Towards Text Generation with Adversarially Learned “Neural Outlines” NeurIPS 2018

Sandeep Subramanian, Sai Rajeshwar, Alessandro Sordoni, Adam Trischler, Aaron Courville, Christopher Pal

Mila Université de Montréal McGill

Microsoft

Generative models for text

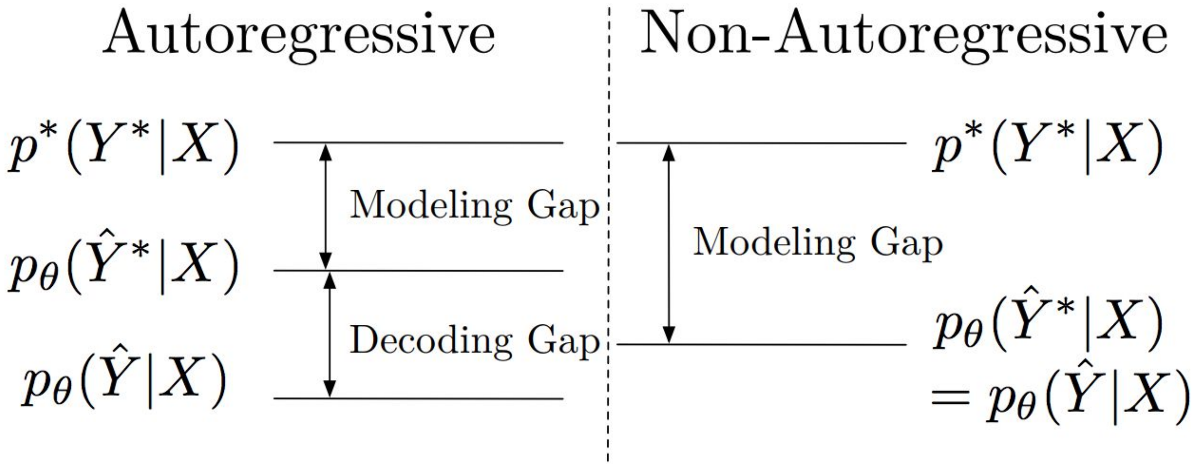
- Autoregressive models - n-gram, recurrent, convolutional, transformer etc.

$$p(x_1, x_2 \ldots, x_n) = p(x_1) \prod_{i=2}^{n} p(x_i | x_{<i})$$

- Conditional probabilities are estimated via counts (n-gram LMs) or parameterized by neural nets and learned.
- They are trained by “teacher forcing” wherein the model sees the ground truth observation at the next time step, independent of what the model may have predicted.
- At inference, the model needs to condition on its own outputs produced greedily (picking the most likely word/char) or by sampling from the softmax.

Potential Issues

- Exposure bias - Ranzato et al. (2015), model hasn't learned to condition on it's own outputs. Remedies (non-exhaustive) include schedule sampling (Bengio et al. 2015), Professor forcing (Lamb et al. 2016) etc.
- Model encodes uncertainty over possible generations only in it's output distribution.



Maybe VAEs?

- Bowman et al. (2015) explore VAEs in the context of language.
- The decoder in the VAE is conditioned on the reparameterized or sampled z. Some of the uncertainty in the output is therefore “pre-determined” by z.
- However, they have been notoriously hard to train due to the “posterior collapse problem”.
- Although several solutions have been proposed, VAEs are still limited by the simplicity of the parameterization of the prior and posterior distributions.

