

Learning the Visual Interpretation of Sentences

Supplementary Material

C. Lawrence Zitnick
Microsoft Research
larryz@microsoft.com

Devi Parikh
Virginia Tech
parikh@vt.edu

Lucy Vanderwende
Microsoft Research
Lucy.Vanderwende@microsoft.com

1. Frequent Nouns

The 200 nouns found most frequently in our dataset along with their frequencies are listed below.

- Mike (26980)
- Jenny (26052)
- (18548)
- ball (4430)
- hat (4070)
- bear (3217)
- dog (3192)
- cat (2709)
- tree (2642)
- table (1825)
- owl (1817)
- balloon (1732)
- snake (1535)
- sandbox (1429)
- sunglass (1353)
- duck (1315)
- pie (1216)
- slide (1177)
- football (1177)
- hamburger (1146)
- pizza (1143)
- Frisbee (1006)
- ground (991)
- sun (909)
- play (885)
- baseball (822)
- hot dog (818)
- kite (794)
- grass (775)
- helicopter (751)
- cloud (721)
- it (690)
- cap (681)
- sky (652)
- park (637)
- and (633)
- beach ball (628)
- swing (627)
- fire (624)
- tent (616)
- rain (598)
- airplane (583)
- bat (559)
- she (549)
- hotdog (540)
- he (539)
- glasses (538)
- grill (513)
- jenny (506)
- hand (471)
- air (467)
- crown (457)
- rocket (426)
- soccer (400)
- basketball (384)
- plane (371)
- catch (361)
- soda (358)
- mustard (332)
- apple (331)
- burger (325)
- mike (319)
- campfire (281)
- ketchup (272)
- box (267)
- drink (260)
- glove (256)
- lightning (255)
- head (250)
- they (248)
- shovel (246)
- scare (242)
- food (234)
- be (231)
- day (230)
- arm (223)
- toy (215)
- racket (213)
- bucket (197)
- set (188)
- picnic (175)
- eat (162)
- pail (161)
- green snake (155)
- fun (145)
- kick (136)
- hotdogs (135)
- lunch (131)
- each other (131)
- kid (129)
- helmet (123)
- storm (119)
- sand (111)
- get (110)
- go (103)
- mitt (101)
- throw (101)
- star (99)
- shade (88)
- share (86)
- Kelly (83)
- you (81)
- rocket ship (78)
- shoe (78)
- hello (73)
- person (72)
- tennis (71)
- foot (67)
- swingset (67)
- bottle (66)
- beachball (66)
- cup (62)
- no one (60)
- fly (59)
- leg (58)
- goose (56)
- face (54)
- shirt (53)
- game (53)
- ride (52)
- horn (52)
- racquet (51)
- fetch (51)
- help (50)
- someone (50)
- hot-dog (49)
- joy (48)
- bee (48)
- lighten (47)
- bottom (47)
- see (46)
- man (46)

- thunderstorm (46)
- take (45)
- collar (45)
- ballon (45)
- friend (44)
- dress (44)
- come (44)
- barbecue (44)
- playground (44)
- pink (43)
- run (40)
- beanie (39)
- big tree (38)
- smile (37)
- have (37)
- water (37)
- hit (36)
- blue (36)
- bonfire (35)
- hold (35)
- baloons (35)
- something (35)
- pit (33)
- cook (33)
- give (32)
- Matt (32)
- X (32)
- what (32)
- camp (31)
- look (31)
- slider (31)
- top (30)
- balloon (29)
- wear (29)
- bench (29)
- disc (29)
- wave (28)
- cry (28)
- story (28)
- sad (27)
- want (27)
- bird (27)
- hole (26)
- sit (26)
- pale (25)
- nobody (25)
- mouth (24)
- watch (24)
- ballons (24)
- lighting (23)
- animal (23)
- tail (23)
- distance (23)
- rocketship (23)
- field (23)
- weather (23)
- child (22)
- lap (22)
- leave (22)
- cold drink (22)
- fall (22)
- fruit (22)
- bolt (21)
- make (20)
- stop (20)
- back (20)
- jump (20)
- soccer-ball (20)

2. Frequent Relations

The 300 relations found most frequently in our dataset along with their frequencies are listed below.

- be (8259)
- wear (6637)
- have (3776)
- hold (3670)
- sit (2669)
- kick (2292)
- play (2148)
- want (2026)
- be:pa:happy (1465)
- throw (1370)
- play:p:with (1057)
- see (1040)
- watch (1013)
- be:pa:sad (911)
- eat (838)
- sit:ptcl:on (832)
- wave (817)
- run (703)
- be:pa:angry (695)
- be:pa:mad (681)
- catch (642)
- try (611)
- fly (604)
- scare (582)
- stand:ptcl:by (577)
- be:pa:scared (537)
- stand:p:next to (528)
- sit:p:next to (522)
- jump (450)
- look (433)
- be:pa:afraid (425)
- to (419)
- like (412)
- with (395)
- fly:p:in (360)
- fly:p:over (353)
- have:ptcl:on (325)
- make (277)
- chase (255)
- play:ptcl:in (253)
- share (249)
- play:ptcl:on (242)
- be:pa:upset (226)
- smile (220)
- be:pa:surprised (220)
- cook (219)
- run:p:away from (216)
- fall (214)
- hit (207)
- get (203)
- stand (202)
- camp (200)
- rain (200)
- take (199)
- slide:p:down (198)
- give (191)
- bring (190)
- start (188)
- cry (187)
- drop (181)
- laugh (179)
- be:pa:excited (173)
- talk (170)
- be:pa:sunny (170)
- be:pa:worried (165)
- go:ptcl:down (157)
- stand:ptcl:on (150)
- come (146)
- sit:p:near (140)
- swing (135)
- sit:p:beside (135)
- run:p:towards (129)
- shine (129)
- look:p:at (128)
- scare:p:of (126)
- stand:p:near (122)
- go (122)
- stand:p:beside (120)
- stand:p:in front of (119)
- find (117)
- beJenny:s:be (112)
- of (112)
- offer (109)
- sit:p:with (108)

