

# Situation Recognition with Graph Neural Networks

R Li et al., ICCV 2017

1. Introduction
2. Graph Neural Network
3. Situation Recognition with GNN
4. Experimental Results
5. Conclusion



[Repairing]

Who

What

With

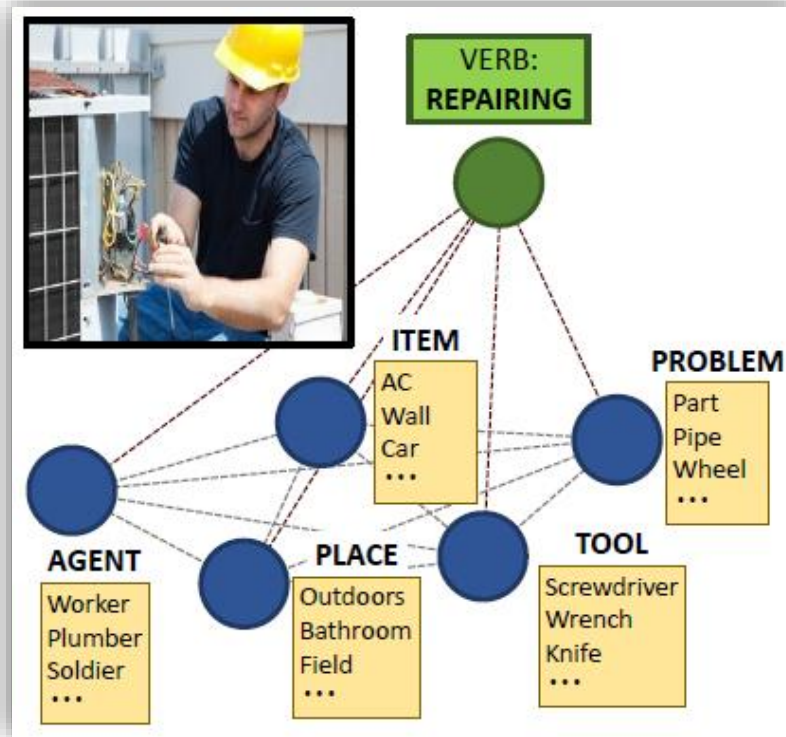
Associate images with verbs and role-noun pairs

- Describe the verb with more details

Visual semantic role labeling task

- Learn dependencies between verbs and nouns with the RNN

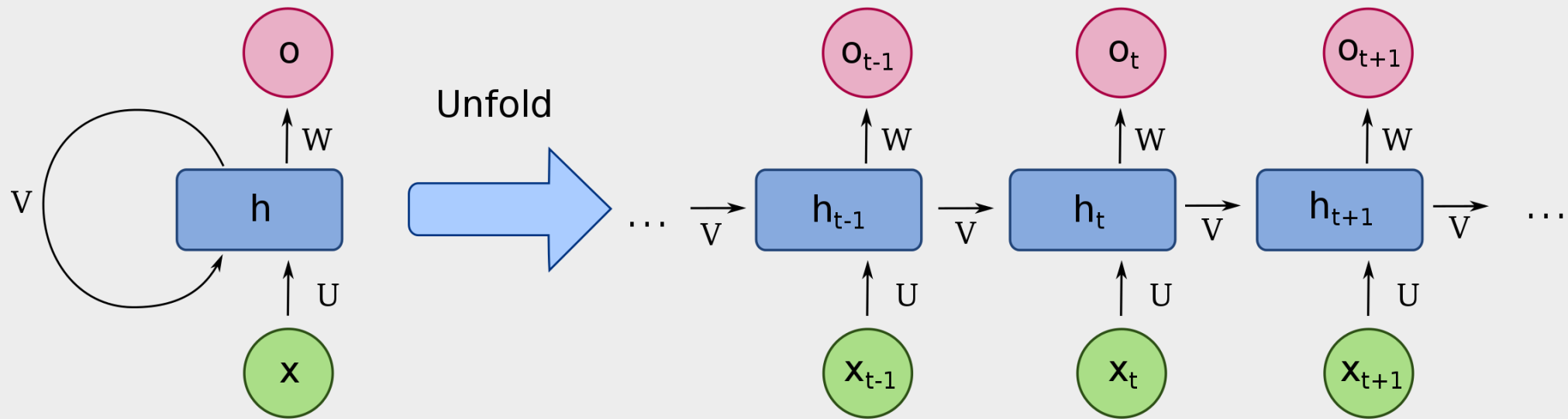
List all important objects with their attributes and relationships