Advances in sentence representation learning

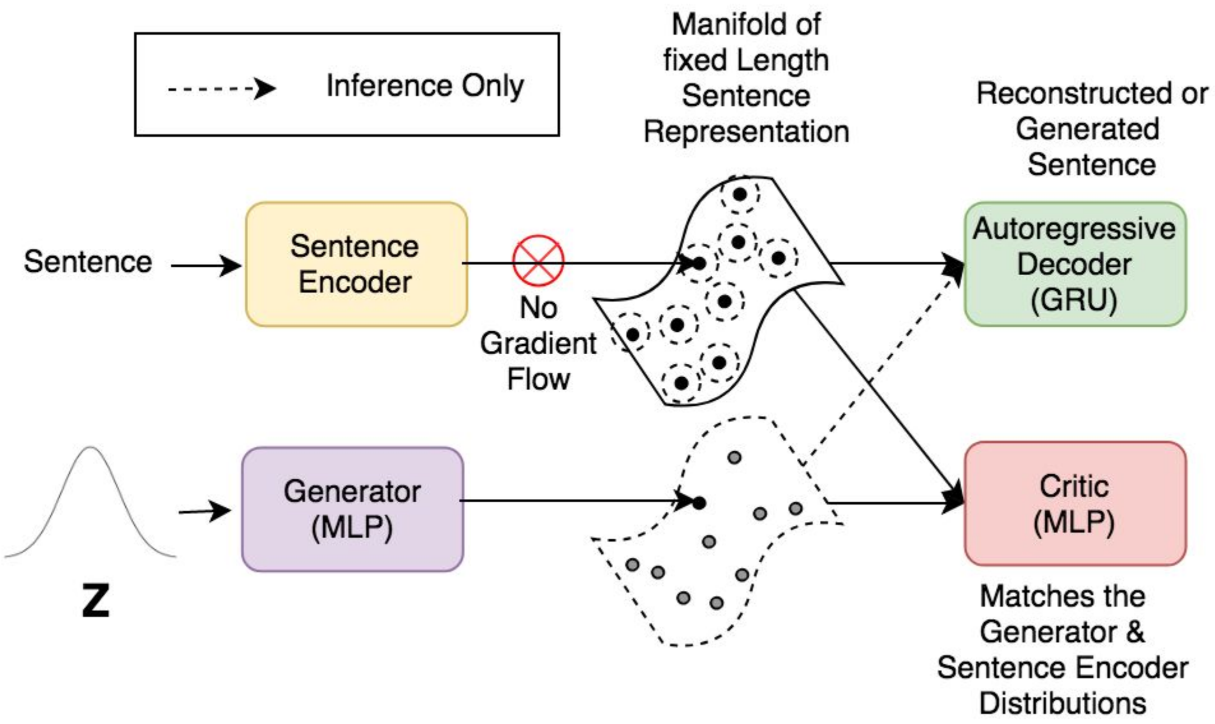
- The latent variable in VAEs doesn't provide enough information to the decoder & the autoregressive signal dominates information from the encoder.
- There have been have been several advancements in representation learning of entire sentences (Kiros et al. 2015, Conneau et al. 2017, Subramanian et al. 2018 etc.).
- What if we interpret general purpose sentence representations as samples from an expressive but unknown prior distribution over a continuous space (from which we cannot sample)?
- By modeling this distribution, we can set up a generative process by
    a. Sampling a sentence vector or a "neural outline"
    b. Decoding it to the observed space (words/characters) using a conditional LM.

This work

Questions we're trying to answer

- Can we model the distribution of learned sentence representations?
- Does that help us learn better generative models of text?
- Can we use this to learn generative models of text directly in the latent space?
- Since the representations are continuous and fixed length, can we use generative adversarial techniques?
- Can we produce meaningful interpolations between sentences in a latent space that doesn't have a simple prior distribution?

Model Architecture



Modeling the distribution of sentence embeddings

- Given a corpus of text, we can encode every **sentence** to $\mathbb{R}^{2048}$ using the sentence encoder.
- Learn a generator (MLP) that maps a random vector (drawn from a simple prior distribution) to one that looks like a sentence embedding.
- The generator is trained adversarially using the Wasserstein GAN (WGAN) objective with gradient penalty (GP) using an MLP as the critic.

