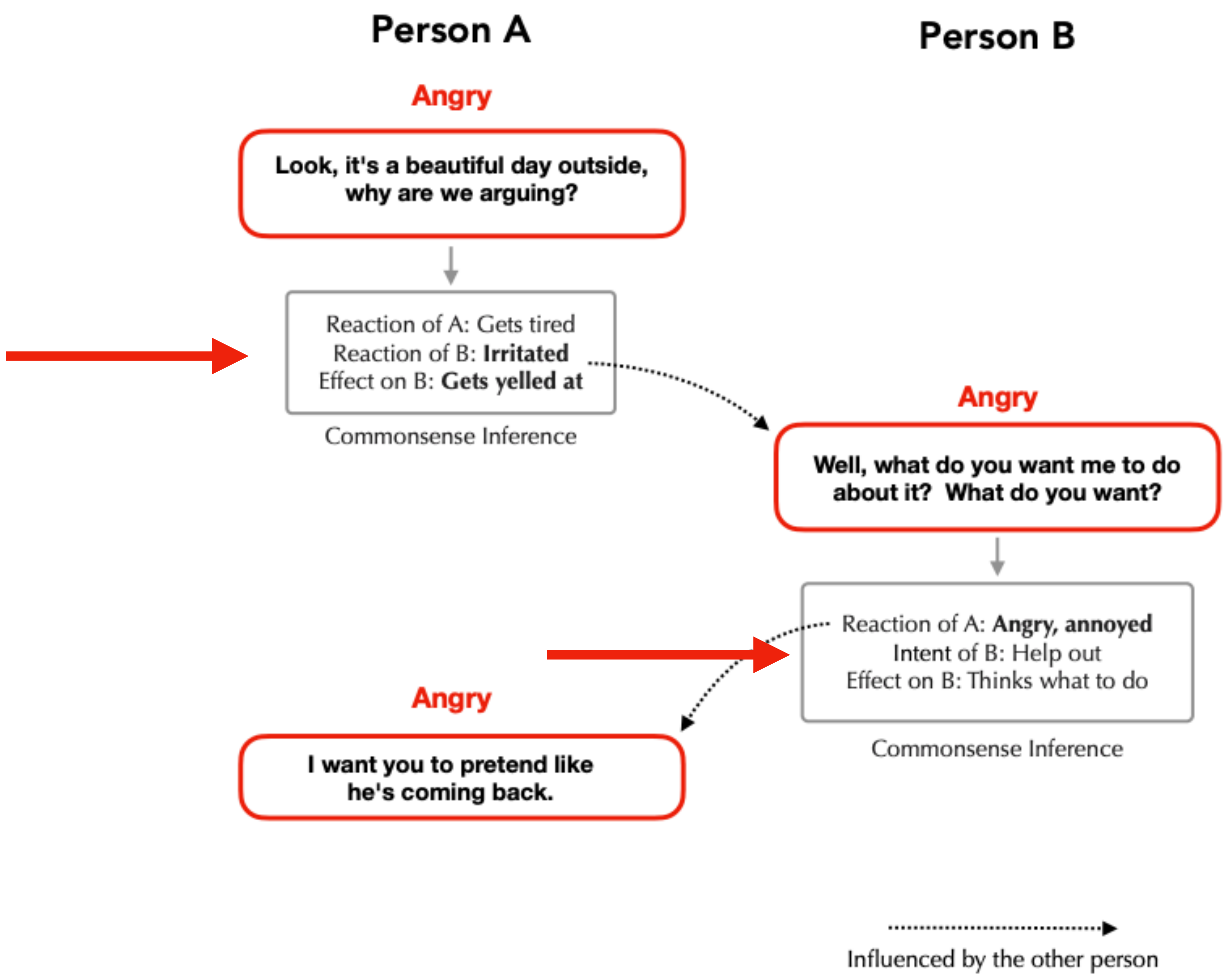
COSMIC

Commonsense knowledge for eMotion Identification in Conversations


Person A
Person B
Angry
Look, it's a beautiful day outside, why are we arguing?
Reaction of A: Gets tired
Reaction of B: Irritated
Effect on B: Gets yelled at
Commonsense Inference
Angry
Well, what do you want me to do about it? What do you want?
Reaction of A: Angry, annoyed
Intent of B: Help out
Effect on B: Thinks what to do
Commonsense Inference
Angry
I want you to pretend like he's coming back.
Influenced by the other person

In this conversation model, only the utterances can be observed as the conversation unfolds, while other variables such as speaker state and intent remain latent as they are not directly observed by the other participants.

$$[(u1, p1),(u2, p2), . . . ,(uN , pN)]$$

$$\big\downarrow$$

$$ui = [ui,1, ui,2, . . . , ui,T] \qquad pi$$

$$\big\downarrow \text{predict}$$

$$ei$$

1. Context independent feature extraction from pretrained transformer language models.

2. Commonsense feature extraction from a commonsense knowledge graph.

3. Incorporating commonsense knowledge to design better contextual representations and using it for the final emotion classification.