Join verbs and their roles using Graph Neural Network (GNN)

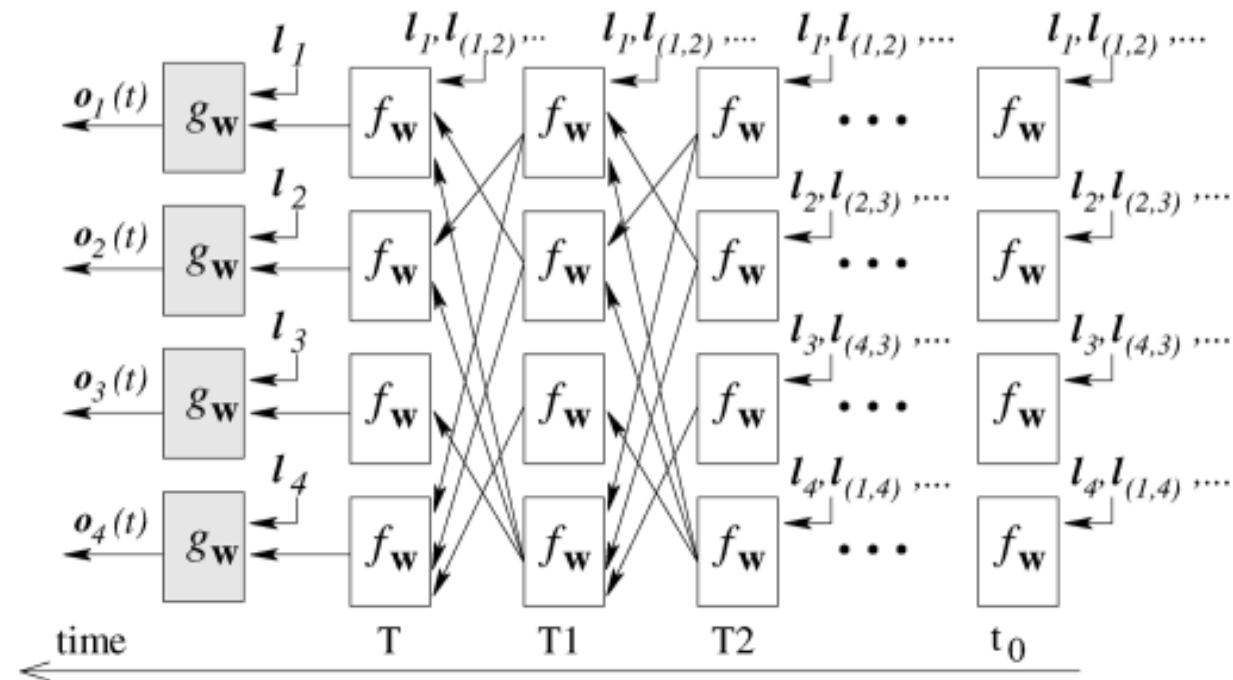