Decoding sentence embeddings to words

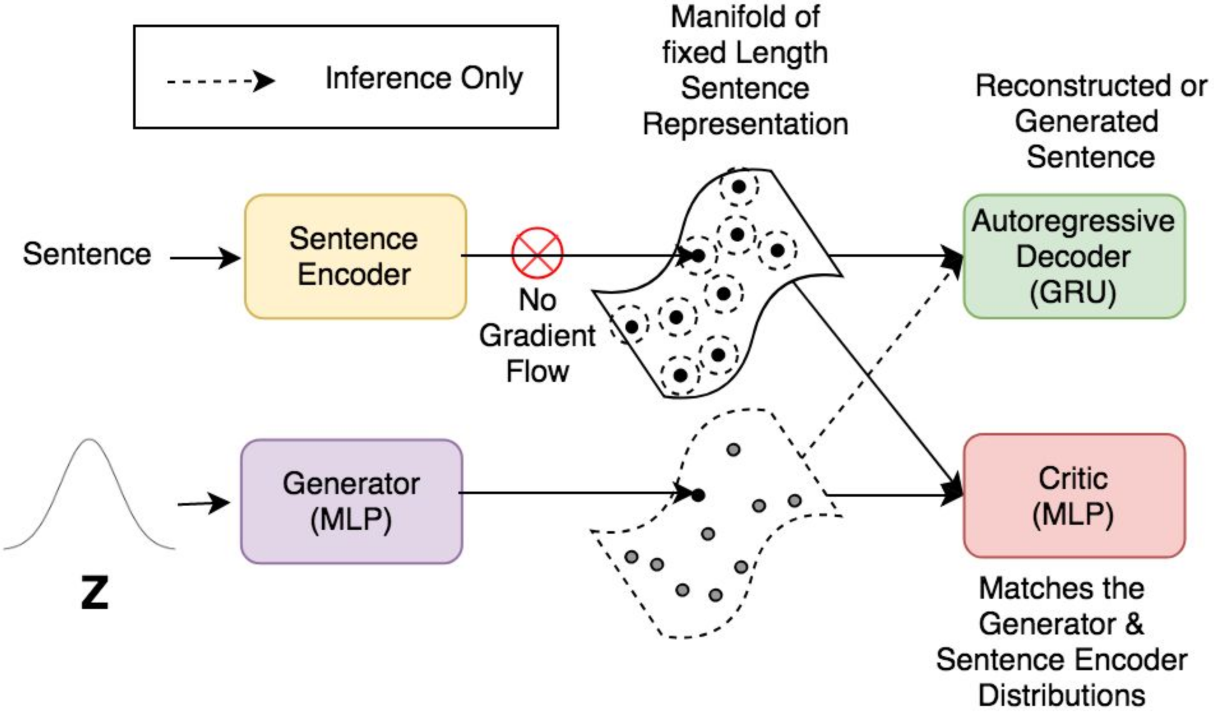
Idea

- Given a sentence embedding, we'd like to reconstruct it's contents to complete the generative process.
- We train an autoregressive (standard) conditional language model that is trained to reconstruct the words in the sentence embedding.
- Can be seen as a sequence autoencoder with a fixed encoder.
- Language model training is **independent** of GAN training.

Importance of adding noise

- We inject gaussian noise into the sentence embedding before reconstruction.
- This locally smooths the manifold so the GAN has a "larger target to aim at".

Model Architecture



Evaluations

What would we like from our generative model?

- Sample quality
- Sample diversity
- Not memorizing the training data
- Density estimation :-(

Forward and Reverse Perplexities

Idea

- Reverse Perplexity (RPPL) - Train a language model on model samples evaluate the likelihood on real data. **(Sample diversity + quality)**
- Forward Perplexity (FPPL) - Train a language model on data samples evaluate the likelihood on model samples. **(Sample quality)**

Justification

- In the limit that the approximation language model models the data distributions, we can recover the perplexities in both directions.
- Scales well with bigger data and more samples.

Simple n-gram statistics & Overfitting

N-gram statistics

- The number of unique n-grams can tell us something about sample diversity.
- The number of n-grams in the samples that can occur in a holdout set can help assess sample quality.

Overfitting

- The corpus-level BLEU score between a sample and it's nearest neighbor in the training set can be used to probe overfitting
- Nearest neighbors computed using word embedding average similarity.

Results (Unconditional) - Qualitative

1	the room was nicely decorated and the two of them were very comfortable and the bathroom was fantastic .
2	all of the information was gathered from the police or the court of justice in the united states .
3	we are working with the elders to tell the story of the ancient egyptian stories of the past .
4	all of our doctors , nurses , and other health care providers have been waiting for me .
5	and this is why it is so important that the health care system be fully understood .
6	this is going to be a good way to get a glimpse of the new york city council .
7	“ what ’s going on with you ? ”
8	i shook my head , not trusting myself .
9	i was too tired to think about it .
10	“ yes , ” he said , nodding .
1	in the mid-1980s , he was appointed as a member of the court of human rights in afghanistan .
2	do you have any other ideas about cooperation between the european union and other countries in the world ?
3	secondly , i am not happy to see that the countries of the european union are in agreement .
4	the main objective of this study is to promote the development of a more comprehensive and accessible information society .
5	we’ve been looking forward to welcoming you to the beach , with a view of the sea .
6	but it is clear that the west and east of the country are not yet fully committed .
7	i would like to point out to the house that there are some amendments to the fisheries act .
8	health and education , research and development are a major factor in the development of health education programs .
9	i therefore ask the commission to cooperate fully with the commission and to parliament to approve this report .
10	i hope that the next step will be to ensure that this agreement is maintained in the eu .