- be:pa:scaring (107)
- leave (107)
- think (106)
- tell (101)
- toss (101)
- run:p:to (100)
- put (99)
- fly:p:above (99)
- drink (99)
- love (97)
- wait (96)
- be:pa:ready (96)
- have:p:on (96)
- enjoy (94)
- surprise (94)
- call (93)
- fall:ptcl:off (93)
- run:p:with (92)
- sit:p:by (90)
- sit:p:in front of (88)
- run:p:toward (83)
- land (81)
- pet (80)
- lose (79)
- help (78)
- yell (76)
- say (75)
- hide:p:behind (73)
- steal (72)
- carry (70)
- float (70)
- grill (70)
- swing:p:on (69)
- be:pa:cloudy (68)
- be:pa:bear (68)
- between (68)
- hold:ptcl:up (67)
- stand:ptcl:in (65)
- show (62)
- walk (62)
- ride (62)
- be:pa:yellow (61)
- be:pa:red (59)
- pretend (56)
- be:pa:pink (55)
- follow (55)
- wave:p:to (55)
- near (54)
- build (54)
- hide (54)
- mike (52)
- stand:p:between (52)
- wave:p:at (52)
- slide (51)
- be:pa:shocked (51)
- hurt (50)
- be:pa:full (49)
- strike (49)
- shine:p:on (49)
- behind (48)
- be:pa:big (47)
- stand:p:behind (47)
- growl (47)
- go:p:to (46)
- next to (46)
- sit:p:on top of (45)
- know (44)
- fall:ptcl:on (44)
- for (43)
- jump:p:for (43)
- do (43)
- roar:p:at (43)
- growl:p:at (42)
- scare:ptcl:away (42)
- feel (41)
- notice (39)
- burn (39)
- sit:p:between (39)
- roar (39)
- cheer (38)
- grow:ptcl:on (38)
- sit:p:behind (38)
- stare (37)
- sit:p:in (37)
- be:pa:unhappy (37)
- let go:p:of (37)
- fly:ptcl:in (37)
- rest (37)
- raise (36)
- jump:p:into (36)
- fly:p:into (36)
- be:pa:high (36)
- climb (36)
- in (36)
- be:pa:empty (35)
- be:pa:hungry (35)
- bring:p:to (35)
- wait:p:for (34)
- cover (34)
- be:loc:sandboxJenny:s:be (34)
- miss (34)
- be:pa:orange (33)
- come:p:out of (33)
- shine:p:in (33)
- be:pa:brown (33)
- be:pa:hot-air (33)
- be:pa:hot (32)
- lighten (32)
- be:loc:bottombottom:p:of (32)
- use (31)
- let (31)
- run:ptcl:after (31)
- shine:p:above (31)
- hold:ptcl:out (31)
- have:ptcl:up (30)
- be:pa:Verb (30)
- begin (30)
- run:p:from (30)
- away from (29)
- beMike:s:be (29)
- sit:p:around (29)
- sniff (29)
- reach:p:for (29)
- keep (29)
- ruin (29)
- play:p:near (29)
- waive (29)
- be:pa:single (29)
- be:loc:toptop:p:of (29)
- jump:p:off (28)
- on (28)
- look:ptcl:for (28)
- ask (28)
- be:pa:shock (28)
- stick (27)
- slither (27)
- stop (27)
- warm (27)
- come:ptcl:to (27)
- fall:p:off of (27)
- set:ptcl:up (26)
- fly:p:behind (26)
- float:p:above (26)
- shin (26)
- rain:p:on (26)
- dance (25)
- fight (24)

- be:pa:friend (24)
- go:p:into (24)
- put:ptcl:out (24)
- reach (24)
- bake (24)
- be:pa:scary (23)
- fly:p:towards (23)
- be:pa:pirate (23)
- look:p:like (23)
- fight:p:over (22)
- above (22)
- betree:p:behind (22)
- be:pa:bright (22)
- argue (22)
- scream (22)
- pitch (22)
- worry:p:about (22)
- float:p:in (22)
- join (22)
- fall:p:from (22)
- towards (22)
- call:p:for (21)
- in front of (21)
- move (21)
- be:pa:small (21)
- perch (21)
- wear:p:on (21)
- be:loc:parkJenny:s:be (21)
- startle (20)
- be:pa:funny (20)
- put:p:on (20)
- bite (20)
- barbecue (20)
- at (20)
- laugh:p:at (20)
- smile:p:at (20)
- be:pa:move (19)
- be:pa:nearby (19)
- be:pa:blue (19)
- slither:p:towards (19)
- kick:p:into (18)
- sneak (18)
- wish (18)
- chase:p:after (18)
- pick:ptcl:up (18)
- have:p:in (18)
- prepare (18)
- be:pa:green (18)
- hope (17)
- sit:p:on (17)
- be:loc:tablemustard:s:be (17)
- lightning (17)
- stay (17)
- hate (17)
- wonder (17)
- feed (17)
- pass (16)
- walk:p:towards (16)
- frighten (16)
- catch:p:with (16)
- come:ptcl:down (16)
- put:ptcl:up (16)
- stare:p:at (15)
- grab (15)
- attack (15)
- share:p:with (15)
- hand (15)
- fly:p:away from (15)
- go:p:behind (15)

3. Qualitative Results

Fifty random example scenes generated by our approach are shown below. These are in addition to the three shown in Figure 5 in the main paper. The input description along with the tuples extracted are provided in the caption. The five images shown are (from left to right): the ground truth scene (GT) from our dataset corresponding to the input description, the scene *generated* by our approach (Full-CRF), the scene retrieved by a bag-of-words baseline (BoW), the scene generated by a CRF model whose potentials are updated with only the nouns (Noun-CRF), and a random scene from the dataset (Random).

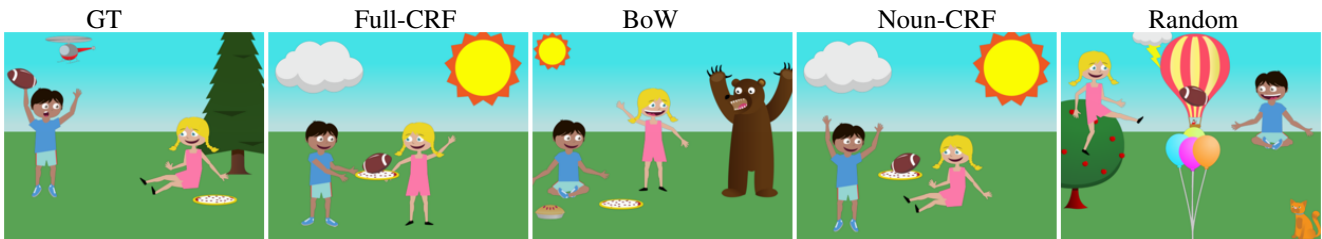


Figure 1: **Input description:** There is a pizza next to Jenny. Jenny is waving to Mike. Mike has a football. **Tuples:** pizza be:pizza:p:next to Jenny; Jenny wave ; Mike have football;



Figure 2: **Input description:** Mike tripped over the cat and fell. The cat was by the table. Jenny ran to help Mike up after he fell. **Tuples:** Mike trip:p:over cat; Mike fall ; cat be table; Jenny run ; Jenny help ; he fall ;

