

Learning Interactions and Relationships between Movie Characters



Anna Kukleva

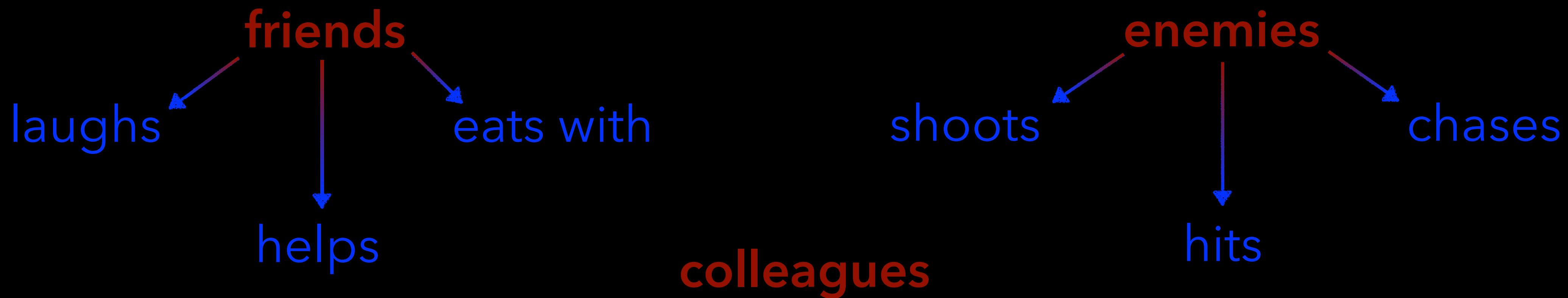


Makarand Tapaswi



Ivan Laptev

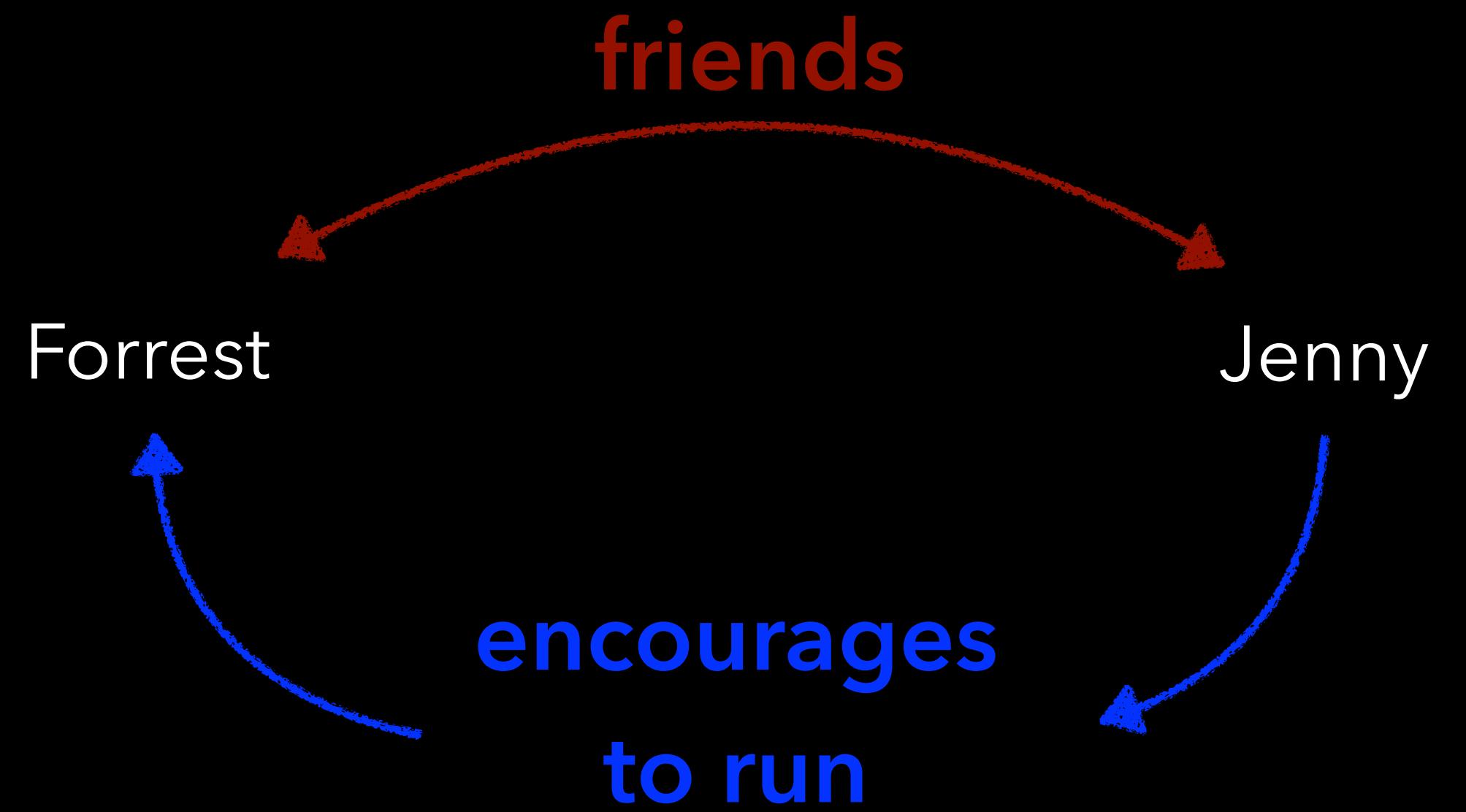
Interactions and Relationships



Assistive robotics



How to understand human behavior?



Our goal:

Who?