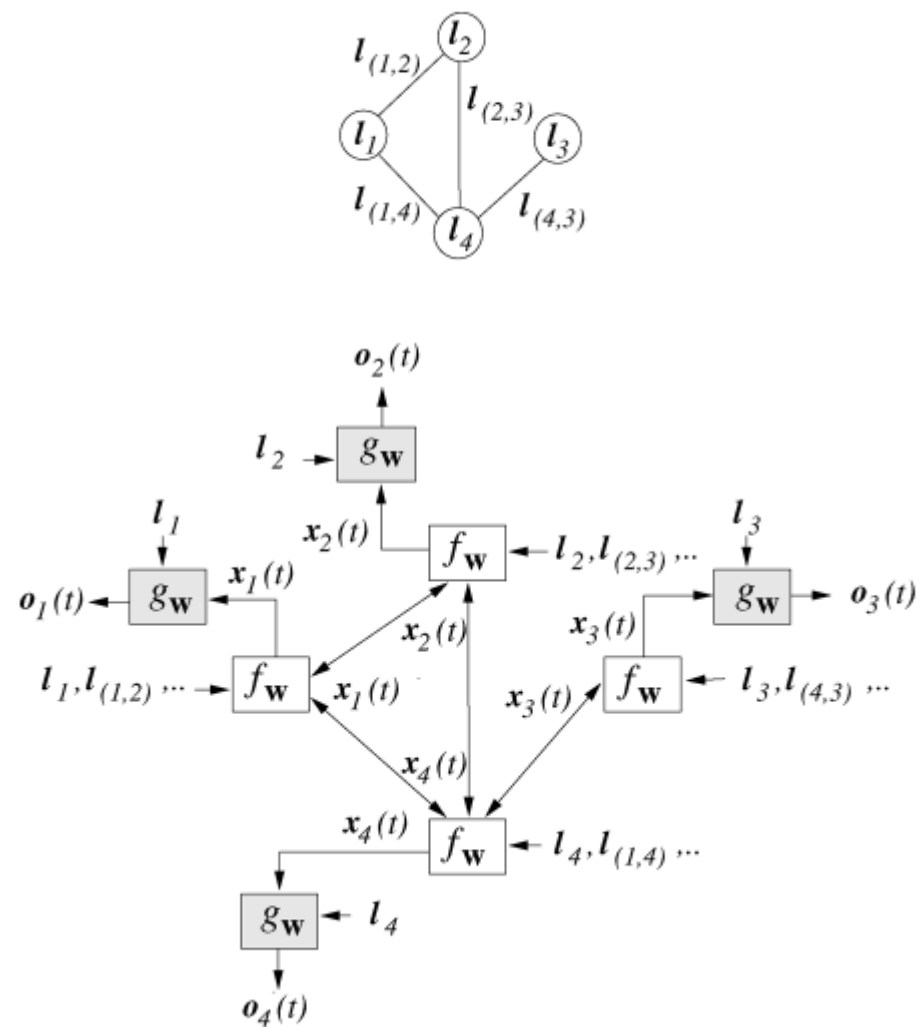
Exploit GNN to model dependencies between roles