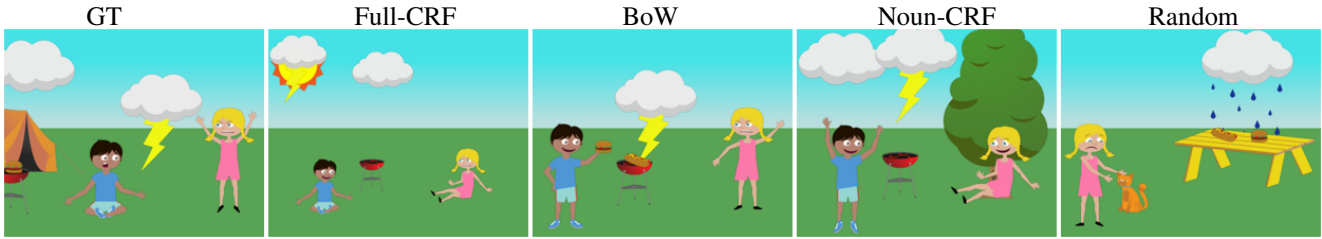


Figure 3: **Input description:** Jenny is mad at the lightning The lightning strikes Mike Mike sits beside the grill **Tuples:** Jenny be:pa:mad ; lightning strike Mike; Mike sit:p:beside grill;

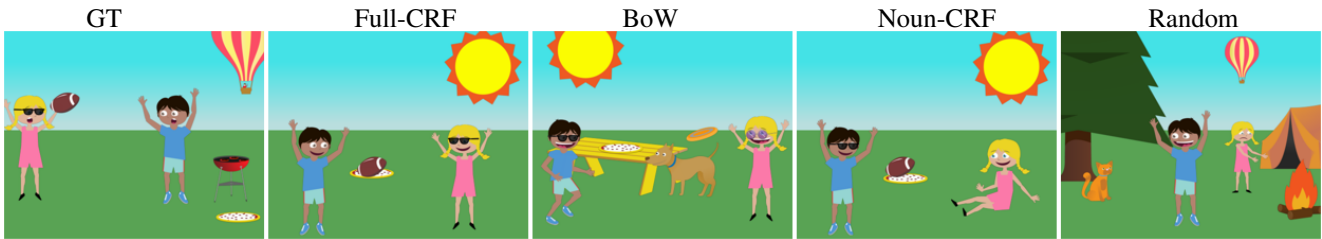


Figure 4: **Input description:** There's a pizza on the ground Jenny catches the football jenny is wearing sunglasses **Tuples:** pizza be ground; Jenny catch football; jenny wear sunglass;

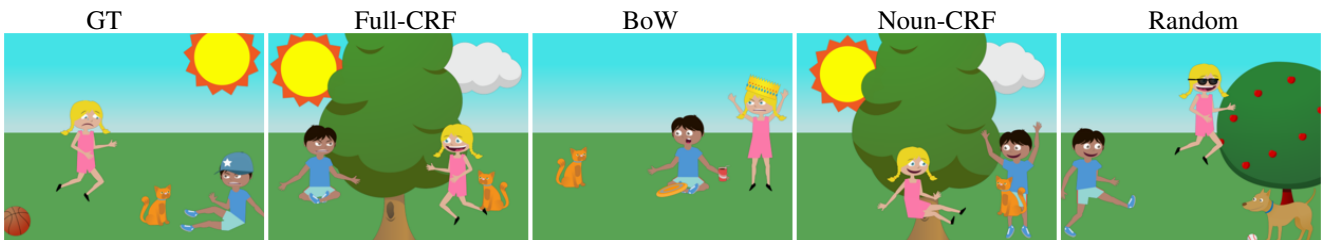


Figure 5: **Input description:** The cat tripped Mike. Mike is angry at the cat. Jenny is running towards Mike. **Tuples:** cat trip Mike; Mike be:pa:angry ; Jenny run:p:towards Mike;



Figure 6: **Input description:** The owl is in the tree. Jenny is angry because it is lightning. Mike was trying to cook hamburgers. **Tuples:** owl be tree; Jenny be:pa:angry ; Mike try cook; Mike cook hamburger;

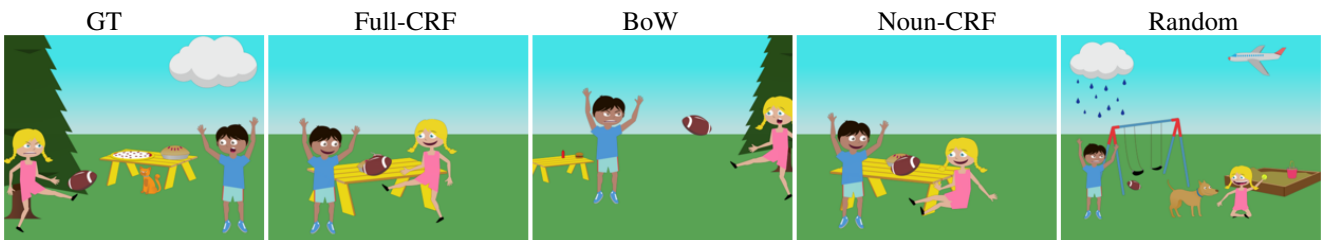


Figure 7: **Input description:** Jenny kicks the football Mike holds up his hands There is pie on the table **Tuples:** Jenny kick football; Mike hold:ptcl:up hand; pie be table;

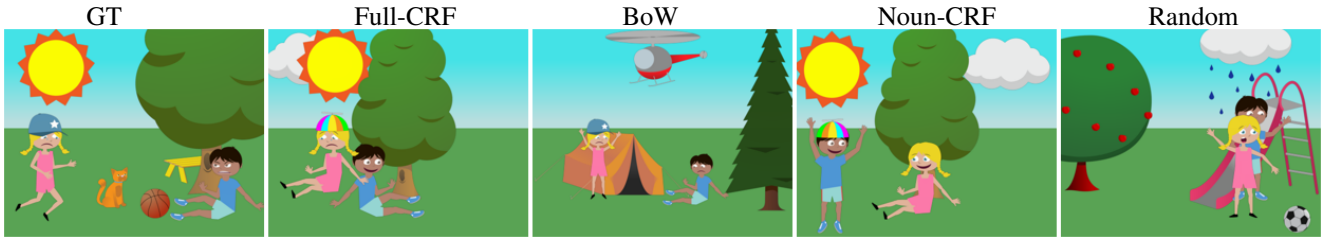


Figure 8: **Input description:** Jenny and Mike are sad. Mike sits on the ground. Jenny is wearing a hat. **Tuples:** and be:pa:sad ; Mike sit ground; Jenny wear hat;



Figure 9: **Input description:** Jenny is going to slap Mike. Mike is wearing a pirate hat. There is an owl above Mike and Jenny. **Tuples:** Jenny slap Mike; Mike wear hat; owl beowl:p:above Mike; owl above Jenny;

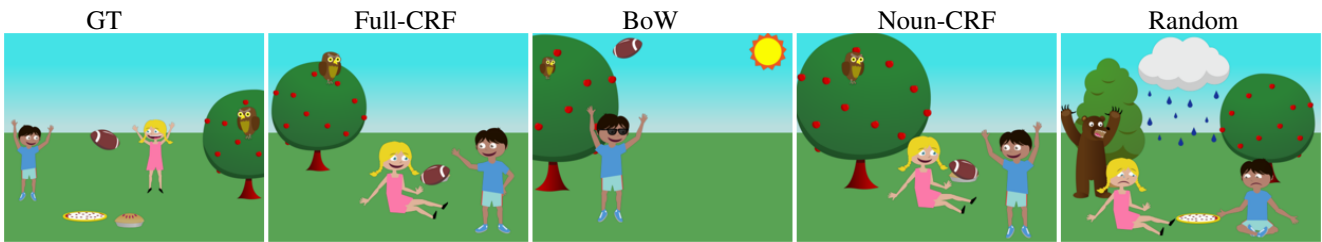


Figure 10: **Input description:** The owl sits in the tree Mike throws the football to Jenny The pie is beside the pizza **Tuples:** owl sit tree; Mike throw football; pie be ;