Jenny → Forrest

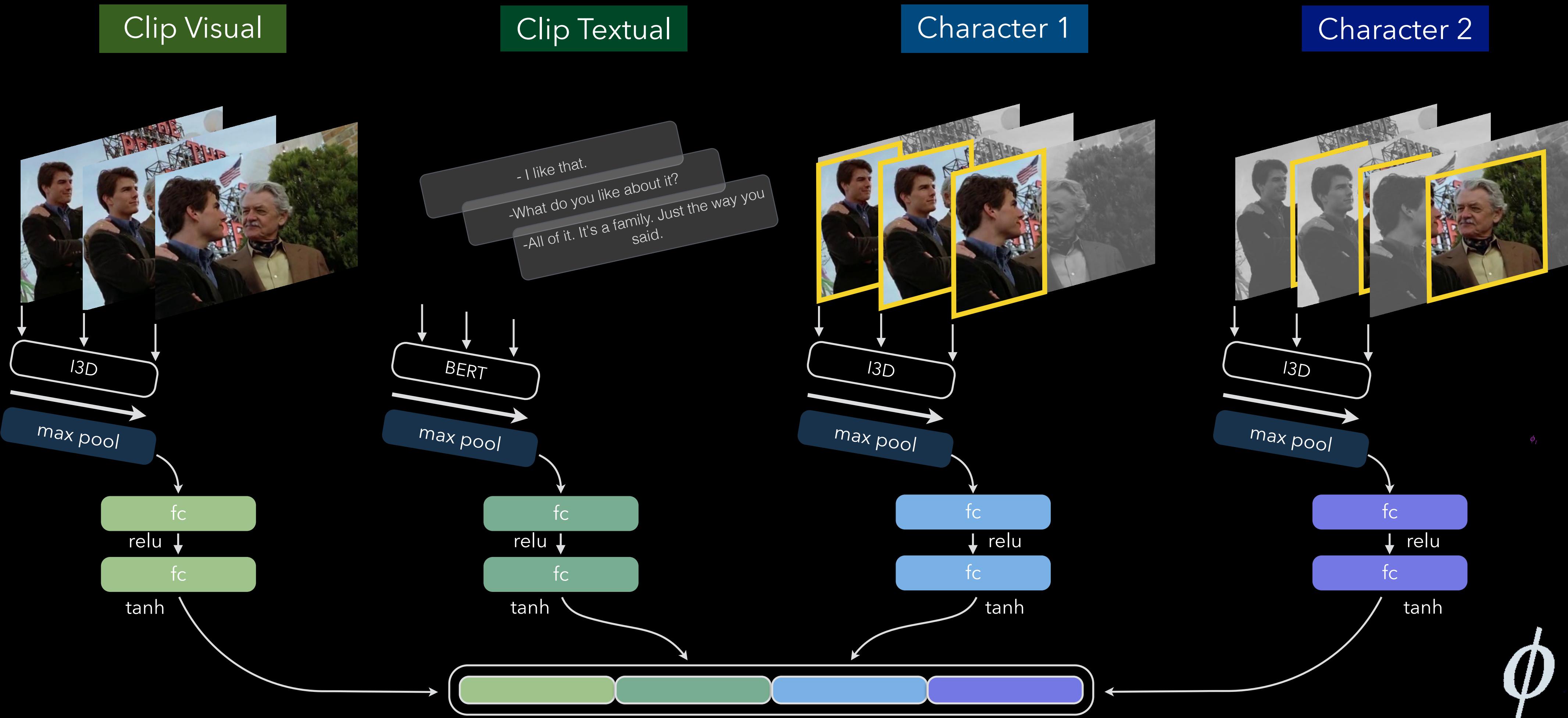
What are they doing?

encourages
to run

What is their relationship?