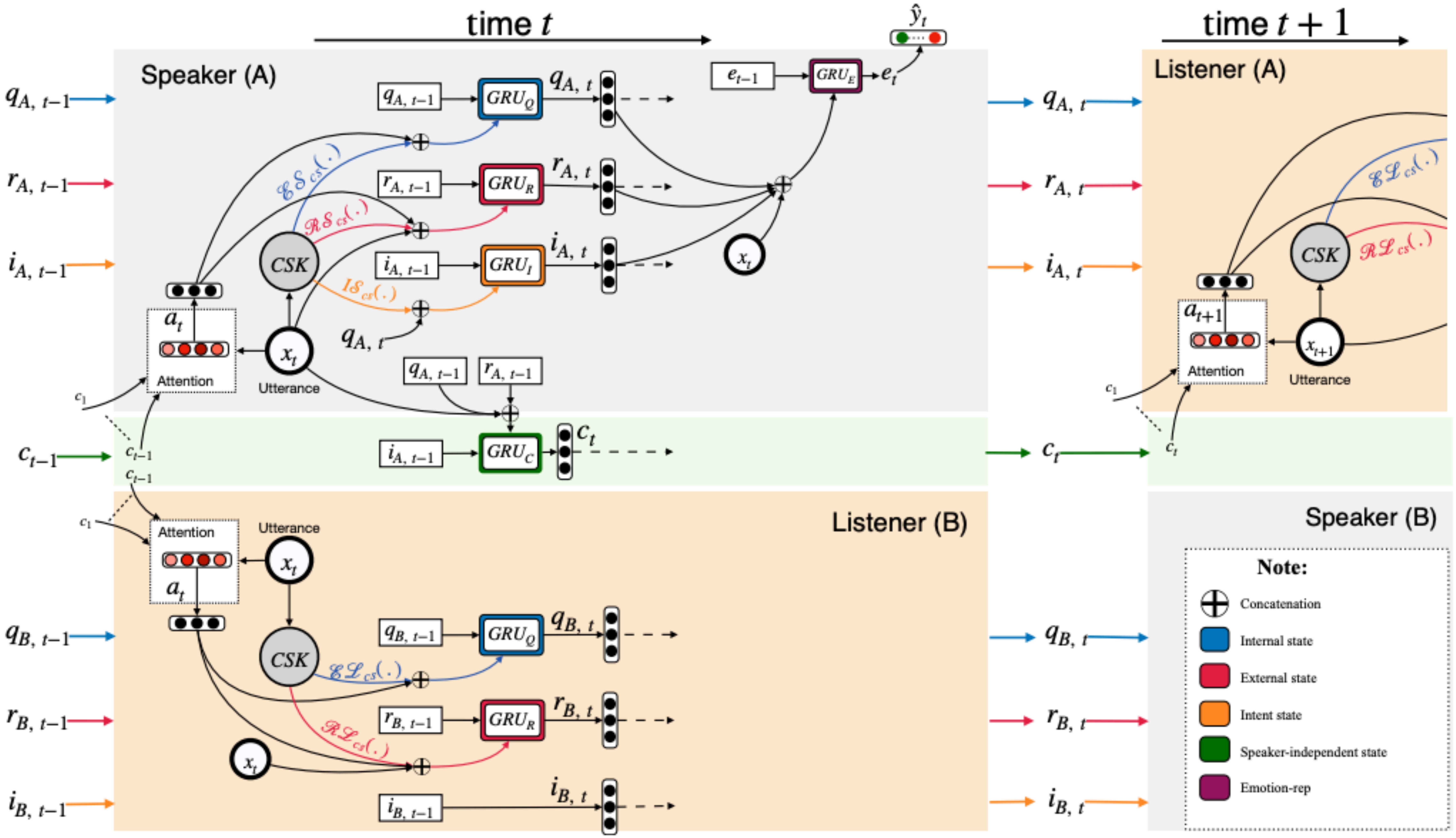


Figure 2: Illustration of COSMIC framework. *CSK*: Commonsense knowledge from COMET. In practice we use Bidirectional GRU cells. However, for clarity unidirectional cells are shown in the sketch.

Context Independent Feature Extraction

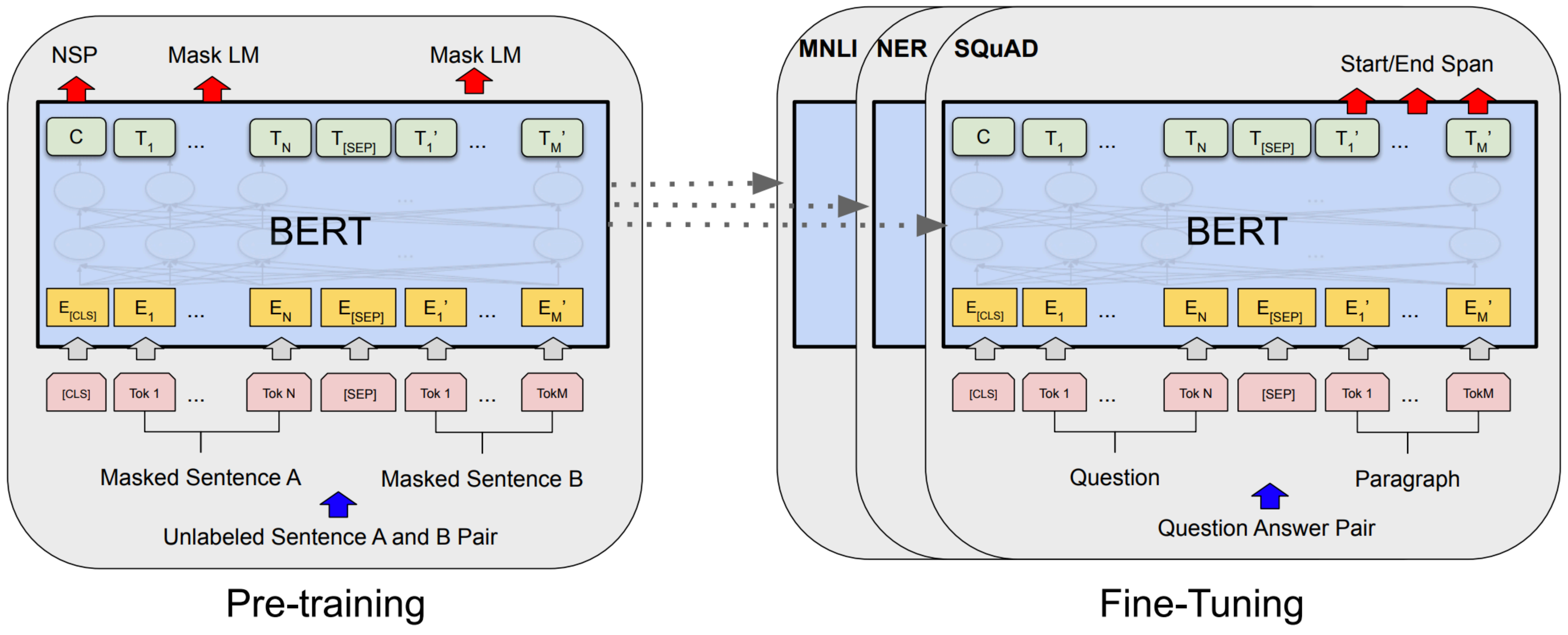


Figure 1: Overall pre-training and fine-tuning procedures for BERT. Apart from output layers, the same architectures are used in both pre-training and fine-tuning. The same pre-trained model parameters are used to initialize models for different down-stream tasks. During fine-tuning, all parameters are fine-tuned. `[CLS]` is a special symbol added in front of every input example, and `[SEP]` is a special separator token (e.g. separating questions/answers).