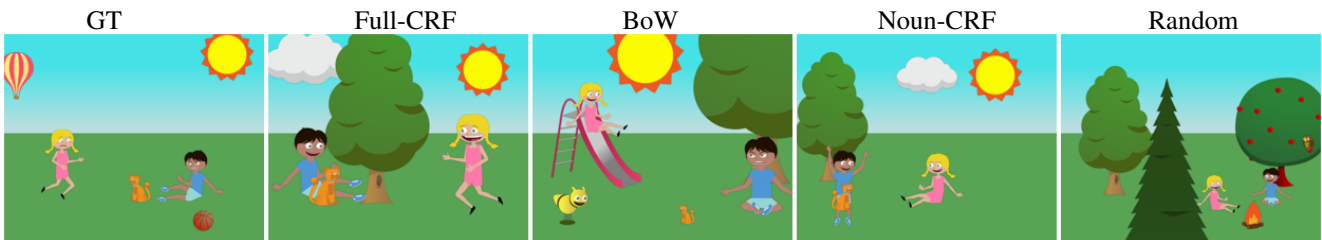


Figure 11: **Input description:** The cats sitting in front of Mike. Mike is sitting on the ground. Jenny is running toward Mike. **Tuples:** cat sit:p:in front of Mike; Mike sit ground; Jenny run:p:toward Mike;



Figure 12: **Input description:** The lightning cloud is between Mike and Jenny Jenny is wearing a crown Mike is sitting next to a burger and drink **Tuples:** cloud be ; Jenny wear crown; Mike sit:p:next to burger; Mike sit:p:next to drink;

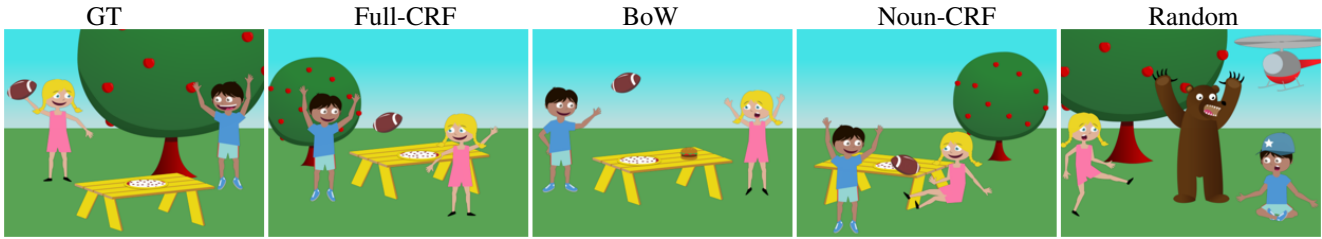


Figure 13: **Input description:** A pizza is on the table. Jenny is throwing a football. Mike is under the apple tree. **Tuples:** pizza be table; Jenny throw football; Mike be tree;



Figure 14: **Input description:** Mike is holding a basketball. Jenny is running towards Mike. The cat is sitting by Mike. **Tuples:** Mike hold basketball; Jenny run:p:towards Mike; cat sit:p:by Mike;



Figure 15: **Input description:** Jenny wears a funny hat and glasses. A lightning strike between Mike and Jenny. Mike was surprised by a lightning. **Tuples:** Jenny wear hat; Jenny wear glasses; lightning strike between Mike; lightning strike between Jenny; lightning surprise Mike;

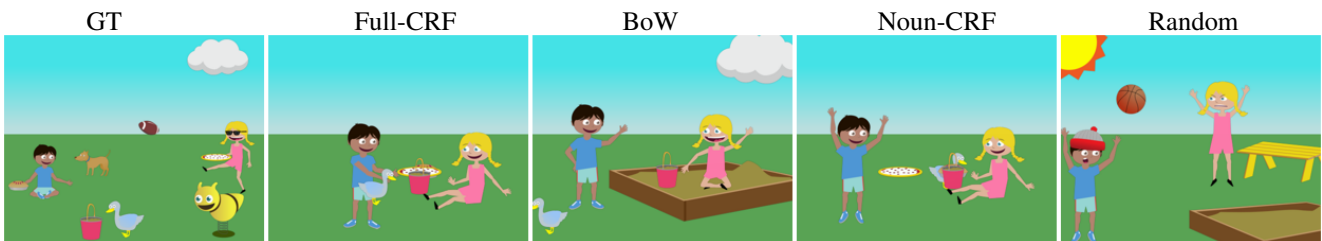


Figure 16: **Input description:** Mike has a pie Jenny is caring a pizza The duck is drinking out of a bucket. **Tuples:** Mike have pie; Jenny care pizza; duck drink:p:out of bucket;



Figure 17: **Input description:** Mike is mad at Jenny. Mike is sitting next to a cat. Jenny looks very upset. The cat is sitting by Mike. Mike is sitting by the basketball. The cloud is moving in front of the sun. **Tuples:** Mike be:pa:mad ; Mike sit:p:next to cat; cat sit:p:by Mike; Mike sit ; cloud move:p:in front of sun;

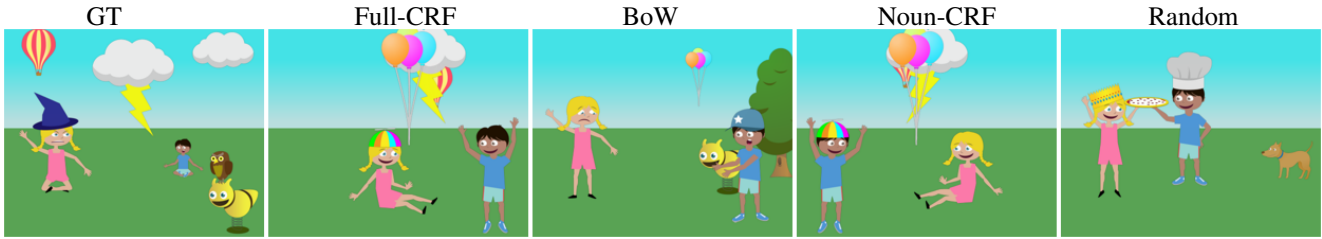


Figure 18: **Input description:** Jenny has a wizard hat on. A balloon is in the sky. Lightning is in the air. **Tuples:** Jenny have:ptcl:on hat; balloon be ; lightning be air;