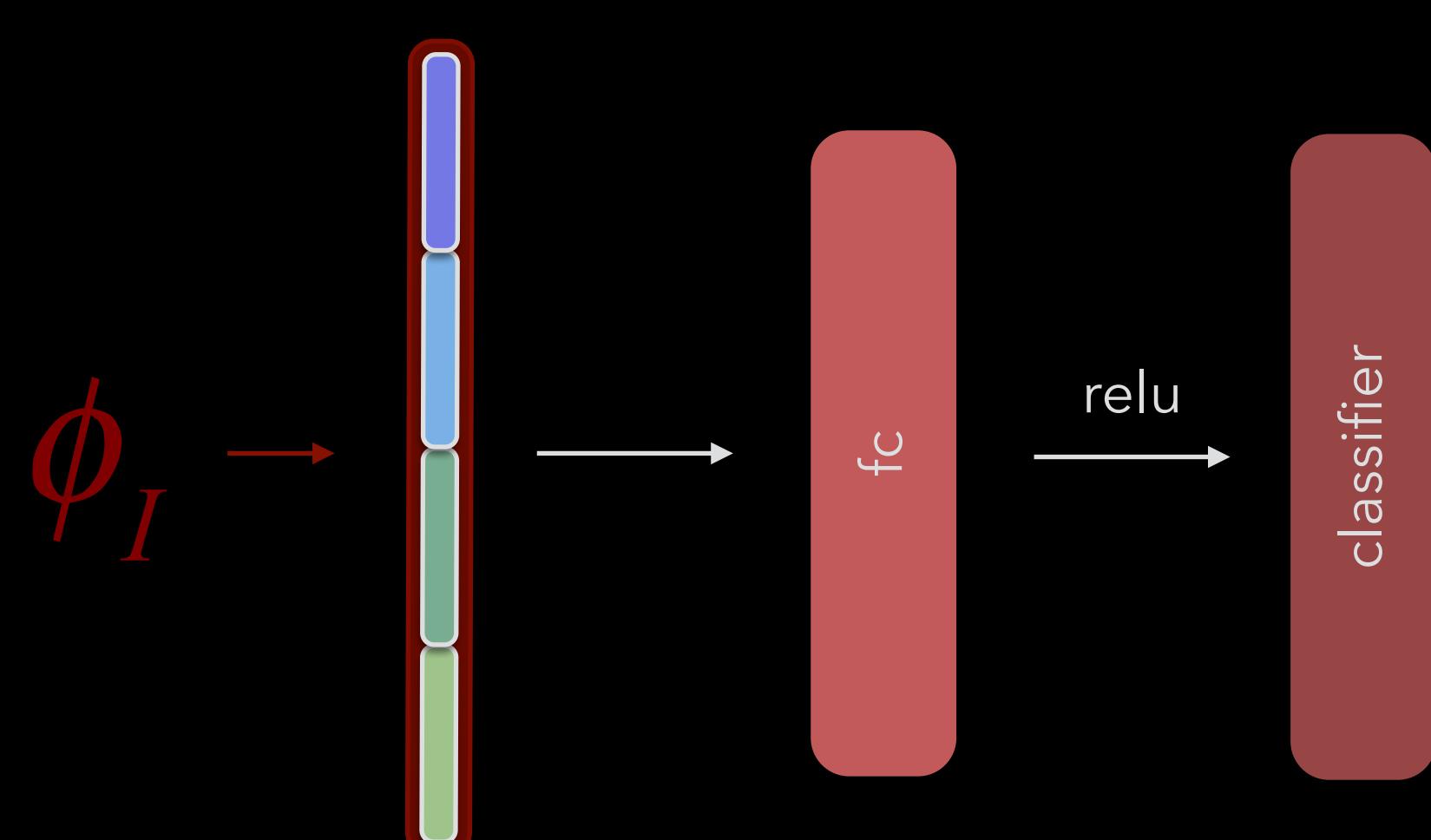
friends

Multimodal clip representation

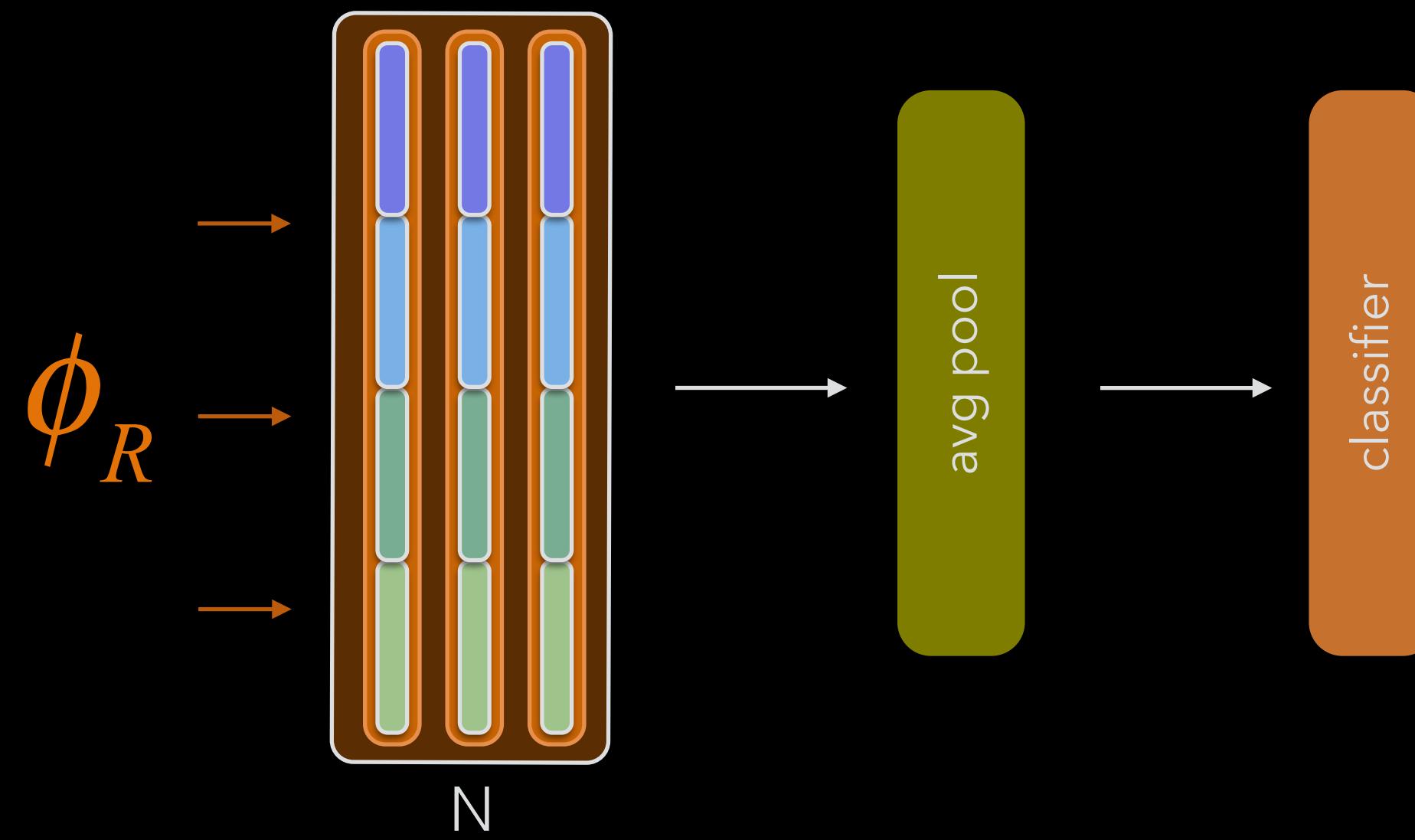


ϕ = clip representation

Interactions & Relationships