Table 3. Generated samples from our model trained on the BookCorpus (top) and WMT15 (bottom)

Results (Unconditional) - Quantitative

Dataset	ARAE						WD-LSTM				Ours					
	FPPL			RPPL			FPPL		RPPL		FPPL			RPPL		
	0.5	1.0	B=1	0.5	1.0	B=1	0.5	1.0	0.5	1.0	0.5	1.0	B=5	0.5	1.0	B=5
BookCorpus	389.6	555.6	364.2	209.2	206.2	213.3	9.4	185.2	280.7	137.2	25.5	66.6	10.5	220.4	152.8	250.9
WMT15	448.7	965.1	385.8	476.2	378.7	626.3	21.4	369.0	528.9	250.5	105.5	212.9	19.9	350.5	254.1	373.2
SNLI	67.5	109.1	62.0	54.8	54.0	59.9	5.9	57.0	86.8	34.5	18.6	35.6	15.3	90.8	49.5	59.8

Approach	Unique n-gram ratio			N-gram validity		
	n=2	n=3	n=4	n=2	n=3	n=4
Ours (Beam)	0.04	0.05	0.07	95.76%	86.44%	65.35%
WD-LSTM (Temp 0.5)	0.05	0.07	0.12	95.22%	84.87%	61.68%
Ours (Temp 1.0)	0.55	0.56	0.57	63.06%	44.34%	26.14%
WD-LSTM (Temp 1.0)	1.38	1.30	1.29	46.20%	31.75%	17.52%

Dataset	BLEU-4			
	WD-LSTM		Ours	
	T=0.5	T=1.0	B=5	T=1.0
BookCorpus	26.8	6.8	47.7	17.92
WMT15	16.6	4.6	13.7	5.5
SNLI	45.8	14.7	34.8	22.1

Results (Unconditional) - Human Evaluation

WMT (Temperature 1.0)	**Grammaticality**	**Topicality**	**Overall**
Ours (Temperature 1.0)	**46.19%**	**48.73%**	**63.95%**
WD-LSTM (Temperature 1.0)	28.90%	25.88%	36.05%
No preference	24.91%	25.39%	-
WMT (Beam Search vs Temperature 0.5)			
Ours (Beam Search)	20.00%	15.20%	**53.20%**
WD-LSTM (Temperature 0.5)	19.30%	24.80%	46.80%
No preference	**60.7%**	**60.00%**	-
BookCorpus (Temperature 1.0)			
Ours (Temperature 1.0)	**50.75%**	**54.27%**	**70.35%**
WD-LSTM (Temperature 1.0)	19.59%	23.61%	29.65%
No preference	29.66%	22.12%	-
BookCorpus (Beam Search vs Temperature 0.5)			
Ours (Beam Search)	22.68%	35.29%	**57.28%**
WD-LSTM (Temperature 0.5)	25.21%	17.64%	42.72%
No preference	**52.11%**	**47.07%**	-