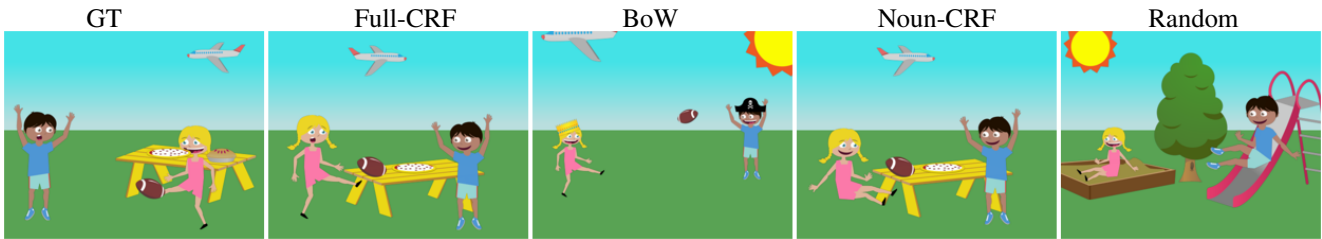


Figure 19: **Input description:** Jenny is kicking the football. The pizza is on the table. The airplane is flying over Jenny. **Tuples:** Jenny kick football; pizza be table; airplane fly:p:over Jenny;

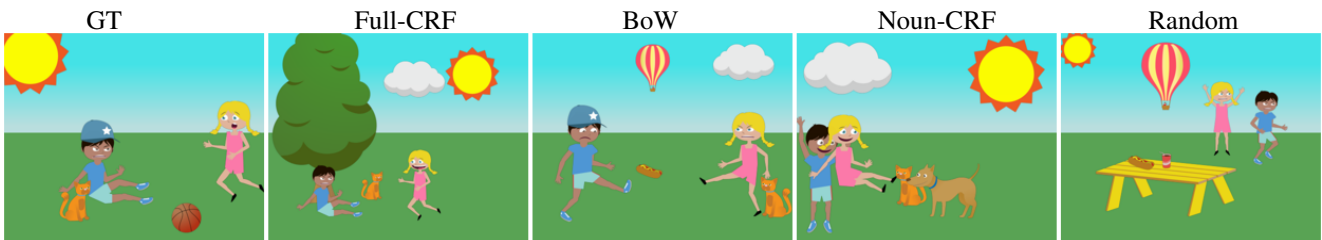


Figure 20: **Input description:** Mike is sitting next to a cat. Mike is angry because he fell down. Jenny is running towards Mike to help him. **Tuples:** Mike sit:p:next to cat; Mike be:pa:angry ; he fall ; Jenny run:p:towards Mike; Jenny help ;



Figure 21: **Input description:** It is lighting out. Mike and Jenny are upset. Mike and Jenny are sitting on the ground with there legs crossed. **Tuples:** it light ; Mike sit ground; Jenny sit ground; ground with leg;

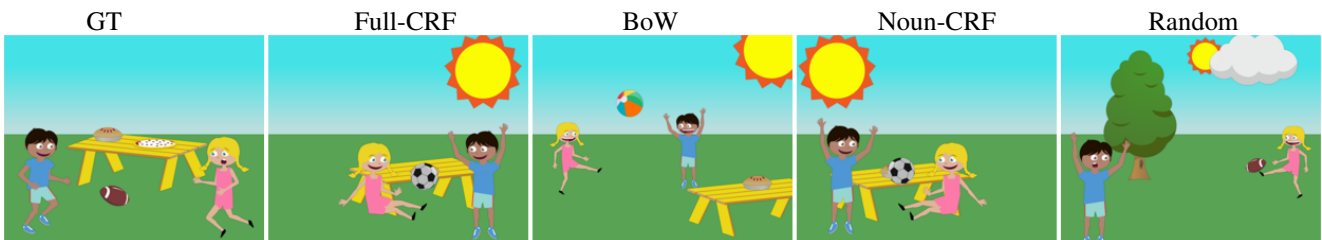


Figure 22: **Input description:** Jenny is throwingthe ball. Mike is catching the ball. The pie is on the table. **Tuples:** Jenny be:pa:ball ; Mike catch ball; pie be table;



Figure 23: **Input description:** Mike fell on the ground. Jenny is worried about Mike. Jenny is running to check on Mike. **Tuples:** Mike fall ground; Jenny be:pa:worried ; Jenny run ; Jenny check Mike;

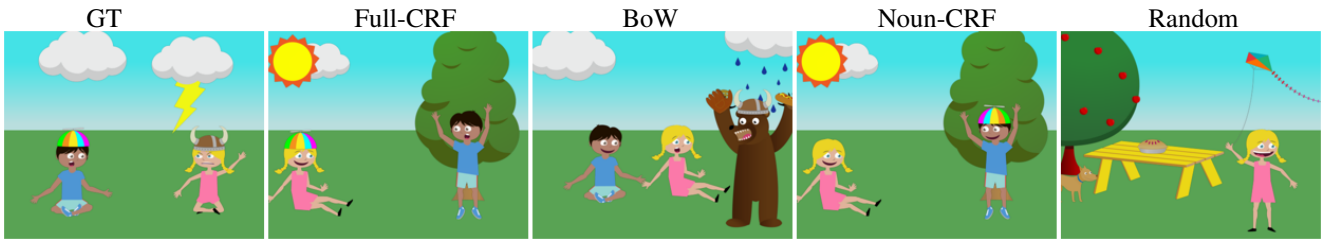


Figure 24: **Input description:** Jenny is wearing a horn hat. Mike is wearing a colored hat. Mike is scared of the lightning. **Tuples:** Jenny wear hat; Mike wear hat; Mike be:pa:scared ;



Figure 25: **Input description:** Jenny is holding a pizza. Mike is holding a football and pie. A tree is between Mike and Jenny. **Tuples:** Jenny hold pizza; Mike hold football; Mike hold pie; tree be ;



Figure 26: **Input description:** Jenny is sad Mike is angry the cat is sitting next to mike **Tuples:** Jenny be:pa:sad ; Mike be:pa:angry ; cat sit:p:next to mike;

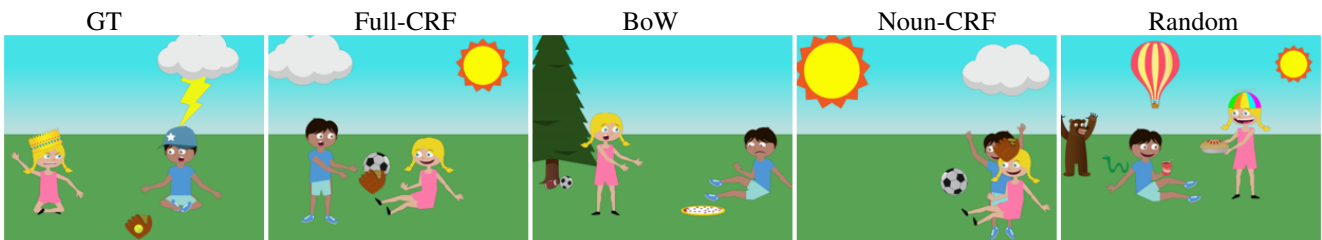


Figure 27: **Input description:** Mike is scared of lightning. Mike dropped his baseball mitt. There is a tennis ball in the mitt. **Tuples:** Mike be:pa:scared ; Mike drop mitt; ball be mitt;