Predict consistent structured outputs

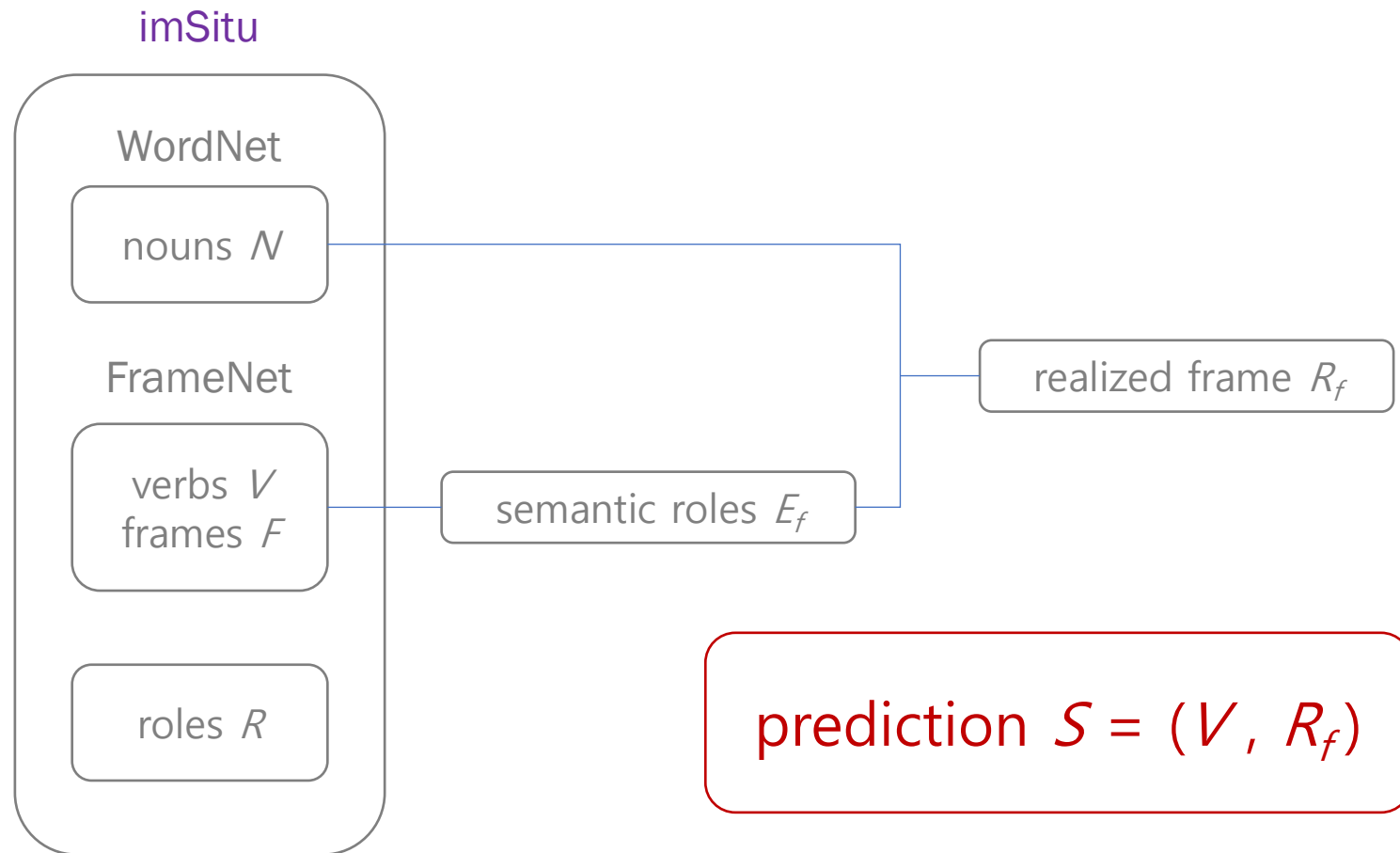


Prior recognition works used RNN network

=> Only moves information in one direction and updates one node per time step



