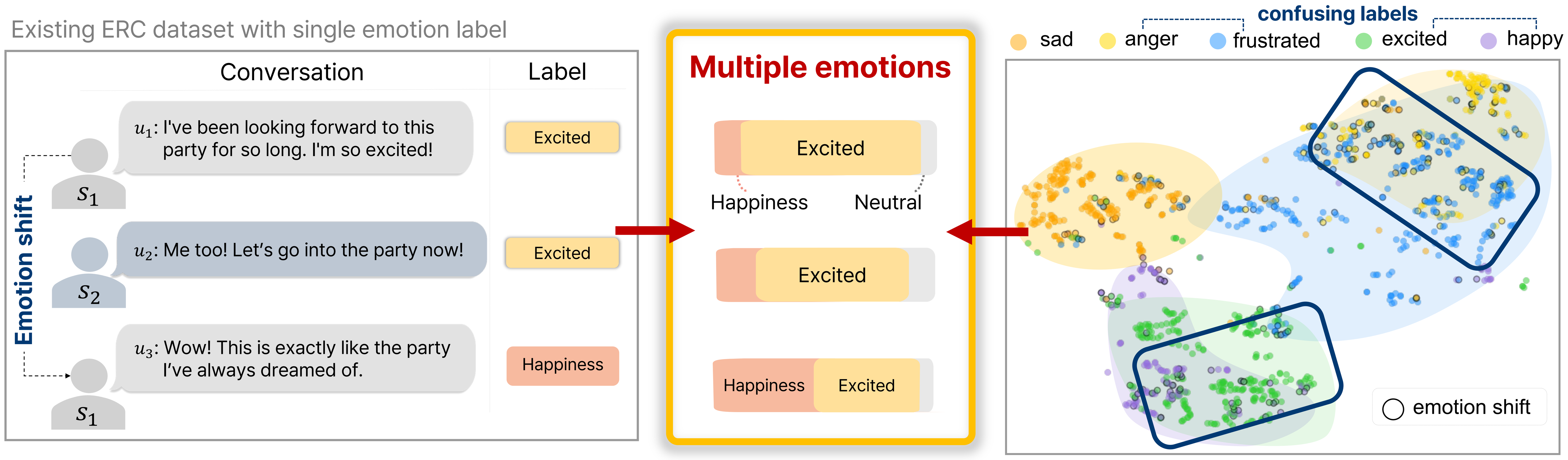
Beyond Single Emotion: Multi-label Approach to Conversational Emotion Recognition