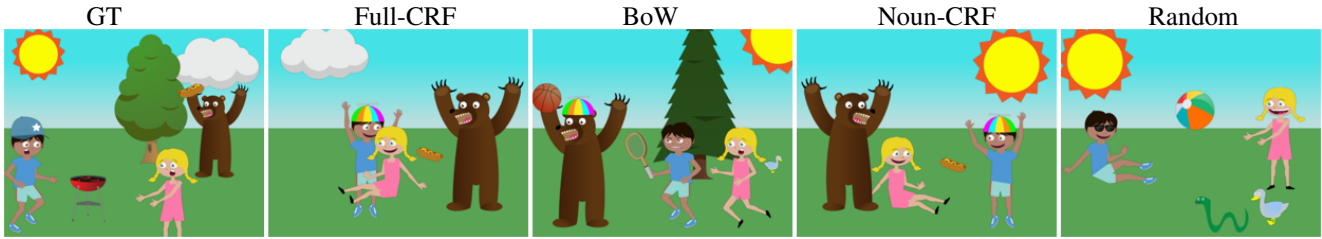


Figure 28: **Input description:** Mike is wearing the hat. The bear is holding the hot dog. Mike is next to the pit. **Tuples:** Mike wear hat; bear hold hot dog; Mike be ;



Figure 29: **Input description:** Mike and Jenny are sitting around the fire. There is a pie on the picnic table. The fire is orange color. Jenny is sitting by the table. The pie is on the table. The table is in front of a tree. **Tuples:** Mike sit:p:around fire; Jenny sit:p:around fire; pie be table; fire be:pa:color ; Jenny sit ; pie be table; table be ;



Figure 30: **Input description:** Mike is sitting on the ground. Jenny wants Mike to come and play with her. A dig is watching Mike and Jenny. **Tuples:** Mike sit ground; Jenny want come; Jenny want play; Mike come ; Mike play:p:with she; dig watch Mike; dig watch Jenny;

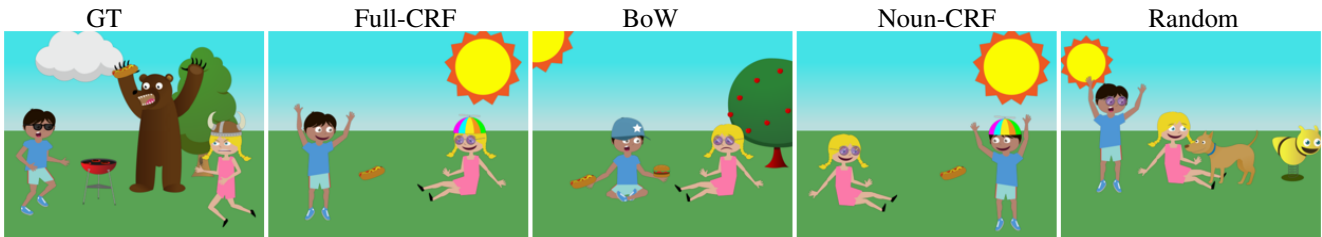


Figure 31: **Input description:** Mike is wearing sun glasses. Mike and Jenny's bear friend eats a hot dog. Jenny is wearing a hat. **Tuples:** Mike wear glasses; Mike bear hot dog; Jenny bear hot dog; Jenny wear hat;

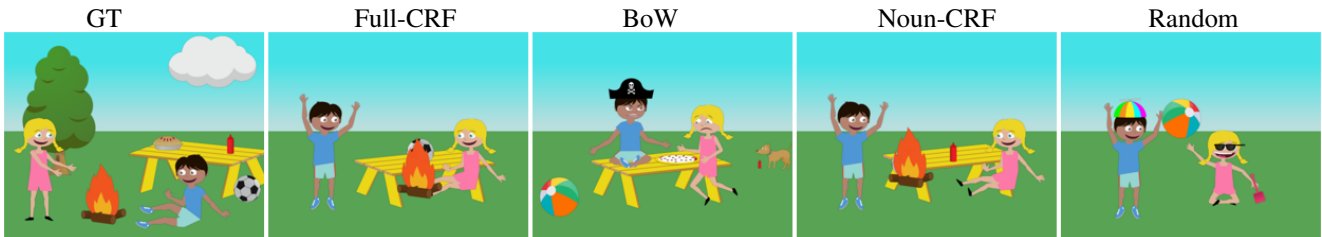


Figure 32: **Input description:** Jenny and Mike are by a campfire. A soccer ball is under the table. The ketchup and pie are on the table. **Tuples:** Jenny be:loc:campfireMike:s:be campfire; ball be table; ketchup be:loc:tablepie:s:be table;



Figure 33: **Input description:** Mike is wearing a pirate hat. The cat is sitting on the slide. Jenny is sitting under the tree. **Tuples:** Mike wear hat; cat sit:ptcl:on slide; Jenny sit tree;

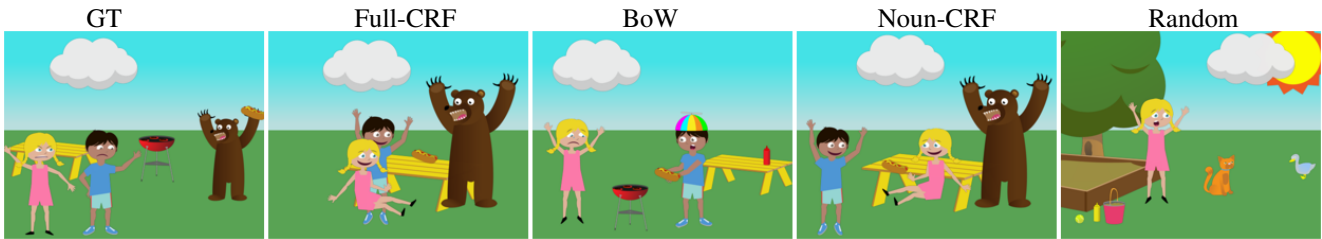


Figure 34: **Input description:** The bear has a hotdog. There is one cloud in the sky. There is nothing on the table. **Tuples:** bear have hotdog; cloud be ; nothing be table;



Figure 35: **Input description:** There is a pie on the picnic table. Mike and Jenny are sitting on the ground. There is one white fluffy cloud in the sky. **Tuples:** pie be table; Mike sit ground; Jenny sit ground; cloud be:pa:fluffy ;