Results - InfoGAN

	Latent Category 1	Latent Category 2	Latent Category 3
1	she turned her head to look at him .	he did n't know how to stop him .	it was the most beautiful thing to do .
2	she had no idea what to do next .	he turned away from the window and left .	it was not a good thing to do .
3	but she knew it would be a lie .	he looked at her , his expression grim .	it had been a long time since then .
4	she did n't want him to hurt her .	he stopped and stared at the closed door .	it was n't fair to say that .
5	she asked , her eyes pleading with concern .	he looked at zane , who was smiling .	it was like a punch to the gut .
6	she wrapped her arms around him and squeezed .	he could n't help but smile at her .	it was a good idea to walk home .
7	her hand trembled slightly , her heart racing .	he wanted to know where he was going .	it was the strangest thing in the world .
8	she went to the bathroom to get dressed .	his voice was low , his eyes narrowing .	it was almost three o'clock in the morning .
9	she twisted her arms around his neck again .	he waited for a moment to regain consciousness .	it was a long way from the city .
	Latent Category 4	**Latent Category 5**	**Latent Category 6**
1	i thought i was in love with him .	no , not your father , he said .	oh , shit , i 'm so sorry !
2	i was in a hurry to find out .	i mean , you know what i mean .	come on , let 's go to bed .
3	i took a deep breath and exhaled slowly .	oh , yeah , well , i guess .	he asked , as if to say something .
4	i was pretty sure i was a fan .	look at me , i love you too .	she asked , her voice dripping with sarcasm .
5	i did n't mean to be with you .	i dont know , she said to herself .	she asked , her eyes wide with annoyance .
6	i heard the door slam shut behind me .	im sorry , she said with a sigh .	she asked , her eyes pleading with concern .
7	my heart was pounding , my chest hurt .	yeah , well , that 's all right .	oh , shit , she said to herself .
8	i got up and went to the kitchen .	when she did , she didnt say anything .	he said , his voice thick with emotion .
9	i swallowed hard and tried to sound stupid .	i mean , what are you doing here ?	yes , of course , i 'm sure .
	Latent Category 7	**Latent Category 8**	**Latent Category 9**
1	" that 's what you 're doing here .	" that 's right , " he said .	" i love you , my love . "
2	" it 's all you have to say .	" i 'm fine , " she said .	" i 'm not a good man . "
3	" i do n't know where to go .	" i do n't , " ashe said .	" i do n't think so . "
4	" i did n't mean to upset you .	" i 'll go , " he says .	" i 'm sure it 's a trick . "
5	" you 've got to be kidding me .	" you 're beautiful , " she said .	" she did n't tell him that ! "
6	" she 's not going to kill you .	" it 's not , " he said .	" that 's why i 'm here . "
7	" that 's not what i 'm saying .	" you 're not , " he said .	" i 'll take care of you . "
8	" i 'm so glad you 're here .	" she 's not , " ethan said .	" maybe you can do that again ? "
9	" you 're going to be a hero .	" it 's okay , " i say .	" i thought you might like that . "

Results - Latent Space Interpolations

1	“ you ’re human , ” she said softly .	1	a girl on a stage holding a guitar .
2	“ you ’re a vampire , ” she said .	2	a girl on a stage holding a scythe .
3	“ you ’re a vampire , ” she said .	3	the girl in the sweat shirt plays a guitar on a stage .
4	“ you ’re a b**ch , ” she said .	4	the girl in the sweat vest is reading a newspaper .
5	“ you ’re angry , ” he said flatly .	5	a woman in a white shirt is taking a shortcut .
6	“ you ’re a jerk , ” he said dryly .	6	a woman in a white shirt is holding a scythe .
7	“ you ’re not going to let me out ? ”	7	a woman is playing the sax .
8	“ i do n’t know why you ’re here ? ”	8	a man is in the firehouse .
9	“ i do n’t know why you ’re here ? ”	9	a man is in the firehouse .
10	“ you do n’t know what to do ? ”	10	two men are in an enclosed room .
1	it is therefore quite impossible to separate the various provisions of the single market with regard to quotas .		
2	i do not therefore think that it is necessary to continue with a different view of the commission .		
3	i am therefore very sensitive to the issue of women in different member states and the social repercussions .		
4	therefore , i am very pleased with the report on women and their social rights in the member states .		
5	i am also very pleased to have the opportunity to discuss with the european parliament on this matter .		
6	i would also like to take the opportunity to comment on the issue of the european economic partnership .		
7	i would not like to mention the commission ’s statement on the issue of the council ’s statement .		
8	i do not want to answer any of the commissioner ’s questions to the commission .		
9	mr president , i am not going to reply to mr santer ’s statement on the lisbon strategy .		
10	mr president .		

Results - Conditional Latent Transformations

Label	Given Hypothesis	Generated Premise
E	the woman is very happy .	a lady wearing a blue white shirt is laughing .
C	no one is dancing .	a group of people playing guitar hero on a stage .
N	the man is reading the sportspage .	a man in a white shirt is sitting in a recliner .
E	a man is in a black shirt	a man in a black shirt stands in front of a store while a man in a blue hat and white shirt stands beside him .
C	an old woman has no jacket .	a woman with a white hat and jacket is playing with a girl in a red jacket .
N	a person is waiting for a train .	person in white and black hat standing in front of a train track .

