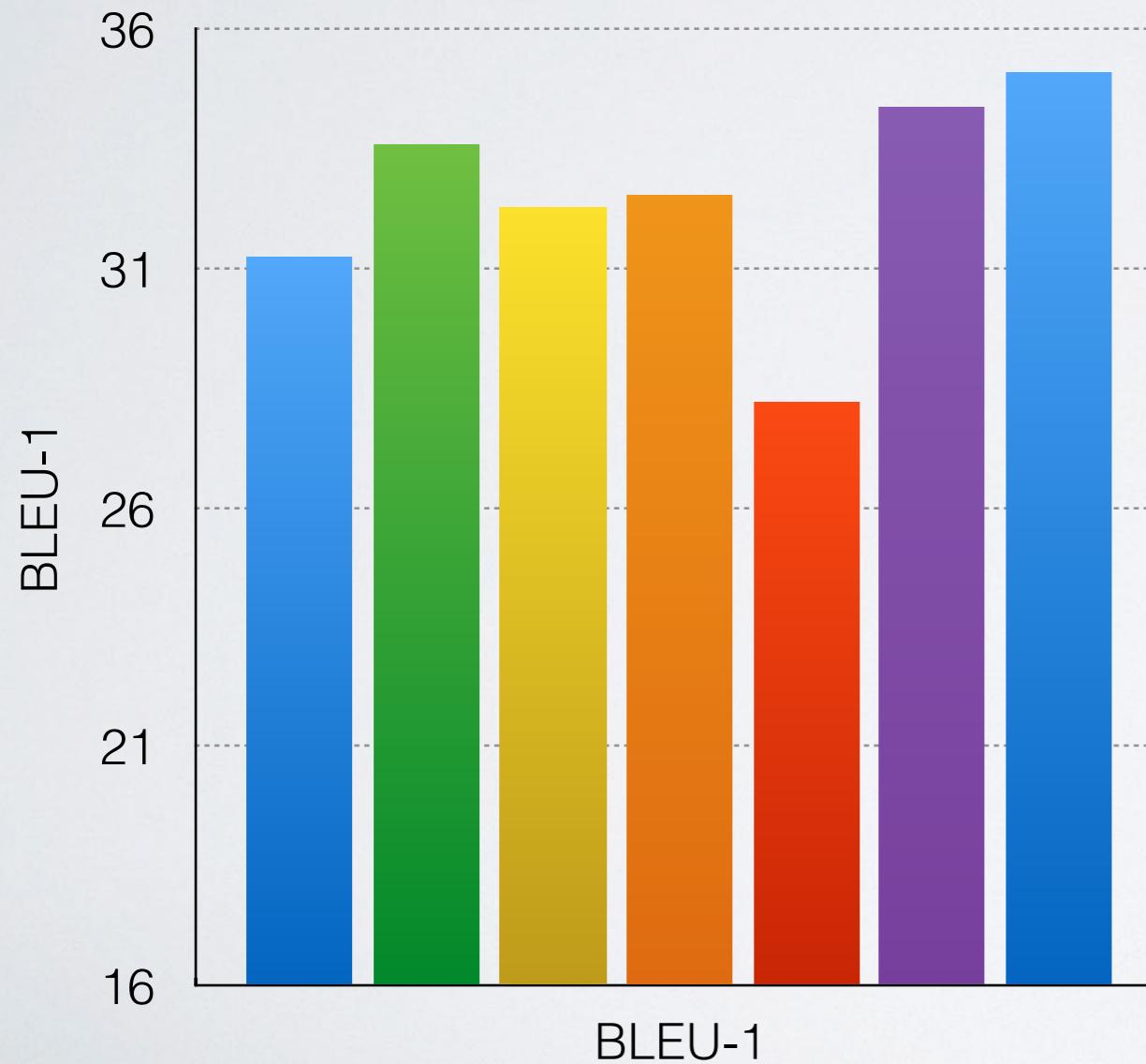
MLE SC_ROUGE B1 B4 AO RO RO+B4



Action Results

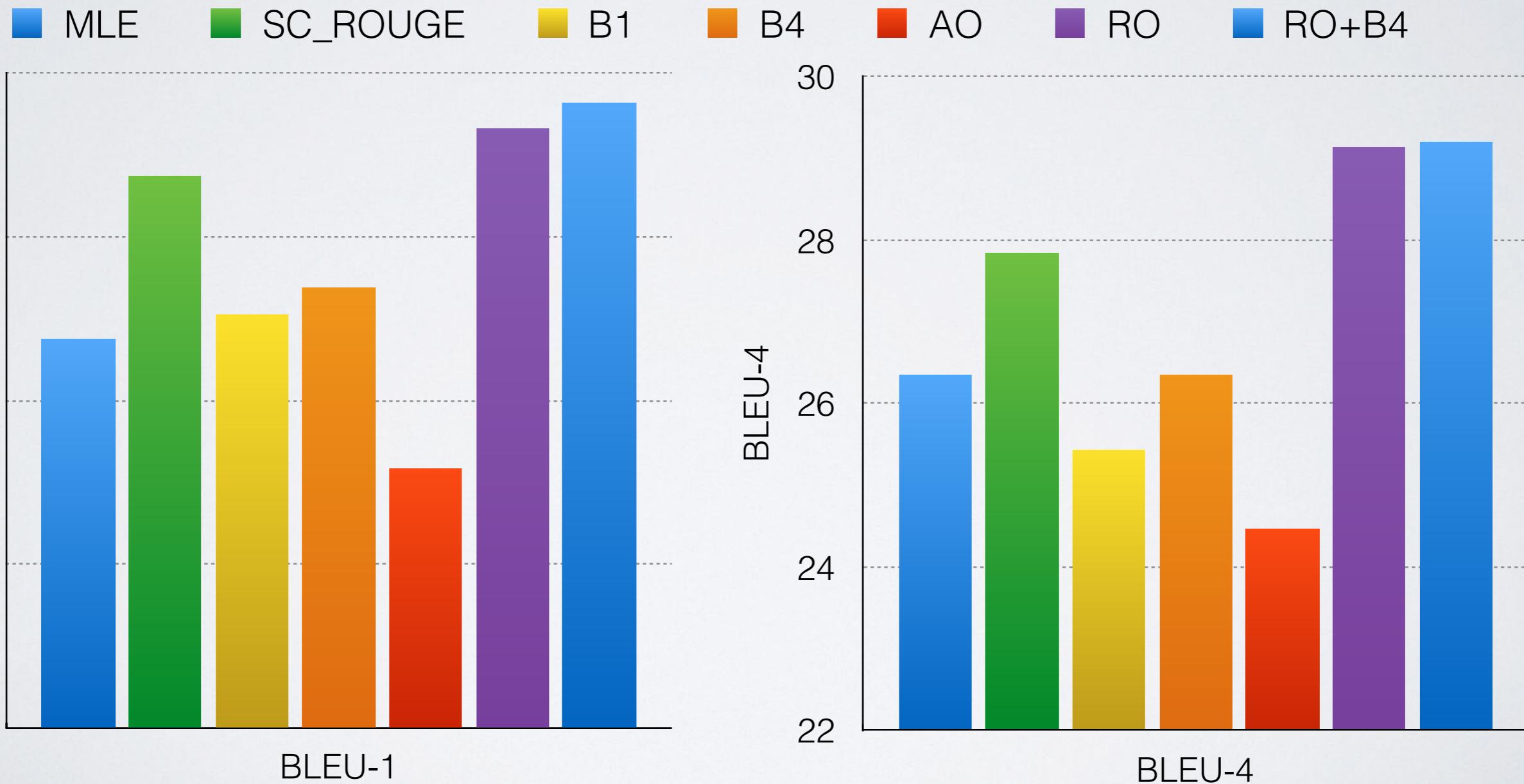
Action Scores

MLE SC_ROUGE B1 B4 AO RO RO+B4

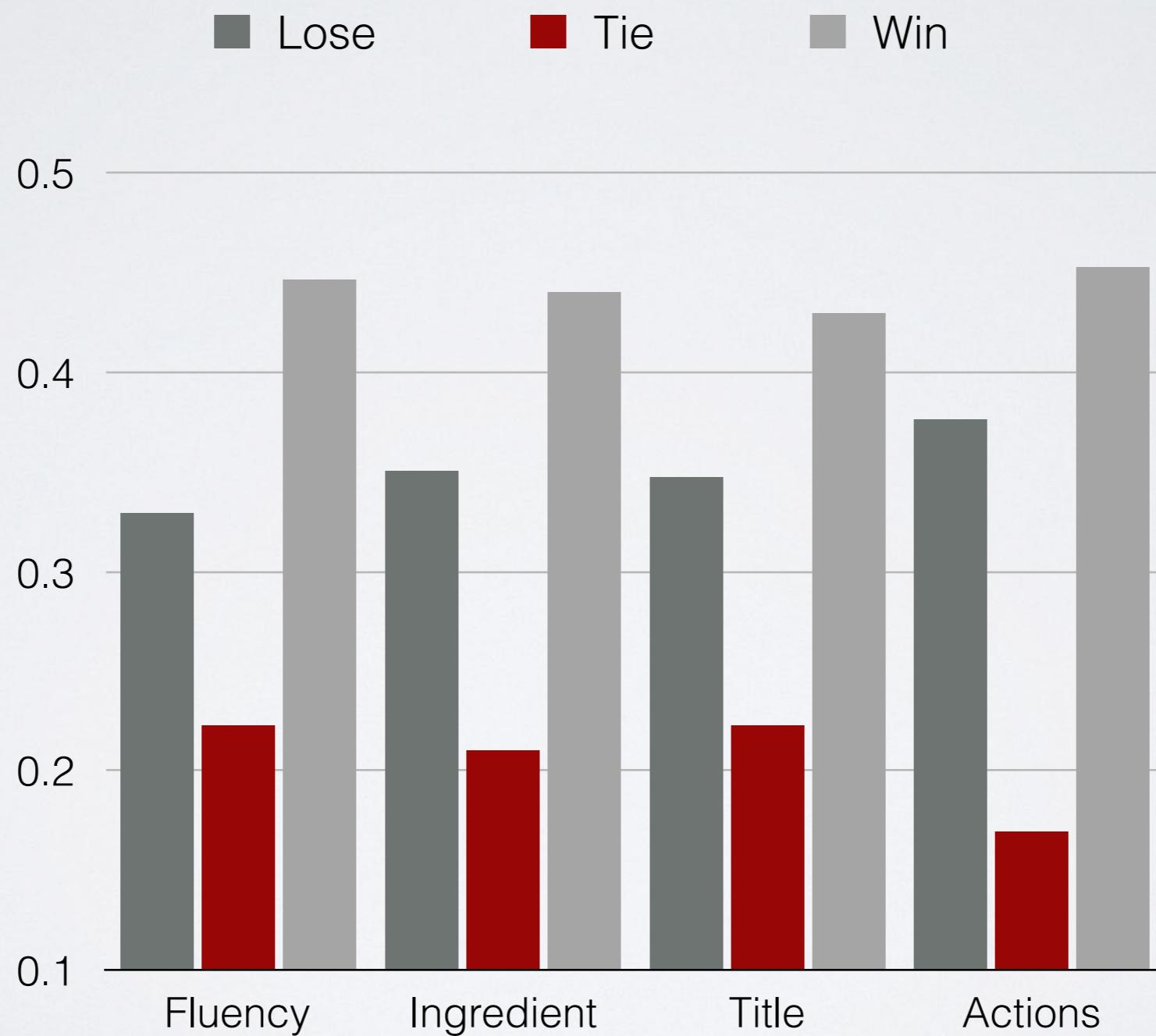


State Change Results

State Change Scores



Human Evaluations



Examples



Maple cream candy

brown sugar
cream
butter

Gold Recipe	MLE	Mixed MLE + RL (Ours)
<ol style="list-style-type: none"> 1. Stir all together. 2. Boil until it drops from spoon. 3. Add vanilla. 4. Beat until it thickens. 5. Pour into buttered pan. 	<ol style="list-style-type: none"> 1. Mix all ingredients. 2. Cook on low for 10 minutes. 3. Remove from fire and add flavoring. 4. Beat until creamy. 5. Beat until creamy. 6. Beat until creamy. 7. Beat until creamy. 8. Beat until creamy. 9. Beat until creamy. 10. Beat until creamy. 11. Beat until creamy. 