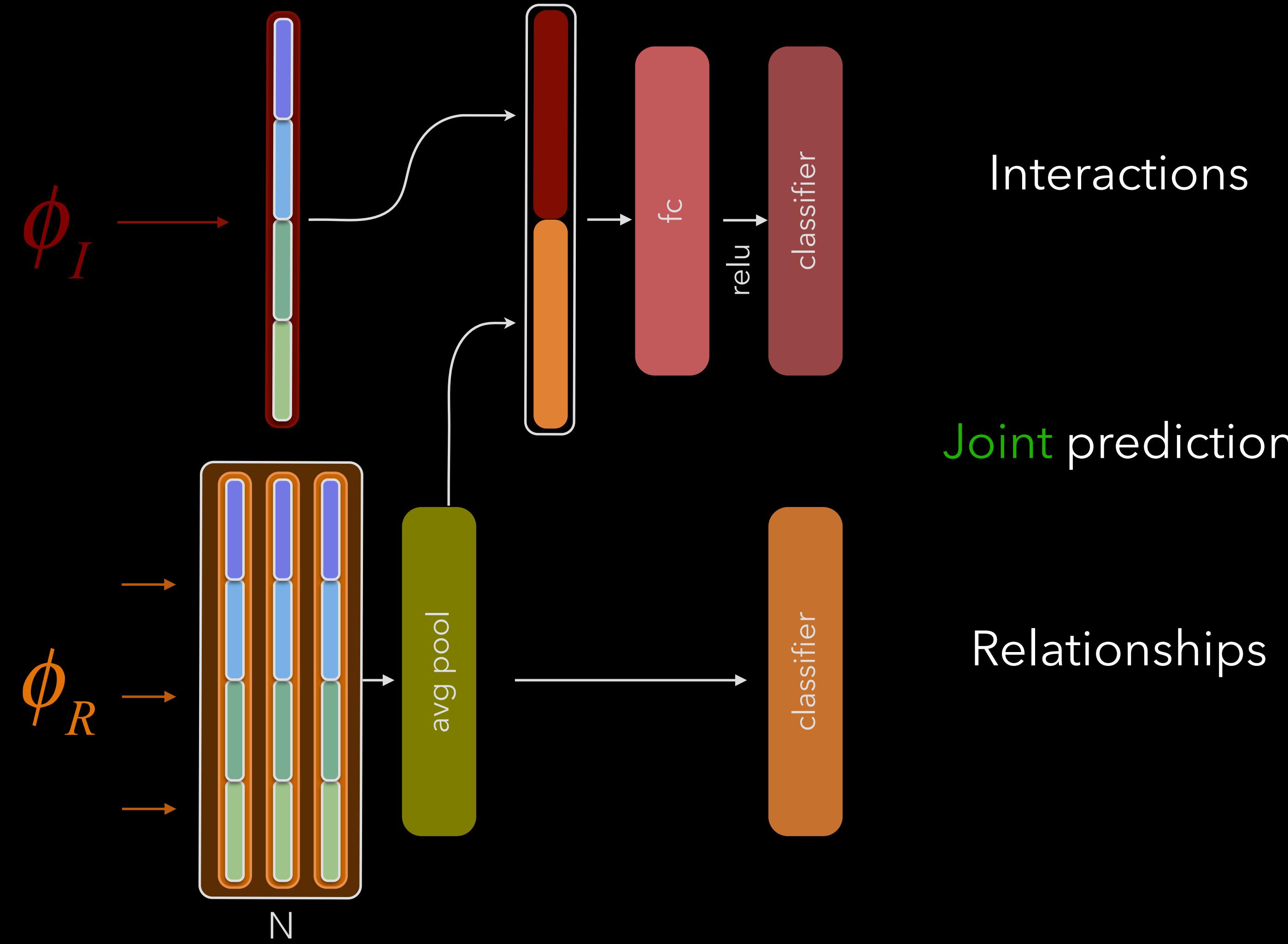
Interactions Network



Relationships Network

we average N clips as relationships last longer

Multitask problem



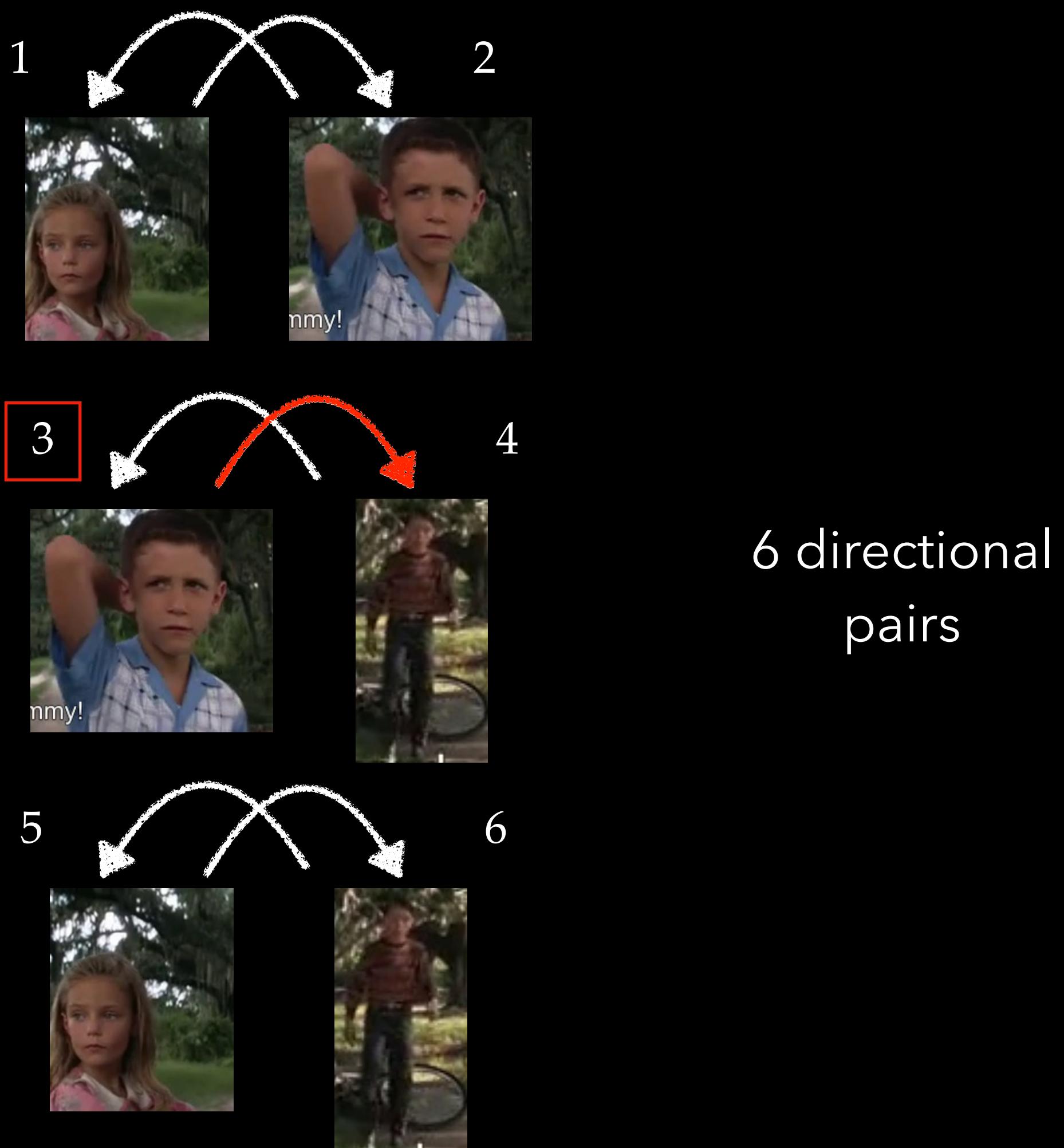