Method	Accuracy
Random	41.1%
Baseline-Seq2Seq (Mean)	59.6%
Baseline-Seq2Seq (MOSM)	62.6%
Shen et. al Mean (N=1)	62.4%
Shen et. al MOSM (N=1)	75.9%
Ours	70.8%

Interpolation with gradient-based optimization

- Given two sentences x_1 and x_2, we formulate interpolation as a gradient based optimization problem that iteratively transforms x_1 into x_2.
- Let h_{x1} be the sentence representation of x_1.
- We interpolate by taking gradient steps along h_{x1} in order to reconstruct x_2.
- We start the optimization process at $h_0 = h_{x1}$ and take gradient steps as follows.

$$h_t = h_{t-1} + \alpha \nabla_{h_{t-1}} \log P(x_2 | h_{t-1})$$

- The probability of reconstruction is given by a fixed and train sentence representation decoder.

Comparison to linear interpolations

1	**of course, i had already made coffee and she headed right for the pot.**	**of course, i had already made coffee and she headed right for the pot.**
2	of course, she had already made coffee.	of course, i had already made coffee and she headed right for the pot.
3	colin had already made a pot of coffee.	of course, i had already made coffee.
4	colin pulled out the coffee pot .	i had a lot of things to do .
5	colin pulled out the file .	colin pulled the file out of his pocket .
6	colin pulled out the myers file .	colin colin pulled colin out of the colin .
7	**colin pulled out the myers file .**	**colin pulled out the myers file .**
1	**“ my mother struggled to make ends meet when i was a child .**	“ my mother struggled to make ends meet when i was a child .
2	“ my mother struggled to make ends meet .	“ my mother struggled to make ends meet when i was a child .
3	“ my mother would make ends meet .	“ my mother struggled to make ends meet .
4	“ my mother would ’ve loved you too .	i ’m so sorry , ” i said .
5	you would ’ve loved her mother ’s child .	i love you , i love you . ”
6	you would ’ve loved her . ”	you loved him , would n’t you ? ”
7	**you would ’ve loved her . ”**	**you would ’ve loved her . ”**

Table 9. Interpolations using gradient-based optimization (Left). Corresponding linear interpolations directly in the sentence representation space (Right). The two randomly selected sentences for each example from the BookCorpus are in bold.

Results - Gradient Based Interpolation

Start	her boyfriend eyed him curiously and gave him a cautious nod , then followed her into the bar , not waiting for a reply .
1	her boyfriend eyed him curiously , not waiting .
2	her boyfriend gave him a cautious nod .
3	her boyfriend never showed up at the bar .
4	grace gave him a cautious nod of approval .
5	grace never met her a boyfriend that night .
6	her boyfriend never showed up at the mall .
7	grace never showed up at work that night .
8	grace never showed up at her job in the mall .
9	grace never showed up at her job in the mall that night .
10	grace never showed up at her job in the mall that started at 5:00 that night .
End	grace never showed up at her job in the mall that started at 5:00 that night .
Start	naturally , he was upset when i said it was over .
1	naturally , he was upset when i said it was over .
2	he was upset when she said it .
3	naturally , he was upset when she said it .
4	naturally , she was upset when she said it .
5	no way , she said .
6	no way , she said , turning slightly .
7	no way , she said , turning her head slightly so she was nose to nose with him .
End	no way , she said , turning her head slightly so she was nose to nose with him .
Start	“ i ’ve been growing it out for a while .
1	“ i ’ve been growing out of it .
2	“ i ’ve been growing for a while .
3	“ i ’ve been listening for a while .
4	“ i ’ve been listening to a voice .
5	peter listen to me , you need it .
6	peter listen , you need to speak out .
7	peter listen , the voice didnt give him a chance to speak , you need to understand something .
End	peter listen , the voice didnt give him a chance to speak , you need to understand something .
Start	it ’s one of the lesser-known failings of the vampire .
1	it ’s one of the lesser-known failings of the vampire .
2	it ’s one of the failings of the vampire .
3	im sure it ’s one of the failings .
4	im sure it s one of their own .
5	im sure they got one of their own .
6	im sure theres a program of their own .
7	im sure they got a program of their own .
End	im sure cascadias got a program of their own , but it wouldnt be the same .