RoBERTa: A Robustly Optimized BERT Pretraining Approach

Commonsense Feature Extraction

COMET : Commonsense Transformers for Automatic Knowledge Graph Construction

ATOMIC Input Template and ConceptNet Relation-only Input Template

s tokens	mask tokens	*r* token	*o* tokens
PersonX goes to the mall	[MASK]	<xIntent>	to buy clothes

ConceptNet Relation to Language Input Template

s tokens	mask tokens	*r* tokens	mask tokens	*o* tokens
go to mall	[MASK] [MASK]	has prerequisite	[MASK]	have money

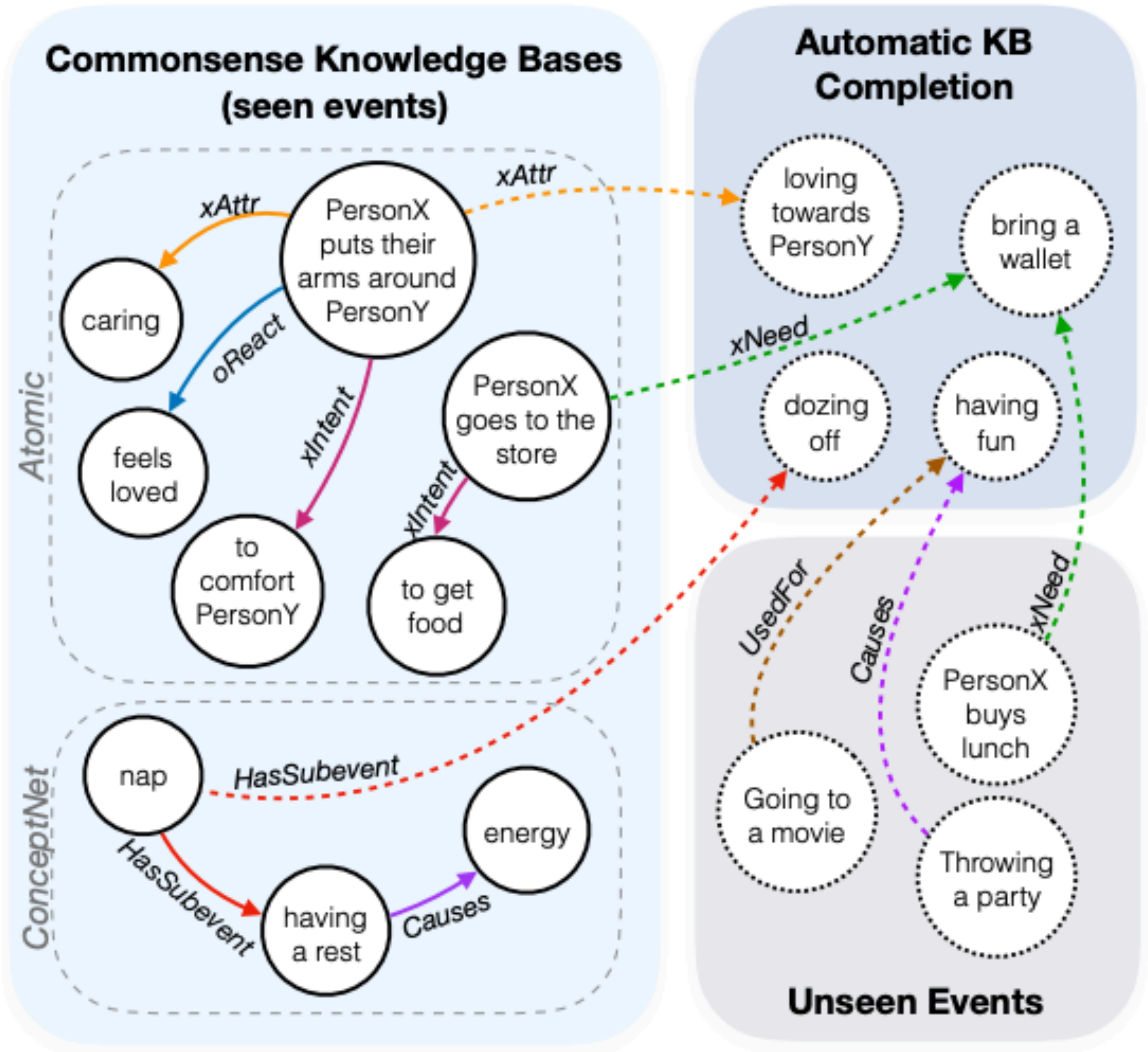


Figure 1: COMET learns from an existing knowledge base (solid lines) to be able to generate novel nodes and edges (dashed lines).