Yujin Kang, Yoon-Sik Cho
Chung-Ang university, South Korea

Sunday (Mar 2) 14:00
Humans and AI session
Oral presentation

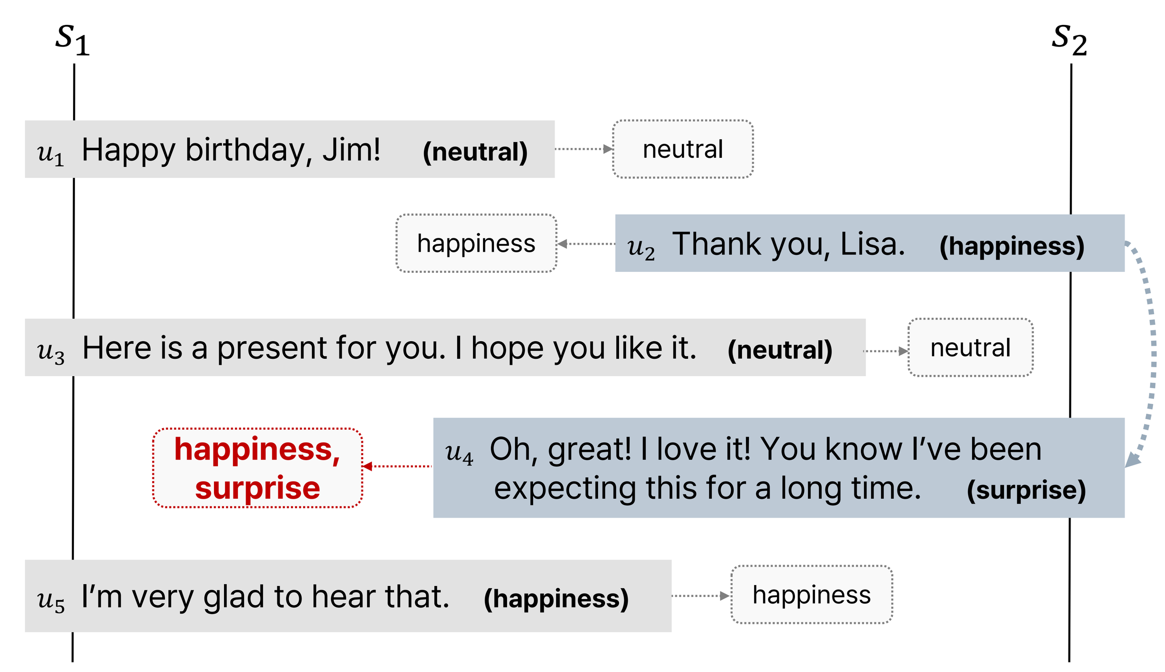
Motivation

- Emotion recognition in conversation(ERC) task aims to identify the emotion at each utterance in a conversation
- ERC suffers from **emotion shift** and **confusing label**
- Annotating a single emotion label to an utterance overlooks the possibility of **multiple emotions**



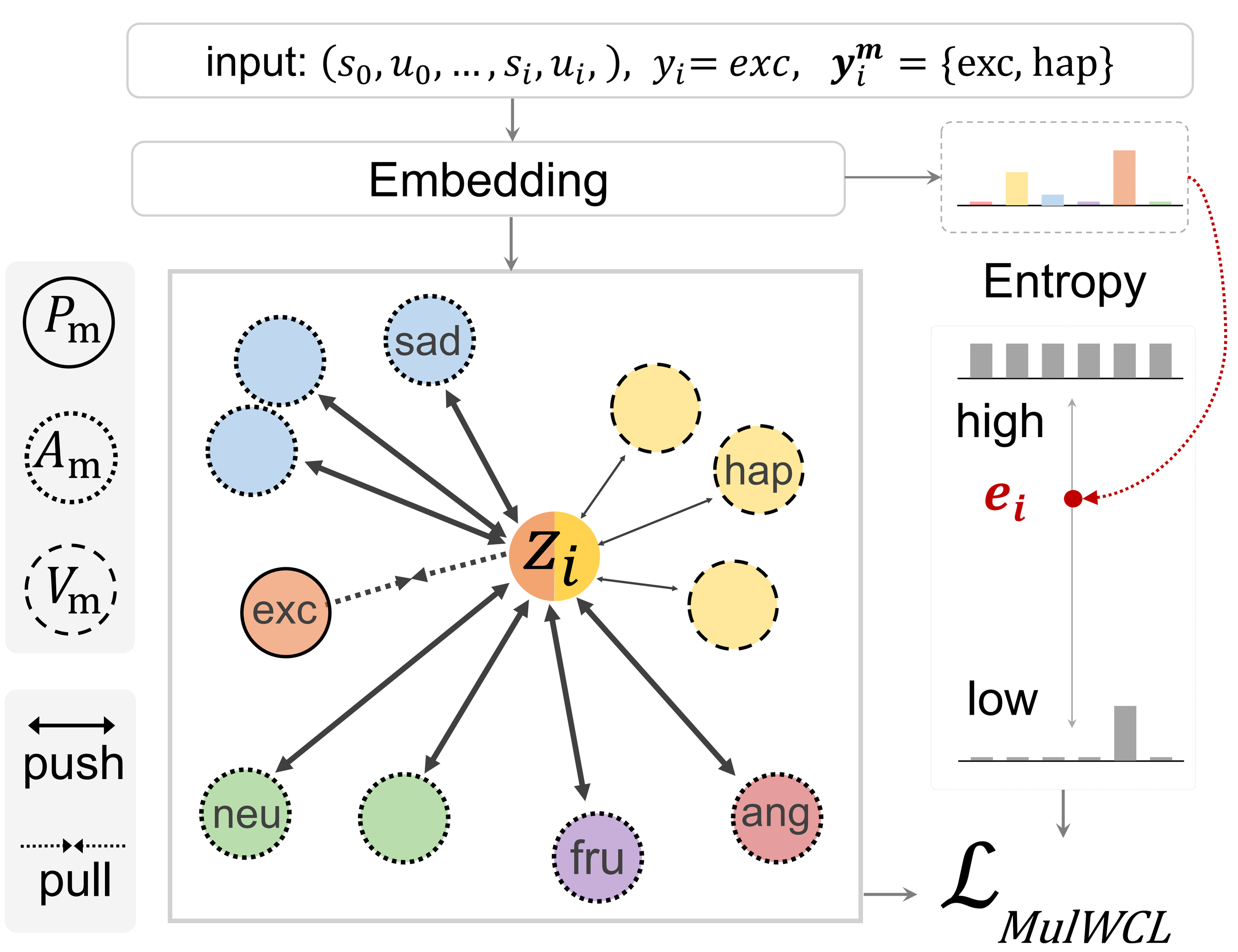
pseudo multi-emotion labels

1) Target an utterance if consecutive utterances of same speaker have different.
2) Combine emotions from consecutive utterances; otherwise, keep single label.