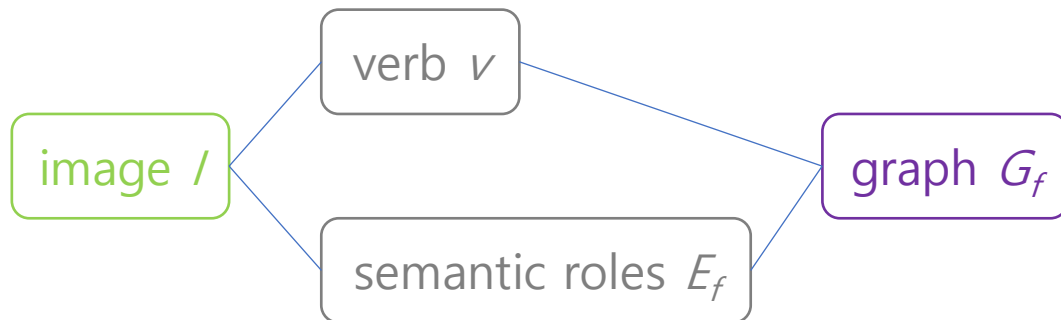
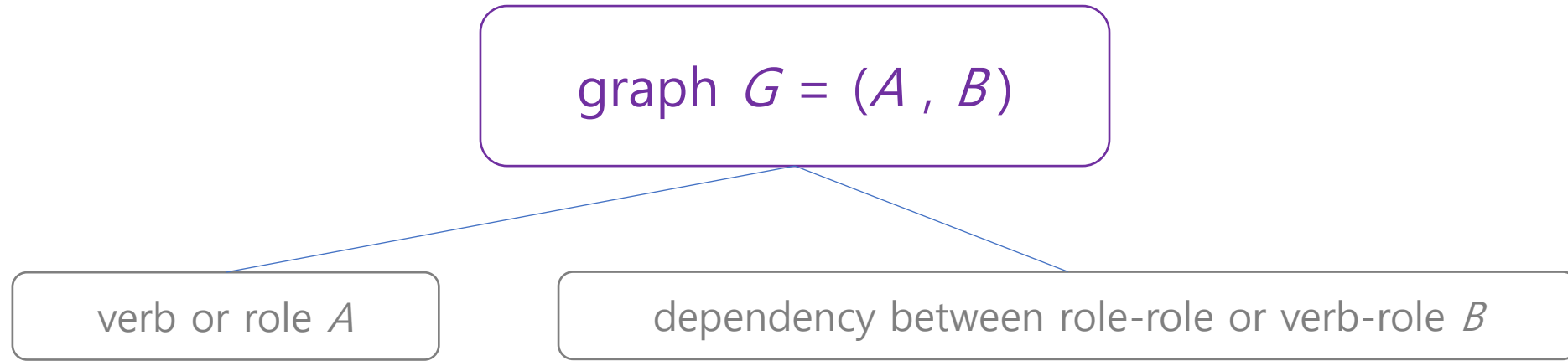
Propagate information by updating the hidden state of each node based on their history and incoming messages from their neighborhood