Commonsense Feature Extraction

$$\{s, r, \longrightarrow o\}$$

Commonsense Feature Extraction

$$\{U, r, \longrightarrow o\}$$

Commonsense Feature	Notation	Nature	Causal Relation
Intent of speaker	$\mathcal{IS}_{cs}(.)$	Mental state	Cause
Effect on speaker	$\mathcal{ES}_{cs}(.)$	Mental state	Effect
Reaction of speaker	$\mathcal{RS}_{cs}(.)$	Event	Effect
Effect of listeners	$\mathcal{EL}_{cs}(.)$	Mental state	Effect
Reaction of listeners	$\mathcal{RL}_{cs}(.)$	Event	Effect

State	Influenced By
Context State	Utterance, Internal state, External state
Internal State	Context state, *Effect on speaker, listener*
External State	Context state, Utterance, *Reaction of speaker, listener*
Intent State	Internal state, *Intent of speaker*
Emotion State	Utterance, Intent state Internal state, External state

Table 2: Different states and the respective variables they are influenced by. *Italic* variables are forms of commonsense knowledge from Table 1.

c **Context state**

q **Internal state**

r **External state**

i **Intent state**

e **Emotion state**

Mental state

Context state

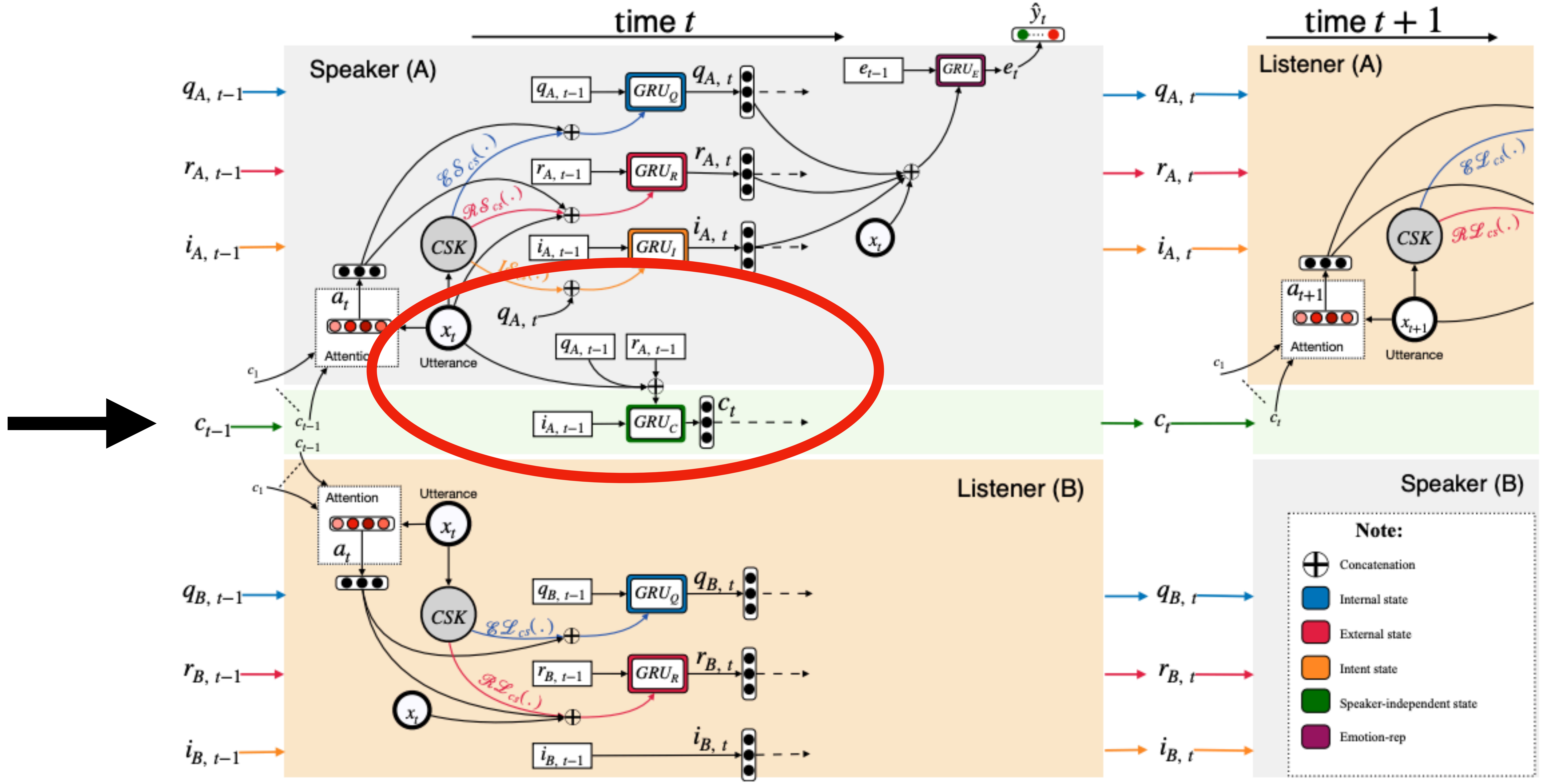


Figure 2: Illustration of COSMIC framework. *CSK*: Commonsense knowledge from COMET. In practice we use Bidirectional GRU cells. However, for clarity unidirectional cells are shown in the sketch.