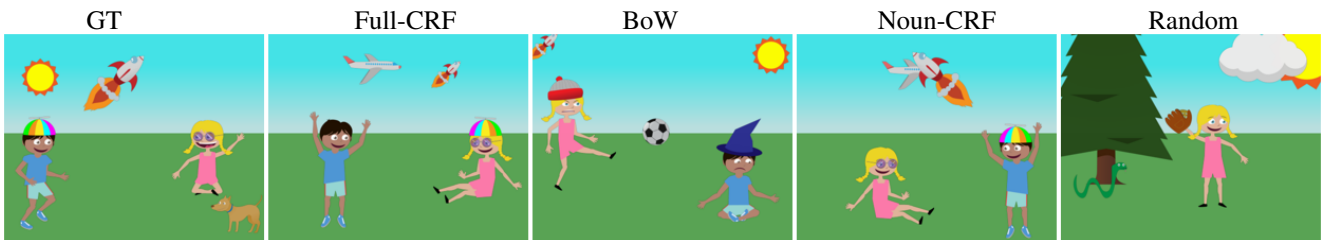


Figure 36: **Input description:** Mike is wearing a propeller beanie hat. Jenny is wearing round glasses. The rocket is flying into the sky. **Tuples:** Mike wear hat; Jenny wear glasses; rocket fly:p:into sky;

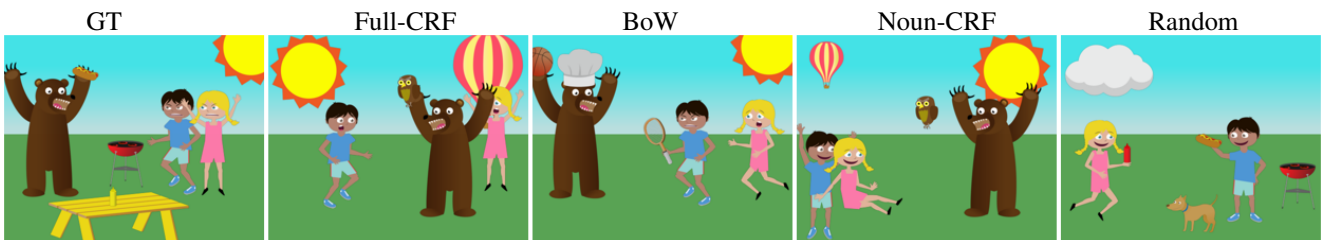


Figure 37: **Input description:** Mike and Jenny are looking at the bear. Mike and Jenny are trying to scare the bear away. Jenny has her arms up in the air while Mike is running at the bear. **Tuples:** Mike look bear; Jenny look bear; Mike try scare; Jenny try scare; Mike scare:ptcl:away bear; Jenny scare:ptcl:away bear; Jenny have:ptcl:up arm; Jenny have air; Mike run bear;

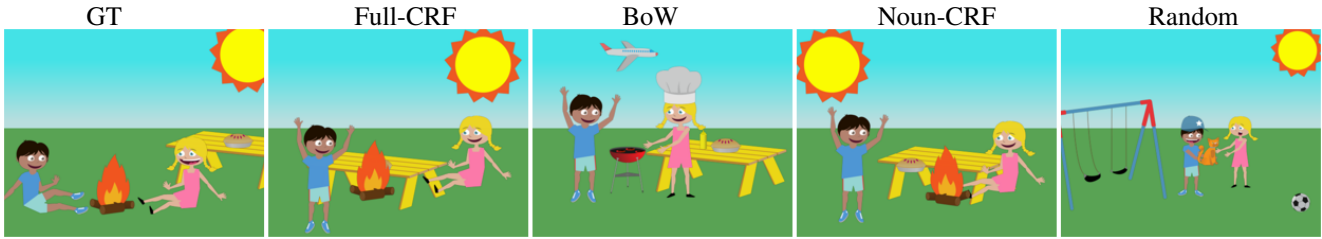


Figure 38: **Input description:** There's a pie on the table Jenny is in front of the table The fire is between Jenny and Mike
Tuples: pie be table; Jenny be ; fire be ;

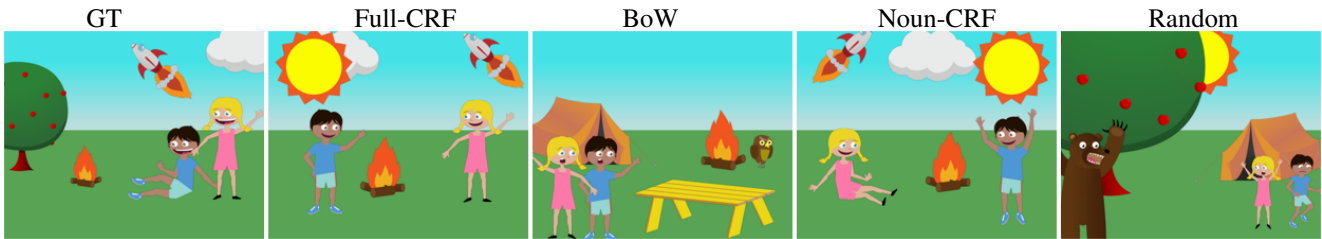


Figure 39: **Input description:** There is a rocket ship in the sky. Mike and Jenny are waving. Mike is sitting far from the fire.
Tuples: rocket ship be ; Mike wave ; Jenny wave ; Mike sit:p:far from fire;



Figure 40: **Input description:** The bear is on the table Jennt and Mike both want the food Mike is in front of Jenny
Tuples: bear be table; Jennt want food; Mike want food; Mike be ;

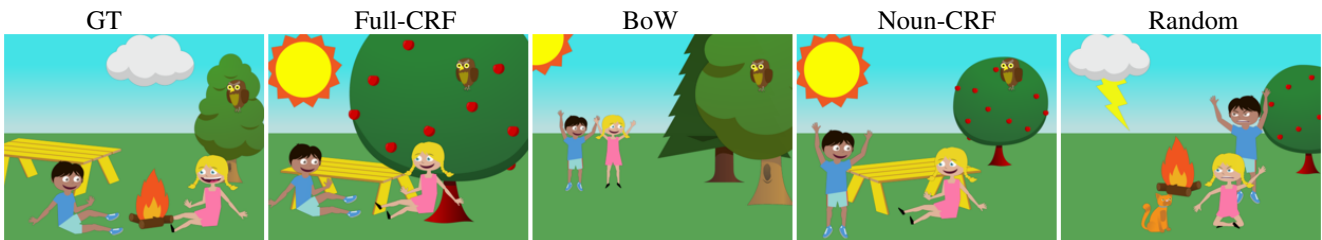


Figure 41: **Input description:** Mike and Jenny are sitting by the fire. An owl is in the tree. The picnic table is behind Mike.
Tuples: Mike sit ; Jenny sit ; owl be tree; table be ;