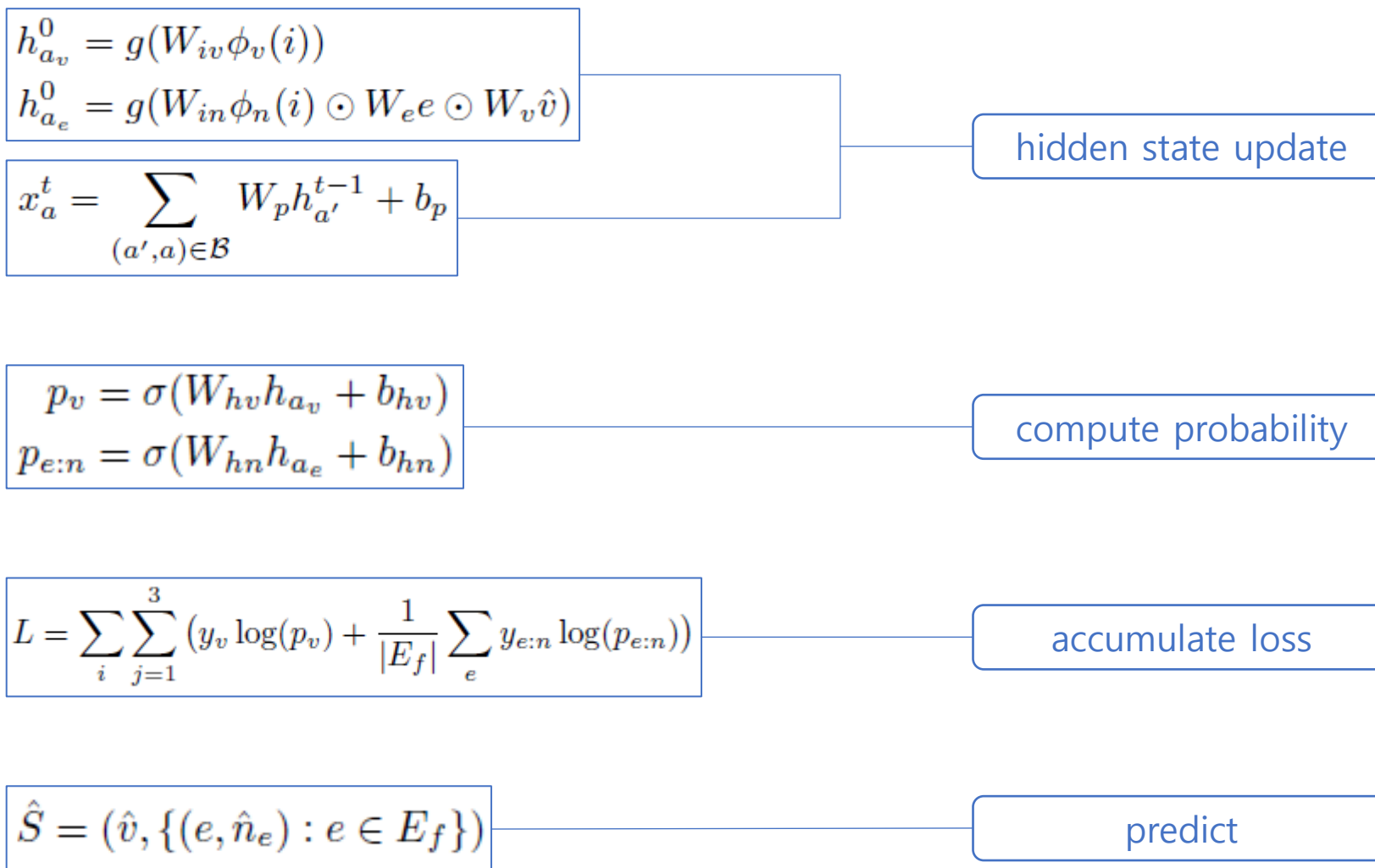






RIDING								
ROLE	AGENT	VEHICLE	PLACE		ROLE	AGENT	VEHICLE	PLACE
VALUE	MAN	HORSE	OUTSIDE		VALUE	DOG	SKATEBOARD	SIDEWALK




Create undirected edges between all pairs of roles















	GT: <b>FISHING</b>			
	AGENT	PLACE	TOOL	SOURCE
	MAN	BOAT	FISHING	LAKE
	PRED: <b>CATCHING</b>			
	AGENT	PLACE	TOOL	CAUGHTITEM
	MAN	BOAT	BODY	FISHING