Methodology

Multi-label ERC (ML-ERC) Model



✓ ***Challenge1:*** finding the positive and negative pairs with multi-label vector is too complex

✅ We **redefine** the positive and negative sets

- $P_m(i) = \{p \mid y_p = y_i, p \neq i\}$
- $A_m(i) = \{a \mid y_p \notin \mathbf{m}_i, a \neq i\}$
- $V_m(i) = \{a \mid y_v \in (\mathbf{m}_i - y_i), a \neq i\}$

✓ ***Challenge2:*** multiple emotions cause overlapping positive and negative pairs

✅ We design **two weights** from different perspectives

- Class-level weighted score

$$\mathcal{P}_{\text{multi}}(i) = \sum_{p \in P_m(i)} \exp(\text{sim}(\mathbf{h}_i, \mathbf{h}_p)/\tau),$$

$$\mathcal{N}_{\text{multi}}(i) = \sum_{a \in A_m(i)} \exp(\text{sim}(\mathbf{h}_i, \mathbf{h}_a)/\tau) + \sum_{v \in V_m(i)} \boxed{\underbrace{\left(1 - \frac{(\text{sim}(r_i, r_v) + 1)}{2}\right)}_{\text{class weight}}} \cdot \exp(\text{sim}(\mathbf{h}_i, \mathbf{h}_v)/\tau)$$

- Instance-level weighted score

entropy

$$\mathcal{L}_{\text{MulWCL}}(i) = \boxed{\underbrace{(1 - e_i)}_{\text{instance weight}}} \left(\frac{-1}{|P_m(i)|} \log \frac{\mathcal{P}_{\text{multi}}(i)}{\mathcal{N}_{\text{multi}}(i)} \right)$$

✓ ***Challenge3:*** Our pseudo labeling scheme may still miss some utterances with multiple emotions

✅ We introduce **soft multi-labels** annotated to the potential utterances

$$\mathbf{p}_i^{\text{soft}} = \begin{cases} \mathbf{o}_i & \text{if } \left(\sum_{j=1}^{|K|} \mathbf{1}(y_j^{\text{pseudo}} \neq 0) = 1\right) \text{ and } (\gamma < e_i) \\ \mathbf{y}_i^{\text{pseudo}} & \text{otherwise} \end{cases}$$

The soft-labels are used as pseudo multi-labels for the data **in the next epoch**

Experiment

Model	EmoryNLP M-F1(↑)	EmoryNLP W-F1(↑)	EmoryNLP AUC(↑)	EmoryNLP HL(↓)	MELD M-F1(↑)	MELD W-F1(↑)	MELD AUC(↑)	MELD HL(↓)	IEMOCAP M-F1(↑)	IEMOCAP W-F1(↑)	IEMOCAP AUC(↑)	IEMOCAP HL(↓)
DialogueRNN	32.80	40.62	0.583	**0.2388**	34.37	56.61	0.612	0.1622	61.77	62.95	0.758	0.1440
DAG-ERC	36.47	41.33	0.586	0.3647	39.13	55.91	0.649	0.3038	58.34	57.97	0.772	0.2814
CoMPM	37.76	39.74	0.617	0.4018	39.80	55.12	0.714	0.2720	57.06	58.19	0.808	0.2803
MPLP	-	-	-	-	41.71	55.49	0.691	0.2466	59.39	59.99	0.811	0.2421
ML-ERC	**38.69**	**41.56**	**0.630**	0.2535	**50.03**	**63.01**	**0.718**	**0.1250**	**68.58**	**69.17**	**0.815**	**0.1312**

Results in multi-label classification

Loss	EmoryNLP M-F1(↑)	EmoryNLP W-F1(↑)	EmoryNLP AUC(↑)	EmoryNLP HL(↓)	MELD M-F1(↑)	MELD W-F1(↑)	MELD AUC(↑)	MELD HL(↓)	IEMOCAP M-F1(↑)	IEMOCAP W-F1(↑)	IEMOCAP AUC(↑)	IEMOCAP HL(↓)
BCE	34.57	37.12	0.597	0.2888	48.60	62.52	0.716	0.1256	65.41	66.28	0.816	0.1339
+ SupCon	36.63	39.38	0.620	0.2675	48.64	62.40	0.717	0.1258	65.35	67.23	0.796	0.1373
+ SCL	37.03	39.95	0.617	0.2632	48.97	62.76	0.712	0.1251	67.24	68.52	0.807	0.1325
+ JSCL	37.86	40.50	0.619	0.2674	49.88	62.28	0.714	**0.1248**	67.35	68.16	0.820	0.1339
+ JSPCL	33.46	35.77	0.589	0.2949	49.07	62.56	0.716	0.1261	66.18	66.98	**0.821**	0.1478
+ SLCL	17.27	23.23	0.503	0.3134	39.19	58.63	0.668	0.1496	61.12	62.11	0.792	0.1869
+ ICL	34.24	36.56	0.592	0.3058	49.02	63.10	0.720	0.1313	62.34	63.34	0.794	0.1698
+ MulWCL	**38.05**	**40.88**	**0.624**	**0.2566**	**50.16**	**63.28**	**0.721**	0.1260	**68.42**	**68.96**	0.817	**0.1288**