12. Beat until creamy. 13. Beat until creamy. 14. Add vanilla. 15. Beat until creamy. 16. Add vanilla. 17. Beat until creamy. 18. Add vanilla. 19. Shape into balls. 20. Wrap in waxed paper and chill. 	<ol style="list-style-type: none"> 1. Combine sugar, cream and butter. 2. Boil until firm. 3. Remove from fire and add flavoring. 4. Cool. 5. Beat until creamy. 6. Add vanilla and beat until stiff. 7. Form into balls and roll in sugar. 8. Wrap in wax paper and chill.

Examples



Honey baked apples

apples baking

chopped walnuts

raisins

water

honey

3-inch stick cinnamon

lemon juice

Gold Recipe	MLE	Mixed MLE + RL (Ours)
<ol style="list-style-type: none">1. Core apples;2. Peel top third of each.3. Place apples in a shallow baking dish.4. Combine walnuts and raisins;5. Stuff cavities of apples with nut mixture.6. Combine water, honey, and cinnamon in a small saucepan;7. Bring to a boil.8. Reduce heat and simmer 5 minutes.9. Remove from heat and stir in lemon juice.10. Remove cinnamon stick;11. Pour liquid over apples.12. Cover and bake at 350 degrees f for 45 to 50 minutes or until apples are tender, basting occasionally.13. 6 servings.	<ol style="list-style-type: none">1. Core apples and pare.2. Cut apples in half.3. Place in a baking dish.4. Combine remaining ingredients and pour over apples.5. Bake at 350 degrees for 1 hour.	<ol style="list-style-type: none">1. Core apples and pare.2. Place apples in a baking dish.3. Combine remaining ingredients and mix well.4. Spoon into cavities of apples.5. Bake in 350 f oven for 1 hour.6. Serve hot.

Examples



Carrot Chowder

browned and ground beef rained
 celery chopped
 chopped onion
 chopped green pepper
 carrots grated
 32oz juice tomato
 celery cream of soup
 water
 garlic salt
 marjoram
 sugar
 salt
 jack <unk>
 cheese

Gold Recipe	MLE	Mixed MLE + RL (Ours)
<ol style="list-style-type: none"> 1. Bring all but cheese to a boil; 2. Simmer until veggies are tender. 3. Sprinkle each serving with cheese. 	<ol style="list-style-type: none"> 1. In large saucepan, brown hamburger, onion, and celery in margarine. 2. Add remaining ingredients, except cheese. 3. Heat to boiling. 4. Reduce heat and simmer, covered, for 30 minutes. 5. Stir in cheese. 6. Heat to boiling. 7. Reduce heat and simmer, uncovered, for 5 minutes. 8. Serve hot. 	<ol style="list-style-type: none"> 1. Combine all ingredients except cheese, celery, and onion in a large dutch oven; 2. Bring to a boil. 3. Reduce heat; 4. Cover and simmer for 30 minutes. 5. Stir in macaroni, cheese, and cheese. 6. Heat through. 7. Serve over hot noodles.