Internal state

ES
Effect on Speaker
Listener

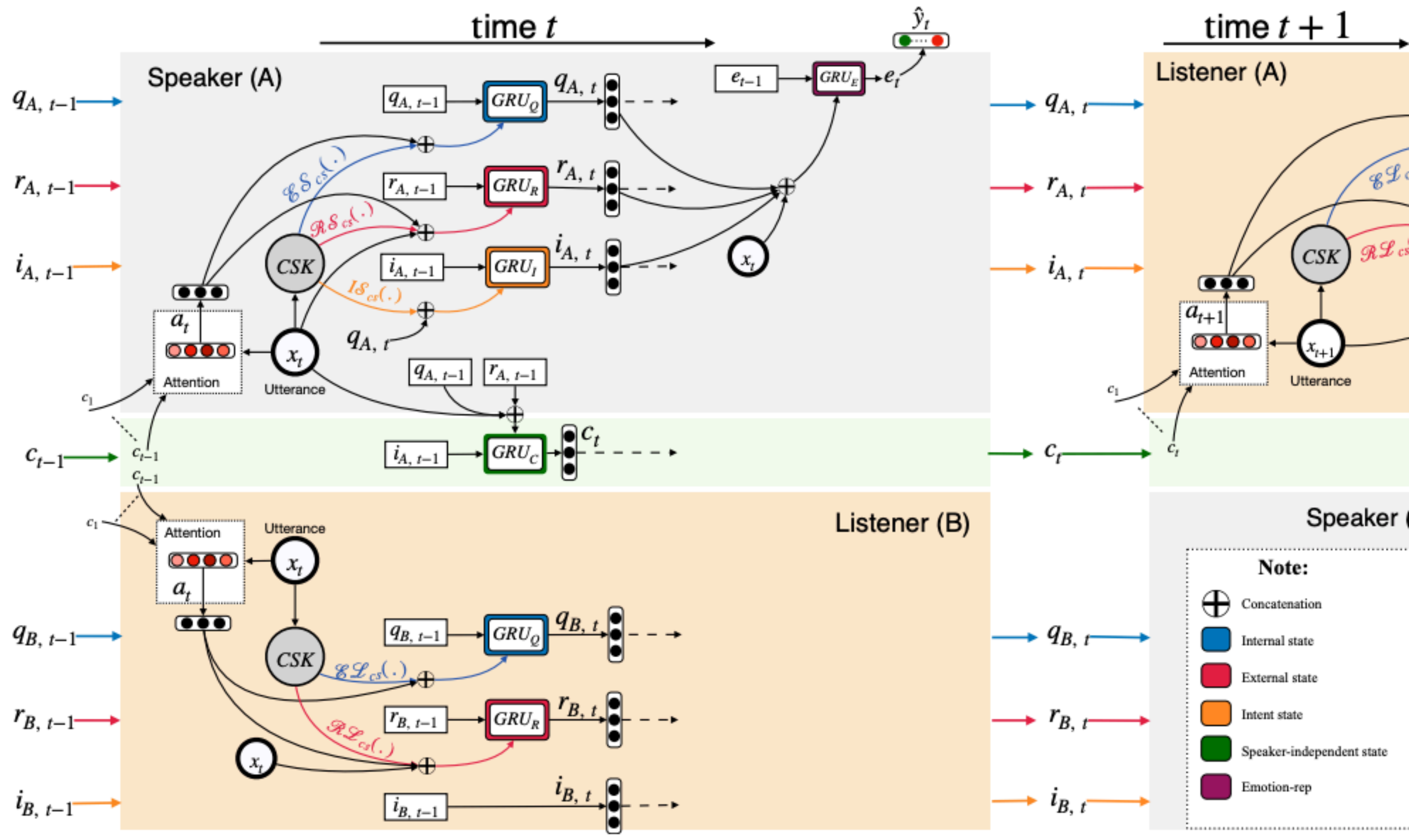


Figure 2: Illustration of COSMIC framework. *CSK*: Commonsense knowledge from COMET. In practice w
Bidirectional GRU cells. However, for clarity unidirectional cells are shown in the sketch.