Character Pair Prediction



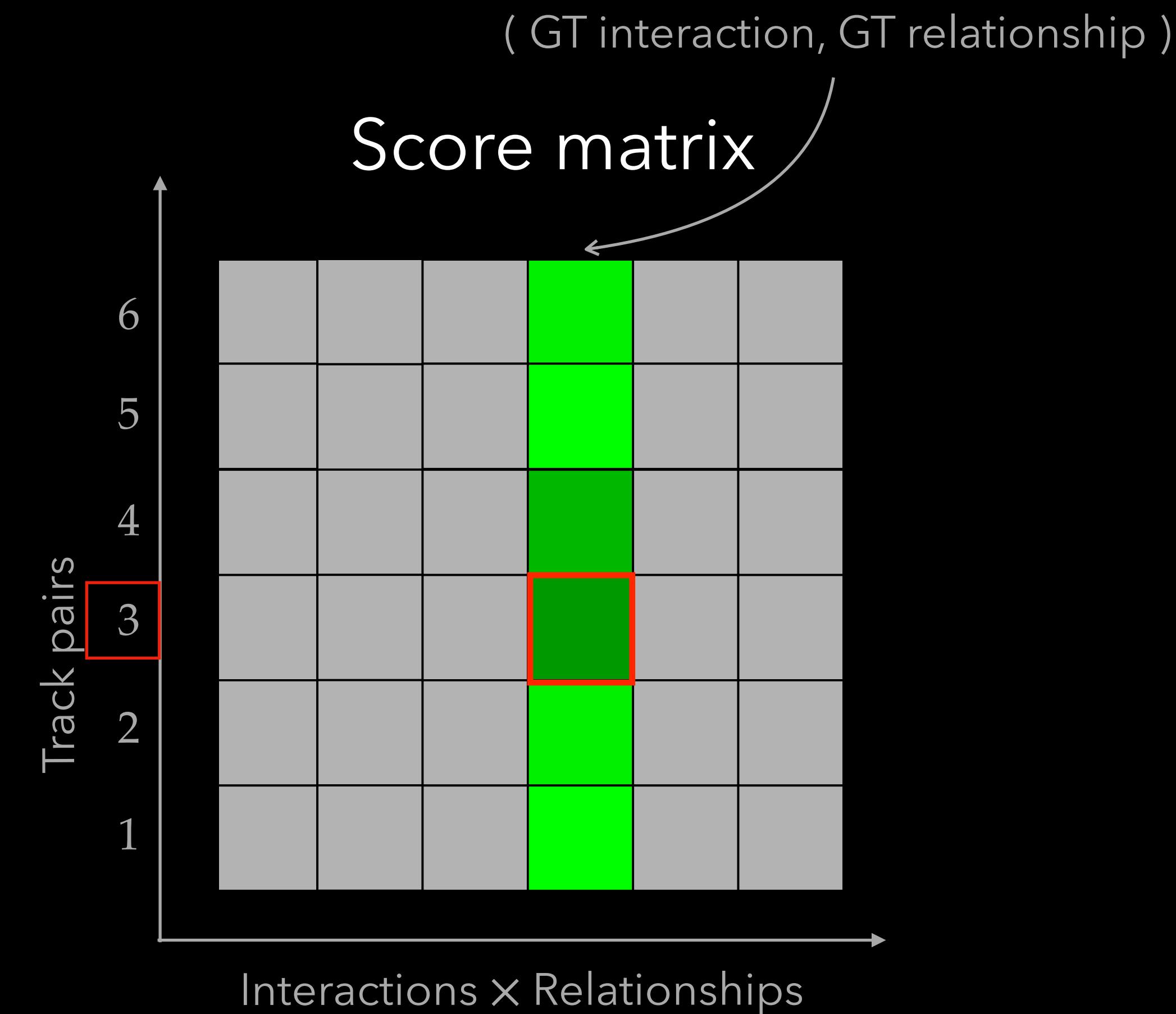
**enemies
run away**



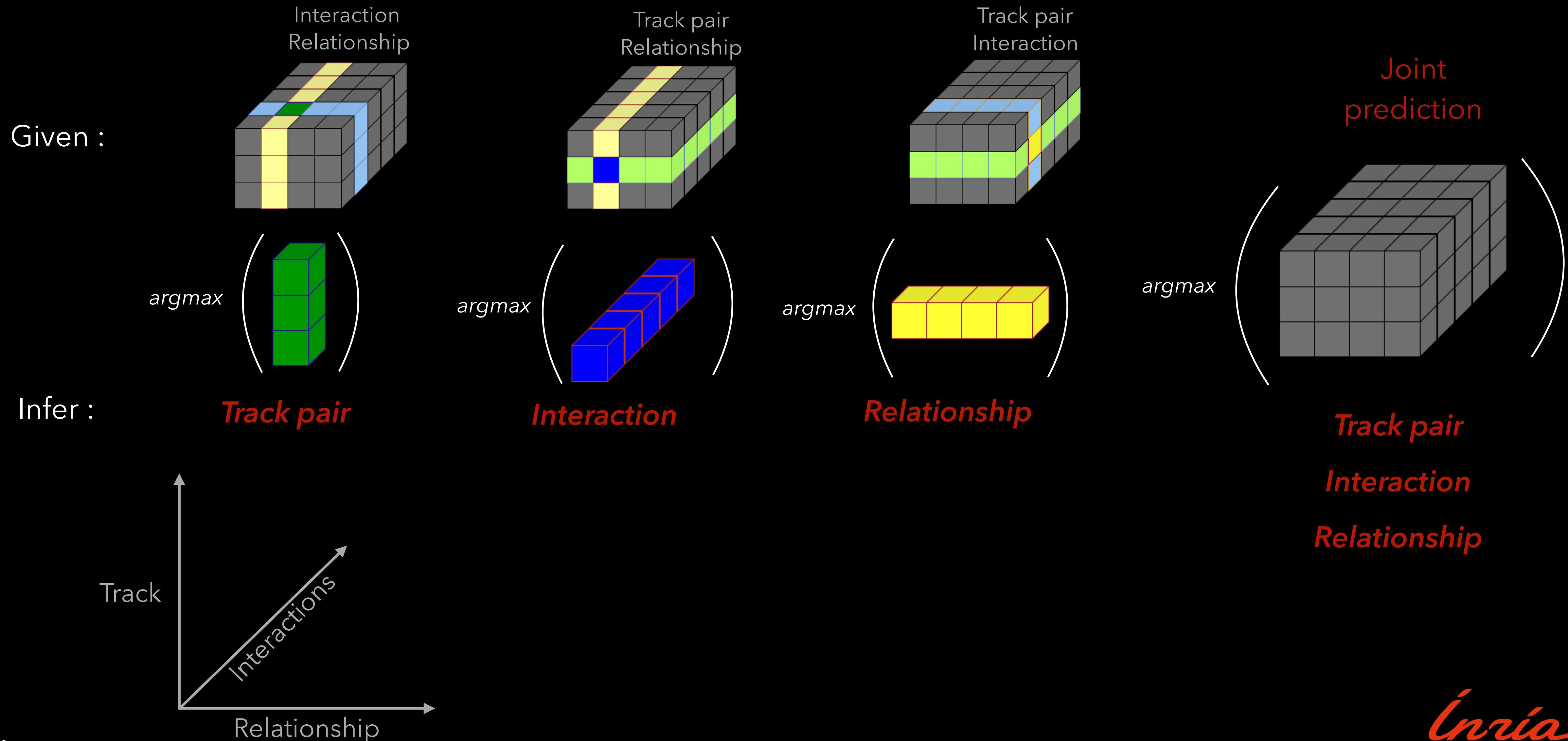
Character pair as latent variable



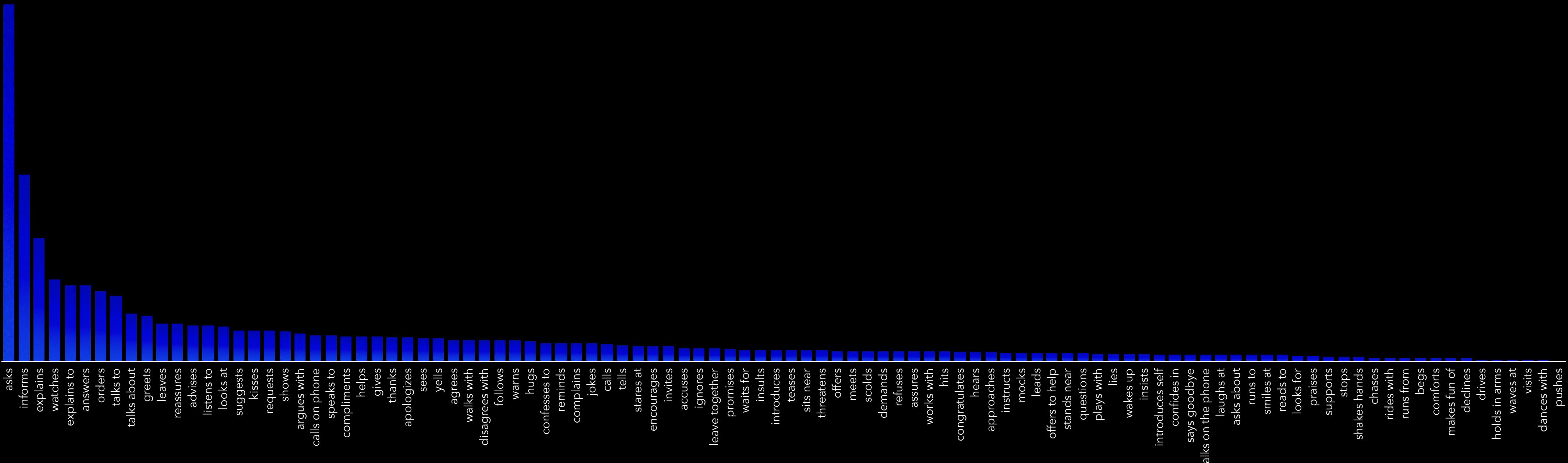