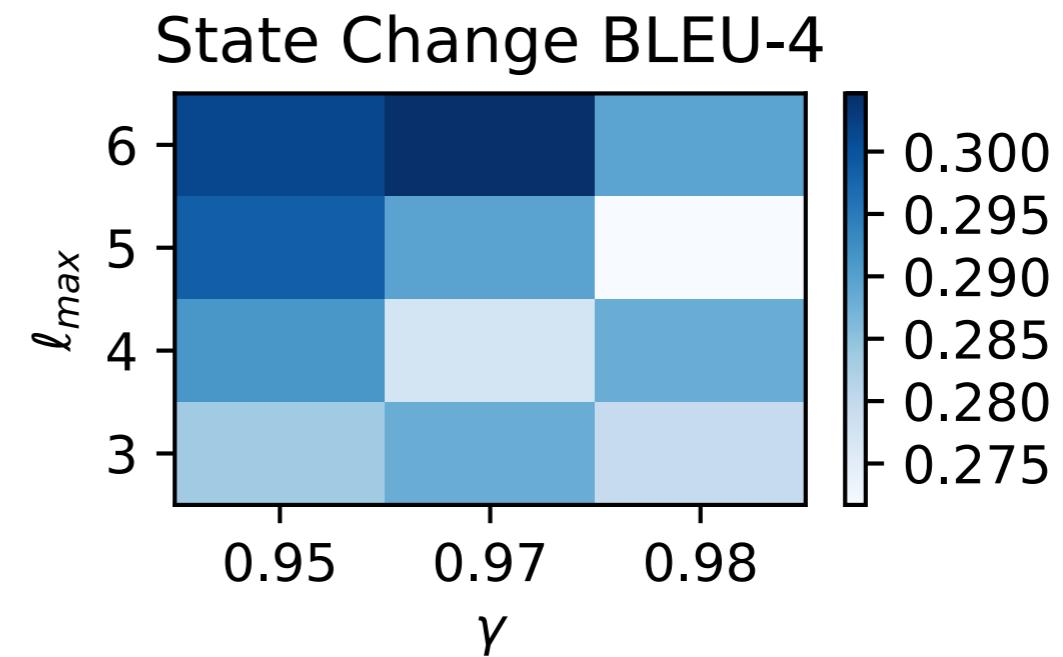
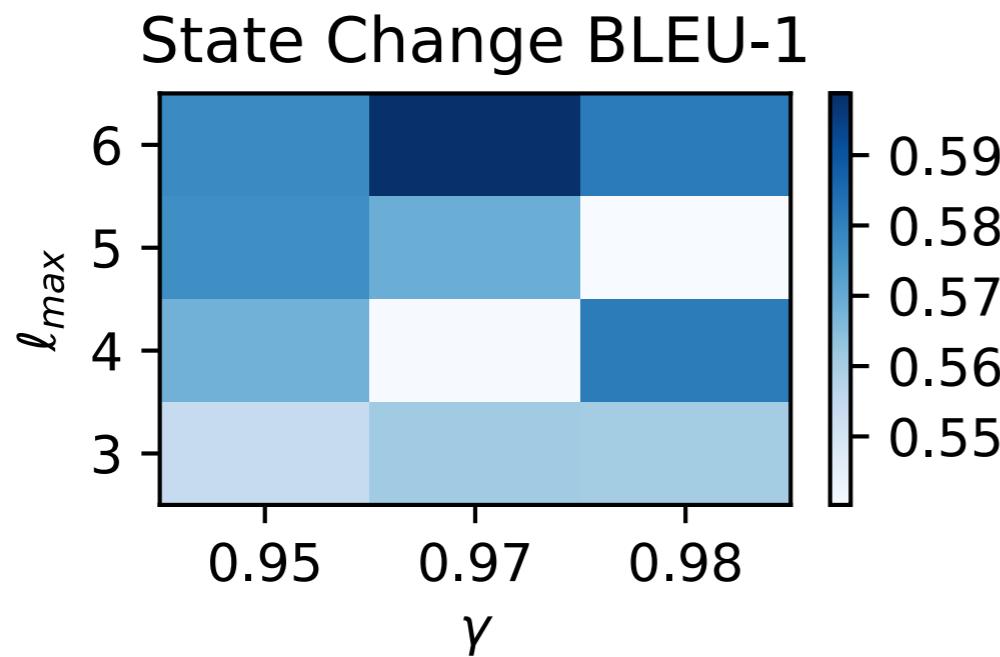
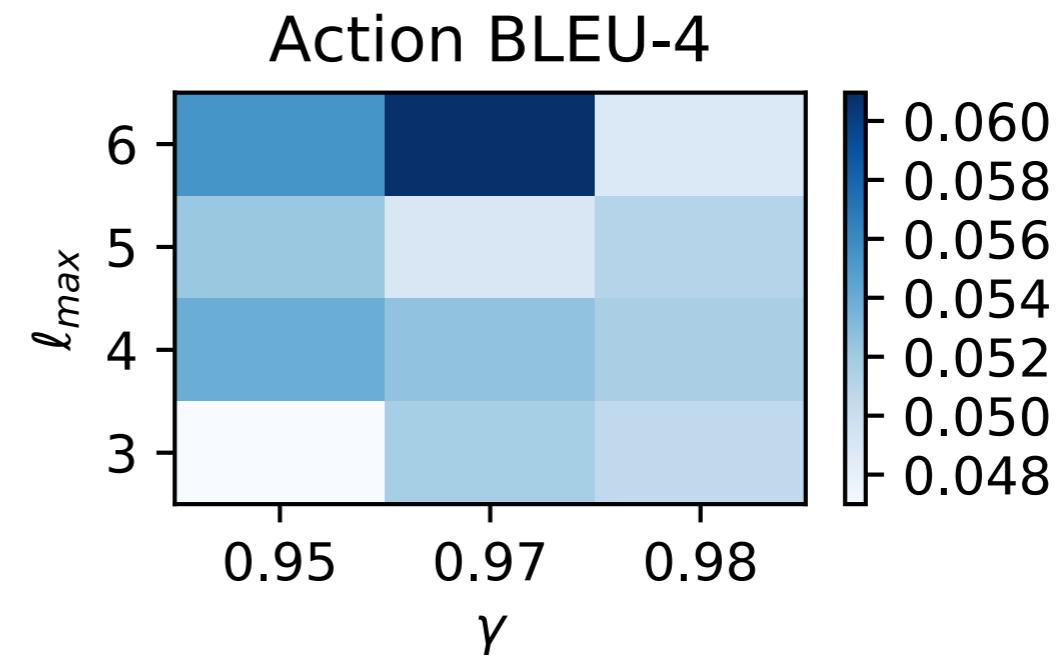
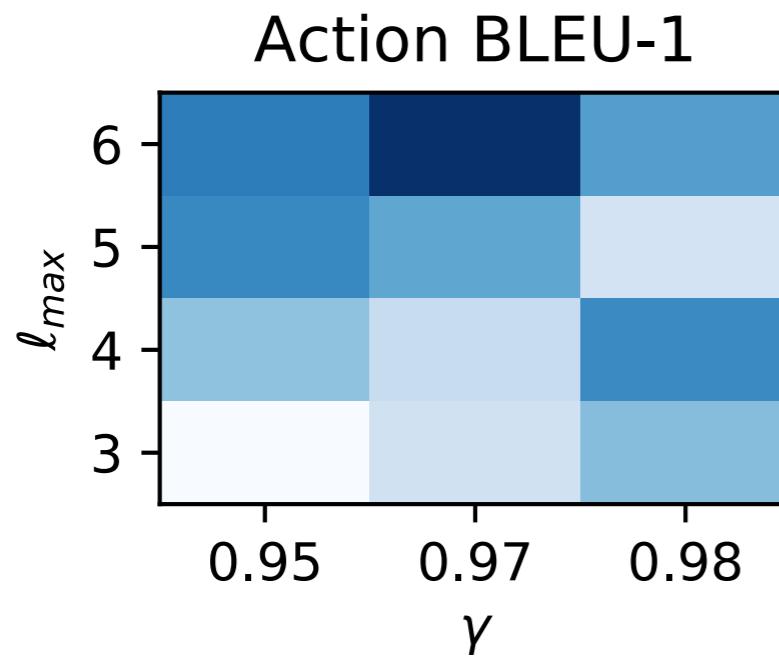
Conclusion

- Neural teacher function that can learn discourse patterns in long form texts
- A method for applying the scores it produces as rewards to neural text generators



<https://homes.cs.washington.edu/~antoineb/>

Hyperparameters



Reverse Reward