External state

RS
Reaction on Speaker
Listener

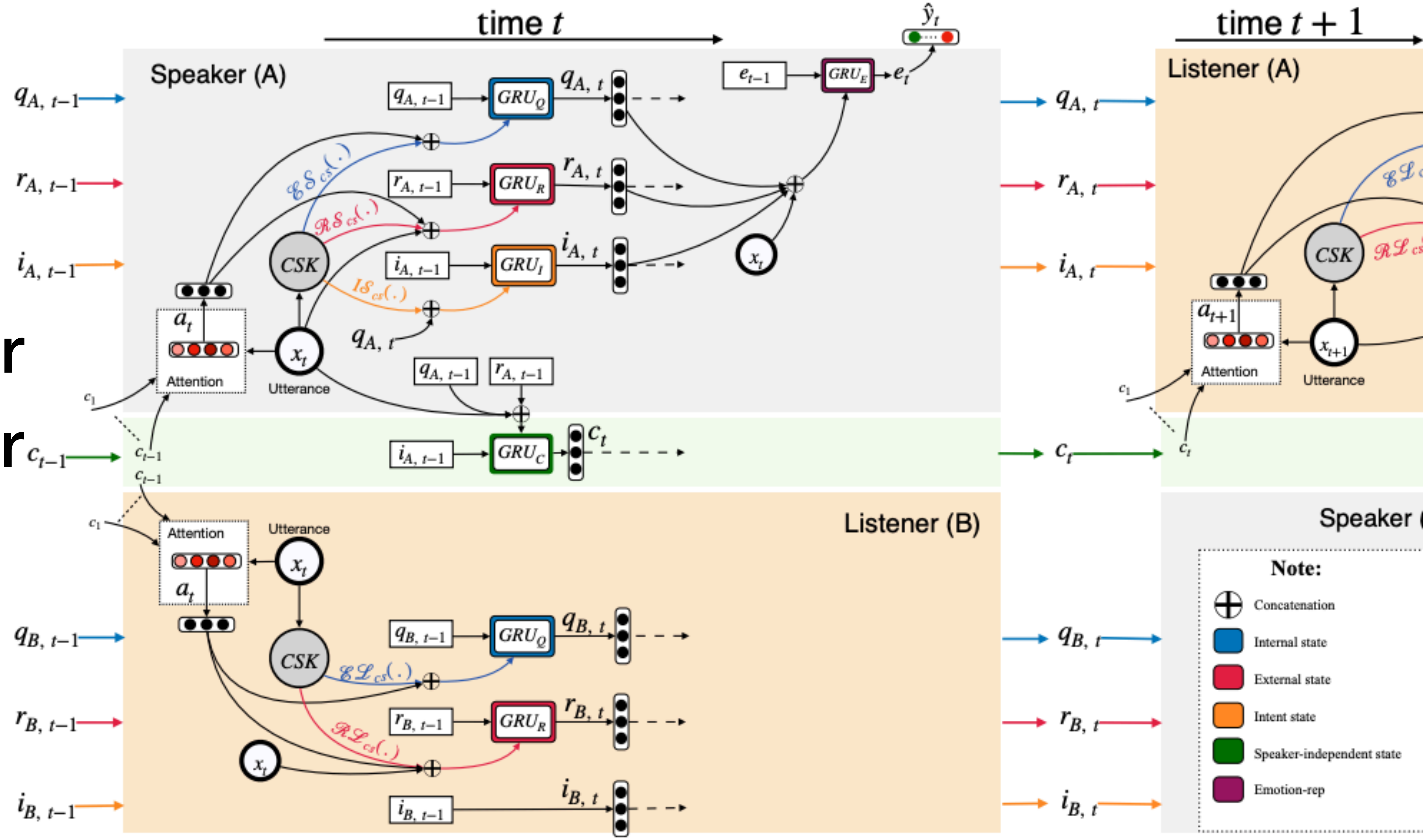


Figure 2: Illustration of COSMIC framework. *CSK*: Commonsense knowledge from COMET. In practice w
Bidirectional GRU cells. However, for clarity unidirectional cells are shown in the sketch.

Intent state

IS

Intent on Speaker

Listener

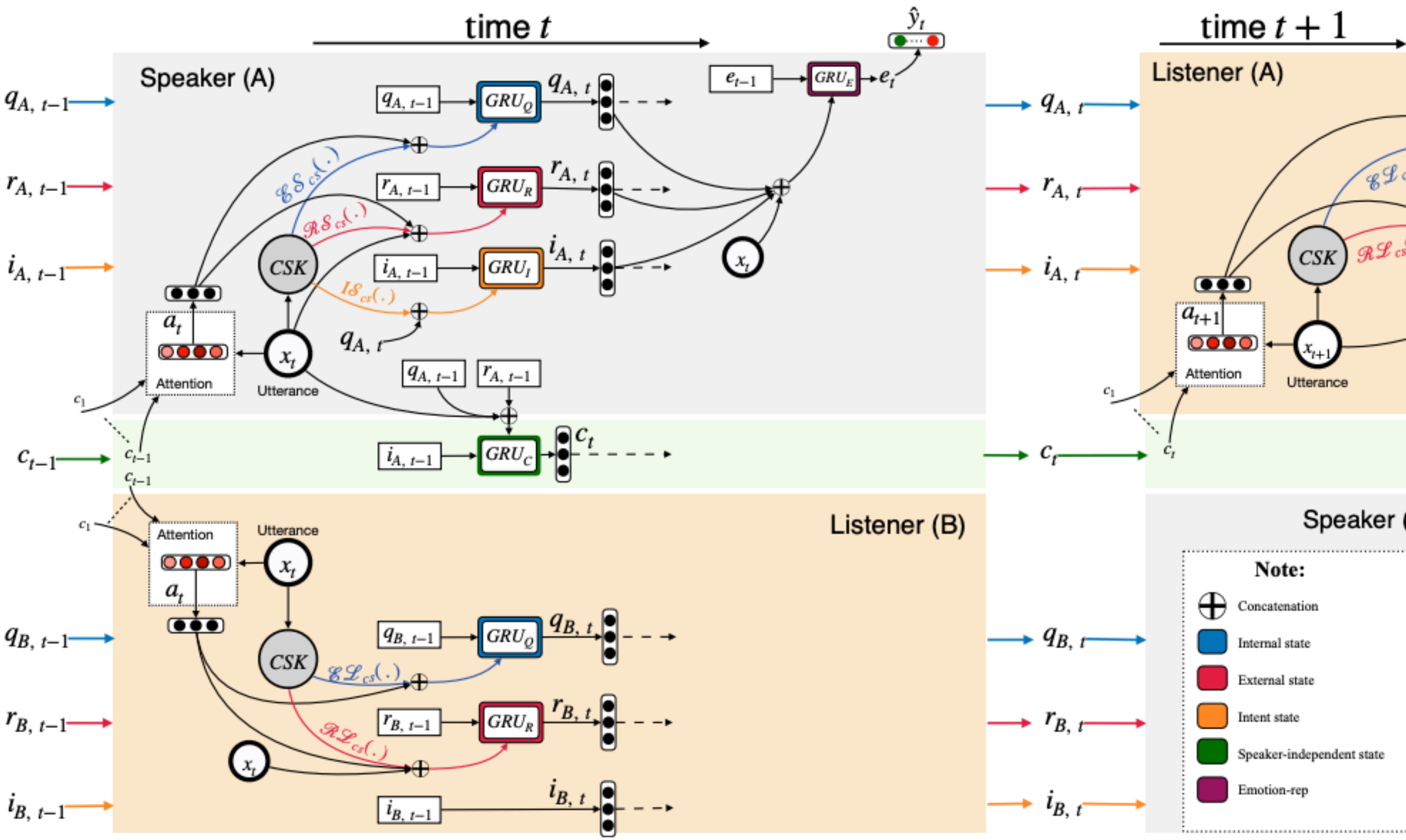


Figure 2: Illustration of COSMIC framework. *CSK*: Commonsense knowledge from COMET. In practice w Bidirectional GRU cells. However, for clarity unidirectional cells are shown in the sketch.