6 directional
pairs



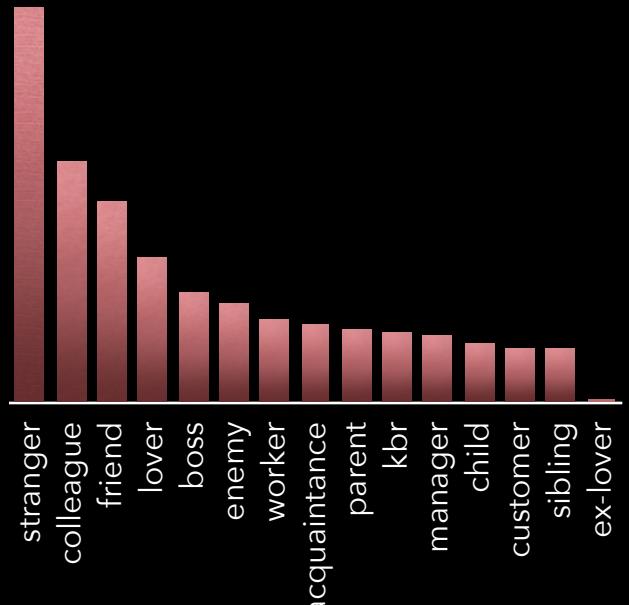
Inference



Interactions



Relationships



- 51 movies

- 101 interactions:

asks, greets, kisses

- 15 relationships:

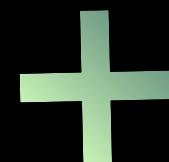
stranger, friend, lover

Modalities influence

Modalities			Interaction Acc	
Visual	Dialog	Tracks	Top-1	Top-5
✓			18.7	45.8
	✓		22.4	50.6
✓	✓		25.0	54.8
✓	✓	✓	26.1	57.3

Qualitative example: Modalities

Clip Visual



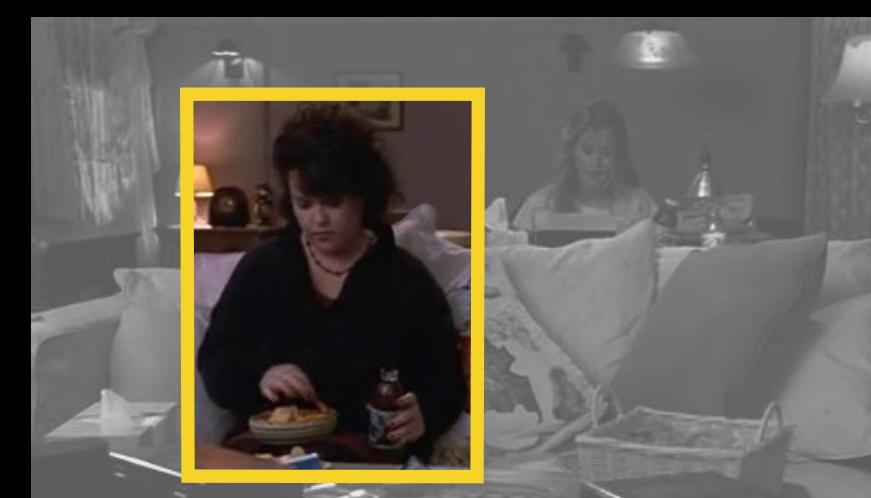
asks	0.77
informs	0.72
explains	0.69
suggests	0.65
compliments	0.63

Clip Textual

- «Dear Sleepless and Son, I have never written a letter like this in my life.»

- That's what everyone writes at the beginning of letters to strangers.

Tracks



explains	0.70
apologizes	0.67
reads	0.65
reassures	0.62
informs	0.62

reads	0.66
explains	0.62
sits	0.60
eats with	0.58
greets	0.58

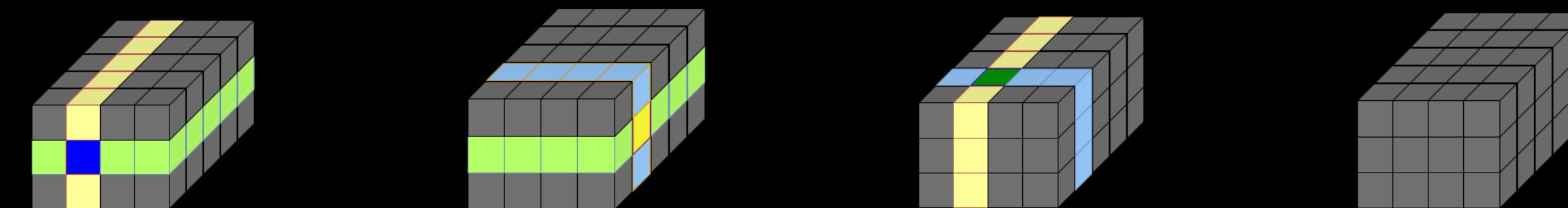
Joint prediction of Interactions & Relationships

Training \ Task	Interactions Acc	Relationships Acc
Random	0.99	6.7
only Interactions	26.1	-
only Relationships	-	26.8
Joint	26.3	28.1

Change of Acc with Joint training	Interaction	Associated Relationship	
↑	hugs	specific	lover, child
↑	introduces	specific	stranger
↑	runs	specific	enemy, lover
↓	talks to	shared	friend
↓	accuses		colleague
↓	greets		stranger

Change of Acc with Joint training	Relationship	Associated Interaction	
↑	sibling	specific	hugs
↑	acquaintance	specific	greets
↑	lover	specific	kisses
↓	boss	shared	asks
↓	manager		orders

Localizing Characters



Given	Relationship Character Pair	Interaction Character Pair	Interaction Relationship	-
Predict	Interactions	Relationships	Character Pair	Joint Acc
Random	0.99	6.67	15.42	0.01
Full	25.8	16.6	88.3	2.71
Weak *	25.8	12.0	42.0	0.86
Weak **	20.8	21.8	33.9	2.14

Weak *

Weak ** : negative mining for max margin loss with $\frac{\text{sum}}{\max}$ operator