Comparisons against multi-label contrastive losses

Base Model	Objective	Confusing labels (↓) Peaceful - Happy	Powerful - Happy	Sad – Fear	Angry - Frustrated	Excited - Happy
RoBERTa	$\mathcal{L}_{\text{ERC}}$	34.14	38.96	28.67	26.39	40.77
	$+\mathcal{L}_{\text{ML-ERC}}$	**33.60 (-0.54)**	**32.37 (-6.59)**	**24.81 (-3.86)**	**23.24 (-3.15)**	**24.44 (-16.33)**
CoMPM	$\mathcal{L}_{\text{ERC}}$	33.25	34.32	30.89	28.20	29.10
	$+\mathcal{L}_{\text{ML-ERC}}$	**26.94 (-6.31)**	**33.91 (-0.41)**	**24.97 (-5.92)**	**22.16 (-6.04)**	**25.66 (-3.44)**
DAG-ERC	$\mathcal{L}_{\text{ERC}}$	**24.07**	35.94	25.52	19.16	33.78
	$+\mathcal{L}_{\text{ML-ERC}}$	24.76 (+0.69)	**32.15 (-3.79)**	**25.42 (-0.10)**	**18.95 (-0.21)**	**33.01 (-0.77)**

The misclassified rate (lower the better) as confusing labels on ERC datasets

Base Model	Loss	Dataset EMORY	MELD	IEMOCAP
RoBERTa	$\mathcal{L}_{\text{ERC}}$	33.08	64.27	63.58
	$+\mathcal{L}_{\text{ML-ERC}}$	**35.74**	**65.15**	**63.96**
DialogueRNN	$\mathcal{L}_{\text{ERC}}$	37.44	57.03	62.75
	$+ \mathcal{L}_{\text{ML-ERC}}$	**38.18**	**57.51**	**63.56**
DAG-ERC	$\mathcal{L}_{\text{ERC}}$	38.85	63.38	67.07
	$+ \mathcal{L}_{\text{ML-ERC}}$	**39.02**	**63.58**	**68.12**
CoMPM	$\mathcal{L}_{\text{ERC}}$	36.20	64.87	66.47
	$+ \mathcal{L}_{\text{ML-ERC}}$	**37.35**	**65.90**	**68.00**
MPLP	$\mathcal{L}_{\text{ERC}}$	-	65.45	65.03
	$+ \mathcal{L}_{\text{ML-ERC}}$	-	**65.93**	**66.09**

Results in single-label classification

Base Model	Loss	Dataset EMORY	MELD	IEMOCAP
RoBERTa	$\mathcal{L}_{\text{ERC}}$	29.17	51.37	45.01
	$+ \mathcal{L}_{\text{ML-ERC}}$	**31.37**	**55.15**	**48.20**
DialogueRNN	$\mathcal{L}_{\text{ERC}}$	36.81	**41.87**	46.86
	$+ \mathcal{L}_{\text{ML-ERC}}$	**37.97**	41.82	**53.41**
DAG-ERC	$\mathcal{L}_{\text{ERC}}$	**35.81**	50.88	63.25
	$+ \mathcal{L}_{\text{ML-ERC}}$	35.01	**50.95**	**63.78**
CoMPM	$\mathcal{L}_{\text{ERC}}$	22.50	50.03	**46.14**
	$+ \mathcal{L}_{\text{ML-ERC}}$	**29.79**	**53.45**	45.60
MPLP	$\mathcal{L}_{\text{ERC}}$	-	52.45	49.34
	$+ \mathcal{L}_{\text{ML-ERC}}$	-	**55.28**	**53.78**

Results in emotion shift data

Conclusion

- We propose ML-ERC, for Multi-label classification to overcome the limitations of single-label emotion assignment.
- ML-ERC utilizes *MulWCL*, an objective function specialized for multi-emotion scenarios, and integrated a soft-labeling technique.
- The empirical results on existing task with single label support the efficacy of our approach, which is more effective in the most challenging settings: emotion shift or confusing labels.