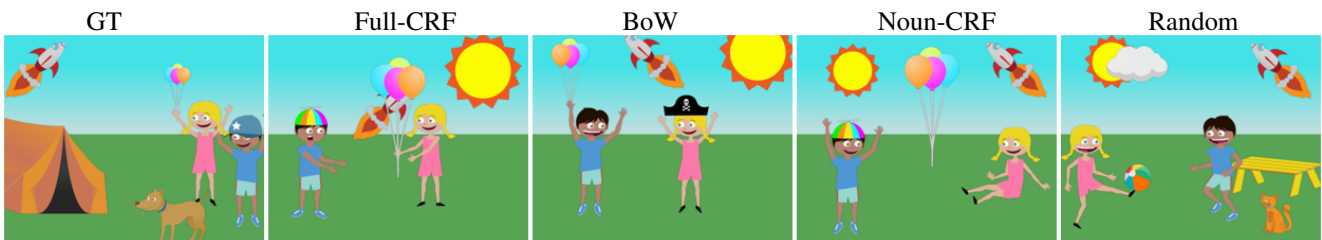


Figure 42: **Input description:** Mike and Jenny are looking at the rocket. Mike is wearing a blue hat. Jenny is carrying balloons.
Tuples: Mike look rocket; Jenny look rocket; Mike wear hat; Jenny carry balloon;

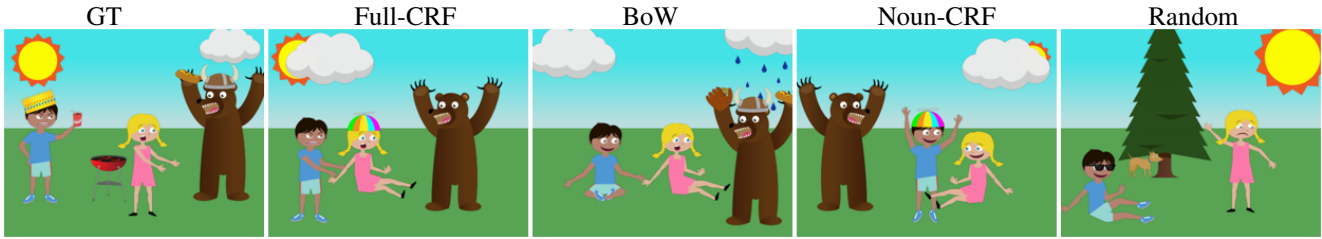


Figure 43: **Input description:** Mike is mad his ice melted. Jenny is scared of the bear. The bear is wearing a viking hat. **Tuples:** Mike be:pa:mad ; Jenny be:pa:scared ; bear wear hat;



Figure 44: **Input description:** Jenny is sitting by the fire. Mike is wearing the hat. The pie is on the table. **Tuples:** Jenny sit ; Mike wear hat; pie be table;

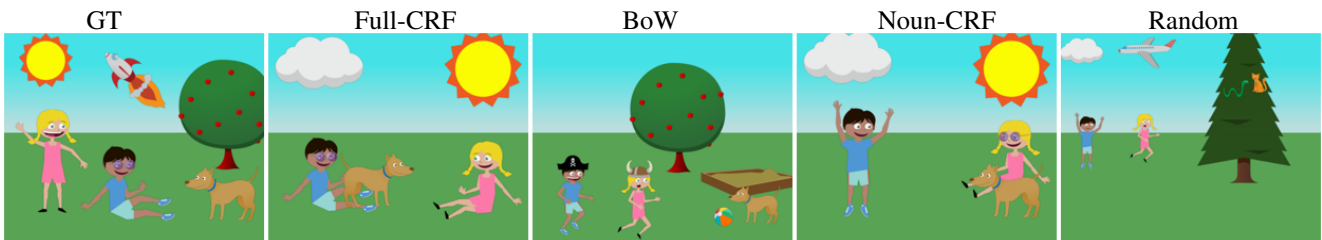


Figure 45: **Input description:** Mike is sitting near the dog. Mike is wearing classes. Jenny is smiling at the dog. **Tuples:** Mike sit:p:near dog; Mike wear class; Jenny smile dog;



Figure 46: **Input description:** The bear came to join the cookout Jenny reminds him mustard gives him a tummy ache Mike wishes the bear would just take his food and go **Tuples:** bear join cookout; Jenny remind give; mustard give ache; Mike wish take; Mike wish go; bear take food; bear go ;



Figure 47: **Input description:** Mike and Jenny are sitting by a fire. Mike and Jenny are having a picnic. Mike is sitting under a tree. **Tuples:** Mike sit ; Jenny sit ; Mike have picnic; Jenny have picnic; Mike sit tree;

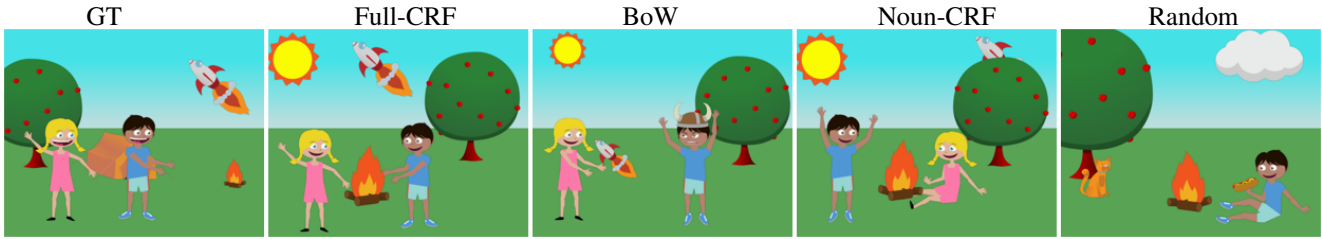


Figure 48: **Input description:** Jenny waved to the rocket. Mike is standing near the fire. There are apples on the tree. **Tuples:** Jenny wave:p:to rocket; Mike stand:p:near fire; apple be tree;



Figure 49: **Input description:** A bear scared Jenny. The bear stole Jenny's hot dog. Jenny has a ketchup bottle. **Tuples:** bear scare ; bear steal hot dog; Jenny have bottle;

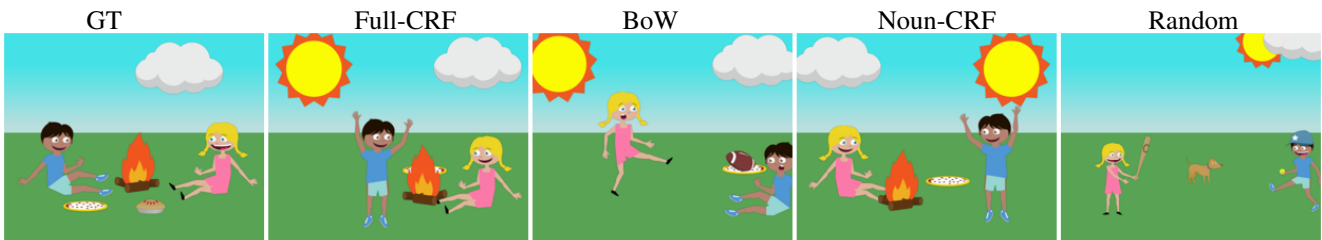


Figure 50: **Input description:** There is a campfire. Mike is near the campfire. There is pizza on the ground. **Tuples:** campfire be ; Mike be ; pizza be ground;