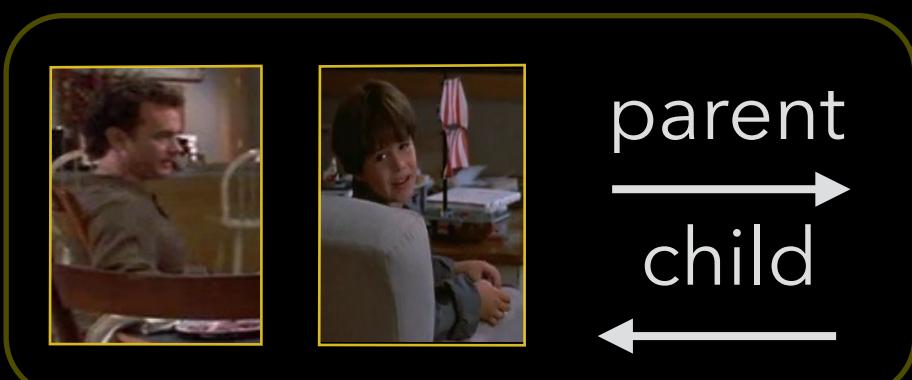
Qualitative example: Joint

Clip

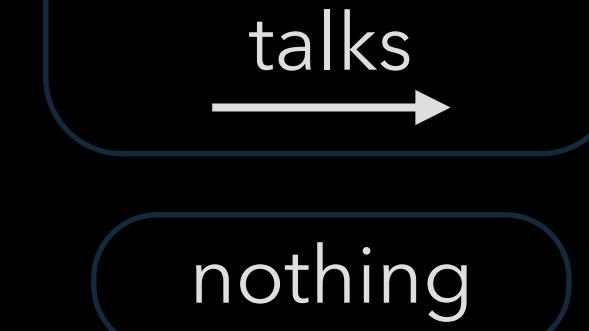
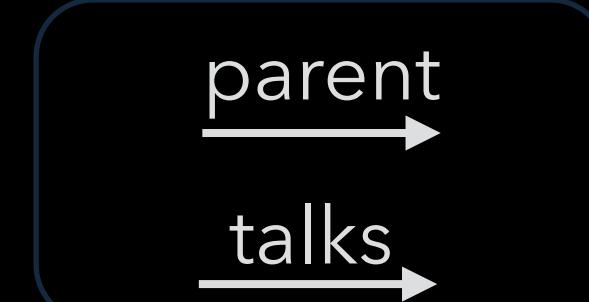


- You don't like Victoria?
- She laughs like a hyena.
- Is this true?
- Sort of, a little bit. Tell them what you did.

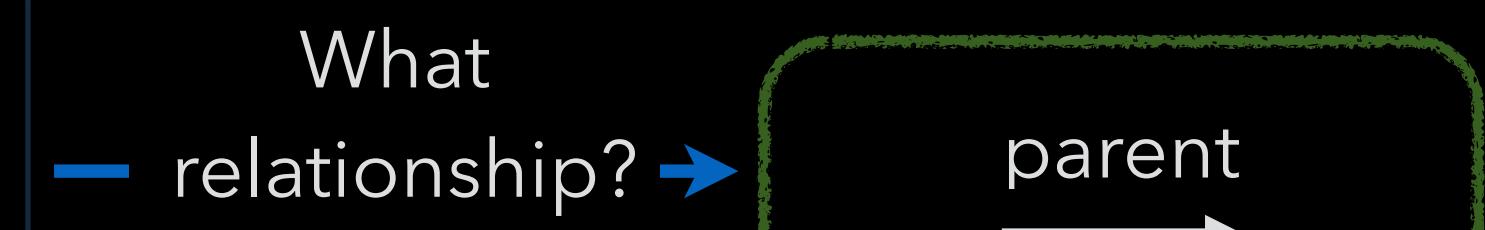
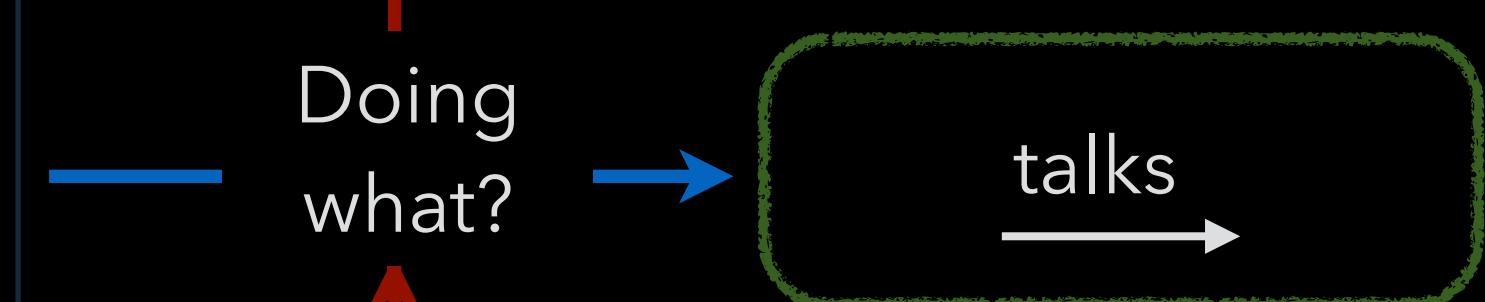
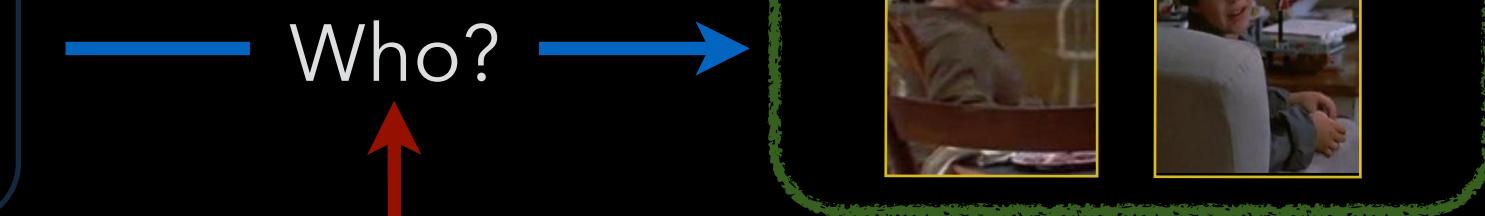
Possible tracks



Given



Prediction



We can do all together
Joint Prediction

Summary



- New task and model for studying social interactions and relationships
- Efficient multimodal clip representation
- Character pairs localization with clip-level interactions and relationships labels