Emotion state

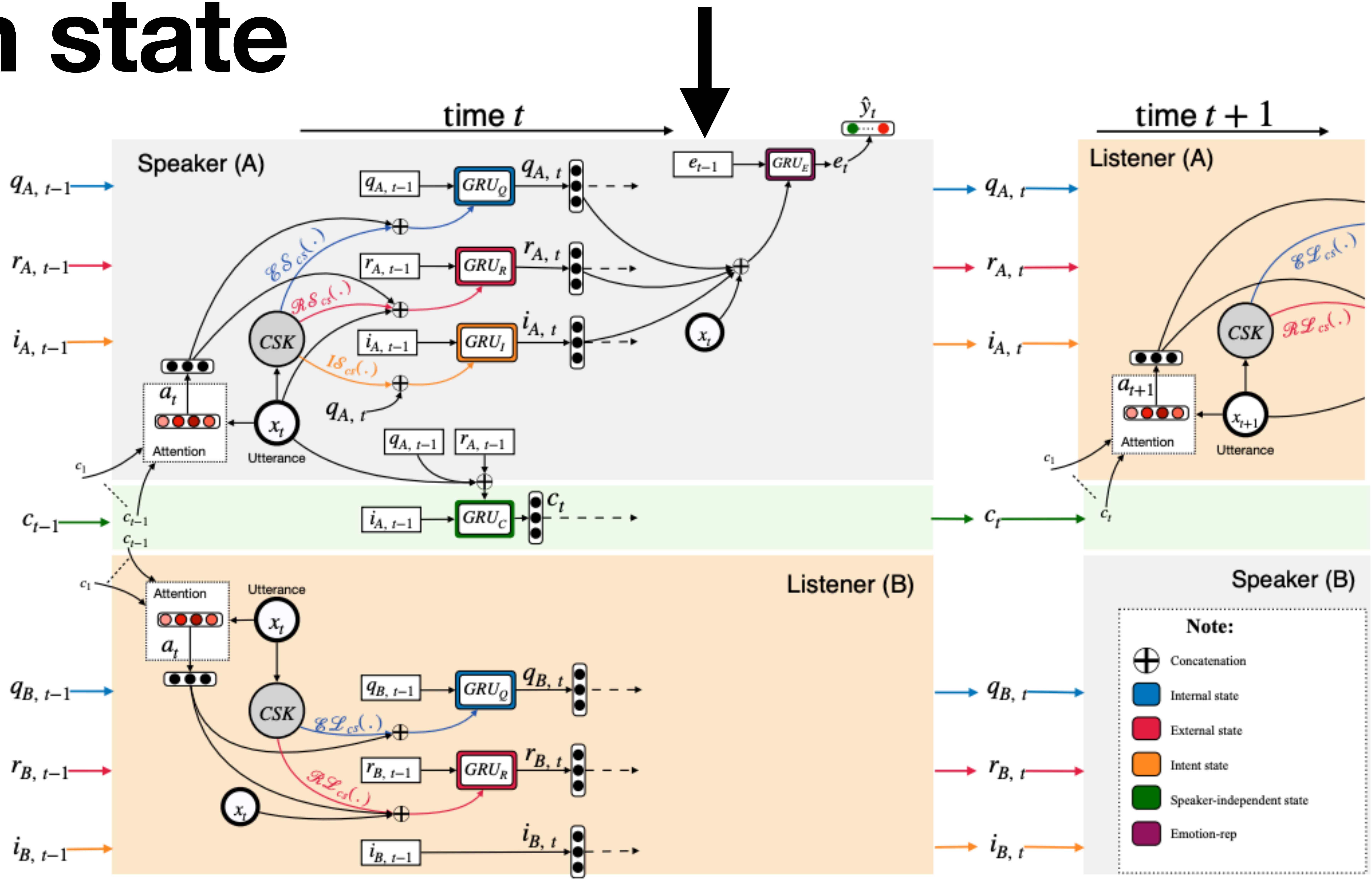


Figure 2: Illustration of COSMIC framework. *CSK*: Commonsense knowledge from COMET. In practice we use Bidirectional GRU cells. However, for clarity unidirectional cells are shown in the sketch.

Dataset

IEMOCAP

4.1 Datasets

Dataset	# dialogues			# utterances		
	train	val	test	train	val	test
IEMOCAP	120	12	31	5810		1623
DailyDialog	11,118	1,000	1,000	87,832	7,912	7,863
MELD	1,039	114	280	9,989	1,109	2,610
EmoryNLP	659	89	79	7,551	954	984

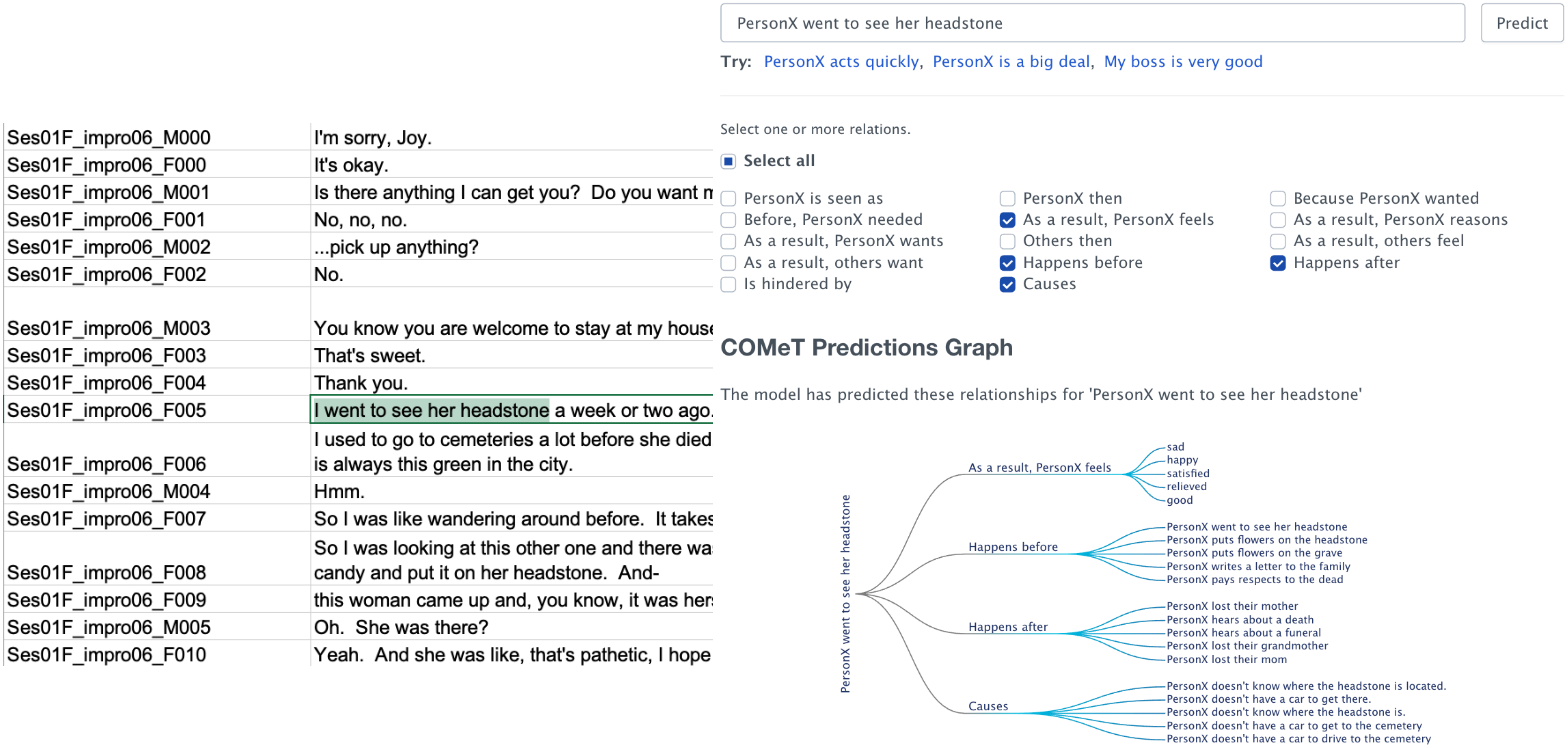
Dataset	# classes	Metric
IEMOCAP	6	Weighted Avg. F1
DailyDialog	7*	Macro F1 and Micro F1
MELD	3 and 7	Weighted Avg. F1 over 3 and 7 classes
EmoryNLP	3 and 7	Weighted Avg. F1 over 3 and 7 classes