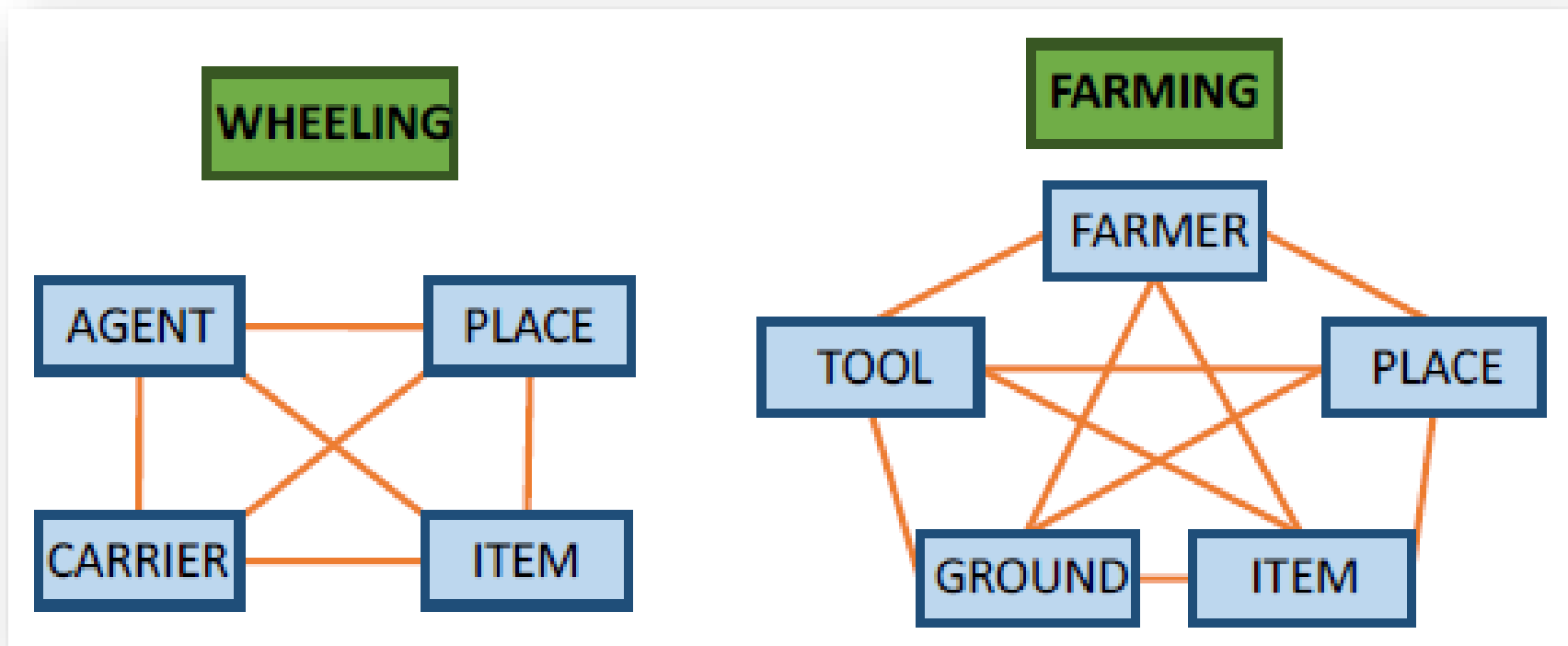
	GT: <b>SLOUCHING</b>		
	AGENT	PLACE	CONTACT
	WOMAN	OFFICE	CHAIR
	PRED: <b>SITTING</b>		
	AGENT	PLACE	CONTACT
	WOMAN	OFFICE	CHAIR

	GT: <b>SHELVING</b>			
	AGENT	PLACE	ITEM	DESTINATION
	WOMAN	LIBRARY	BOOK	BOOKSHELF
	PRED: <b>BROWSING</b>			
	AGENT	PLACE	GOALITEM	
	WOMAN	LIBRARY	BOOK	



											
CLINGING		DYEING		LEAKING		MILKING		DOUSING		DRUMMING	
AGENT	MONKEY	AGENT	PERSON	SUBSTANCE	WATER	AGENT	FARMER	AGENT	MAN	AGENT	MAN
PLACE	OUTDOORS	PLACE	OUTDOORS	PLACE	OUTSIDE	PLACE	OUTDOORS	PLACE	OUTDOORS	PLACE	ROOM
		MATERIAL	FABRIC	SOURCE	PIPE	TOOL	COW	LIQUID	WATER	TOOL	DRUMSTICK
CLUNGTO	MONKEY	DYE	RED	DESTINATION	LAND	SOURCE	COW	UNDERGOER	MAN	ITEM	DRUM
						DESTINATION	BUCKET				

											
HUGGING		PAWING		PICKING		TAXIING		OVERFLOWING		CAMPING	
AGENT	MAN	AGENT	CAT	AGENT	WOMAN	AGENT	AIRPLANE	AGENT	RUBBISH	AGENT	PEOPLE
PLACE	OUTDOORS	PLACE	OUTDOORS	PLACE	OUTDOORS	PLACE	AIRPORT	PLACE	OUTDOORS	PLACE	FOREST
HUGGED	MAN	AGENTPART	PAW	CROP	APPLE						
AGENTPART	ARM	PAWEDITEM	FENCE	SOURCE	TREE	GROUND	RUNWAY	SOURCE	ASHCAN	SHELTER	TENT



Predicting the correct verb along with its corresponding frame consisting of role-noun pairs

Capture dependencies between verb and roles, allowing nouns to inform each other, with GNN

Improve the accuracy compared with RNN network

Q & A