Table 3: Statistics of splits and evaluation metrics used in different datasets. In MELD and EmoryNLP evaluation is performed for 3 class (broad) and 7 class (fine-grained) classification. *Neutral** classes constitutes to 83% of the DailyDialog dataset. These are excluded when calculating the Micro F1 score.

happy
sad
neutral
angry
excited
frustrated state

Result

	Methods	IEMOCAP	DailyDialog		MELD		EmoryNLP	
		W-Avg F1	Macro F1	Micro F1	W-Avg F1 (3-cls)	W-Avg F1 (7-cls)	W-Avg F1 (3-cls)	W-Avg F1 (7-cls)
GloVe-based	CNN	52.04	36.87	50.32	64.25	55.02	38.05	32.59
	ICON	58.54	-	-	-	-	-	-
	KET	59.56	-	53.37	-	58.18	-	34.39
	ConGCN	-	-	-	-	57.40	-	-
	DialogueRNN	62.57	41.80	55.95	66.10	57.03	48.93	31.70
(Ro)BERT(a)-based	BERT DCR-Net	-	48.90	-	-	-	-	-
	BERT+MTL	-	-	-	-	61.90	-	35.92
	RoBERTa	54.55	48.20	55.16	72.12	62.02	55.28	37.29
	RoBERTa DialogueRNN	64.76	49.65	57.32	72.14	63.61	55.36	37.44
	COSMIC	**65.28**	**51.05**	**58.48**	**73.20**	**65.21**	**56.51**	**38.11**
	w/o Speaker CSK	63.27	50.18	57.45	72.94	64.41	55.46	37.35
	w/o Listener CSK	65.05	48.67	58.28	72.90	64.76	**56.57**	**38.15**
	w/o Speaker, Listener CSK	63.05	48.68	56.16	72.62	64.28	55.34	37.10



PersonX went to see her headstone
Predict
Try: PersonX acts quickly, PersonX is a big deal, My boss is very good
Ses01F_impro06_M000 I'm sorry, Joy.
Ses01F_impro06_F000 It's okay.
Ses01F_impro06_M001 Is there anything I can get you? Do you want n
Ses01F_impro06_F001 No, no, no.
Ses01F_impro06_M002 ...pick up anything?
Ses01F_impro06_F002 No.
Ses01F_impro06_M003 You know you are welcome to stay at my house
Ses01F_impro06_F003 That's sweet.
Ses01F_impro06_F004 Thank you.
Ses01F_impro06_F005 I went to see her headstone a week or two ago.
Ses01F_impro06_F006 I used to go to cemeteries a lot before she died is always this green in the city.
Ses01F_impro06_M004 Hmm.
Ses01F_impro06_F007 So I was like wandering around before. It takes
Ses01F_impro06_F008 So I was looking at this other one and there wa candy and put it on her headstone. And-
Ses01F_impro06_F009 this woman came up and, you know, it was her
Ses01F_impro06_M005 Oh. She was there?
Ses01F_impro06_F010 Yeah. And she was like, that's pathetic, I hope
Select one or more relations.
Select all
PersonX is seen as
Before, PersonX needed
As a result, PersonX wants
As a result, others want
Is hindered by
PersonX then
As a result, PersonX feels
Others then
Happens before
Causes
Because PersonX wanted
As a result, PersonX reasons
As a result, others feel
Happens after
COMeT Predictions Graph
The model has predicted these relationships for 'PersonX went to see her headstone'
PersonX went to see her headstone
As a result, PersonX feels
sad
happy
satisfied
relieved
good
Happens before
PersonX went to see her headstone
PersonX puts flowers on the headstone
PersonX puts flowers on the grave
PersonX writes a letter to the family
PersonX pays respects to the dead
Happens after
PersonX lost their mother
PersonX hears about a death
PersonX hears about a funeral
PersonX lost their grandmother
PersonX lost their mom
Causes
PersonX doesn't know where the headstone is located.
PersonX doesn't have a car to get there.
PersonX doesn't know where the headstone is.
PersonX doesn't have a car to get to the cemetery
PersonX doesn't have a car to drive to the cemetery

COMeT Predictions Graph

The model has predicted these relationships for 'PersonX went to see her headstone'