

## Summary

### 1. Entity in Centroid (Cosine) vs Entity in Words(Group Average)

Method	mAP@1 %	mAP@2 %	mAP@3 %	mAP@4 %	mAP@5 %
WV-P-LC + E-V-C-N10 + T-V-LC + GA	55.29	62.13	64.44	65.24	65.29
WV-P-LC + E-W-N10 + T-V-LC + GA	53.20	59.92	62.98	63.71	63.96
WV-P-LC + E-V-D2V-N10 + T-V-LC + GA	50.68	58.28	61.15	62.55	62.66

### 2. Type in Vector vs Type in Similar Words

WV-P-LC + E-W-N10 + T-V-LC + GA	53.20	59.92	62.98	63.71	63.96
WV-P-LC + E-W-N10 + T-SW30-TV-LC + GA	51.57	59.39	61.96	63.19	63.17

### 3. Different Word Representations.

WV-P-LC + E-W-N10 + T-SW30-TV-LC + GA	51.57	59.39	61.96	63.19	63.17
WV-P + E-W-N10 + T-SW30-TV-LC + GA	48.25	57.37	60.10	60.66	61.43
WV-LC + E-W-N10 + T-SW30-TV-LC + GA	49.42	58.00	60.58	61.82	62.30

## Sentence vs Doc Level

### 1. Taking nearby words of all the entity mentions as doc level context word

WV-P-LC + **E-V-C-N10** + T-V-LC + GA

Method	mAP@1 %	mAP@2 %	mAP@3 %	mAP@4 %	mAP@5 %
Document Level	<b>55.92</b>	61.62	63.98	<b>65.26</b>	<b>65.43</b>
Sentence Level	55.29	<b>62.13</b>	<b>64.44</b>	65.24	65.29

WV-P-LC + **E-V-D2V-N10** + T-V-LC + GA

Document Level	<b>51.74</b>	<b>61.33</b>	<b>63.81</b>	<b>64.52</b>	<b>64.77</b>
Sentence Level	50.68	58.28	61.15	62.55	62.66

WV-P-LC + E-W-N10 + **T-SW30-TV-LC** + GA

Document Level	51.37	<b>59.55</b>	<b>62.79</b>	<b>63.67</b>	<b>64.07</b>
Sentence Level	<b>51.57</b>	59.39	61.96	63.19	63.17

## Word Representation Codes

WV-P-LC: Word Vector, Learned on Corpus, Initialized with  
Pretrained Glove

WV-P : Word Vector, Pretrained Glove

WV-LC: Word Vector, Learned on Corpus

## Entity Representation Codes

E-V-C-N10: Entity in Vector form, Centroid of Nearby 10  
words

E-W-N10: Entity in Words, Nearby 10 words

E-V-D2V-N10: Entity in Vector form, Doc2Vec of Nearby 10  
words

## Type Representation Codes

T-V-LC: Type in Vector form, Learned on Corpus

T-SW30-TV-LC: Type in Similar 30 Words to Type Vector,  
Learned on Corpus

## Ranking Entity Against Role Metrics

COS: COSine distance

GA: Group Average

**Configuration#1.1: WV-P-LC + E-V-C-N10 + T-V-LC + COS**

**Corpus Preprocessing:** None

**Word Representation:** Word vectors trained on corpus initializing with pretrained vector

**Entity Context Window:** Left 10 words, Right 10 words

**Entity Representation Type:** Context words and words in Entity and Entity words Centroid

**Type Representation:** Learned vector by replacing entities by tag where words are initialized with pretrained vector

**Ranking mechanism:** and cosine distance with Type vector

**Results:**

Role	P@1	P@2	P@3	P@4	P@5
LOC_Event	68.78 %	62.17 %	57.85 %	56.48 %	54.60 %
LOC_Accused	28.12 %	20.31 %	19.27 %	17.97 %	19.69 %
LOC_Victim	16.67 %	16.67 %	19.44 %	19.44 %	17.78 %
LOC_Others	59.50 %	59.25 %	58.58 %	58.33 %	58.88 %
ORG_Accused	67.89 %	65.60 %	59.63 %	57.11 %	54.08 %
ORG_Victim	40.00 %	36.67 %	34.44 %	37.22 %	39.22 %
ORG_Others	76.00 %	77.50 %	78.92 %	77.46 %	77.78 %
PER_Victim	59.37 %	51.56 %	53.64 %	52.60 %	52.45 %
PER_Others	85.79 %	83.52 %	81.91 %	81.15 %	81.10 %
PER_Accused	50.79 %	51.59 %	56.08 %	58.07 %	58.86 %
Average	55.29 %	52.48 %	51.98 %	51.58 %	51.44 %

## **Configuration#1.2: WV-P-LC + E-V-C-N10 + T-V-LC + GA**

**Corpus Preprocessing:** None

**Word Representation:** Word vectors trained on corpus initializing with pretrained vector

**Entity Context Window:** Left 10 words, Right 10 words

**Entity Representation Type:** Context words and words in Entity and Entity words Centroid

**Type Representation:** Learned vector by replacing entities by tag where words are initialized with pretrained vector

**Ranking mechanism:** Group Average

### **Results:**

<b>Role</b>	<b>P@1</b>	<b>P@2</b>	<b>P@3</b>	<b>P@4</b>	<b>P@5</b>
LOC_Event	44.97	48.67	49.21	48.41	47.83
LOC_Accused	28.12	20.31	22.39	19.53	19.06
LOC_Victim	41.66	29.17	27.77	25.64	26.11
LOC_Others	64.5	62.25	62.42	62.33	62.88
ORG_Accused	50.46	51.83	52.60	52.29	51.88
ORG_Victim	40.00	40.00	38.89	37.22	37.88
ORG_Others	79.00	77.00	77.08	76.71	76.78
PER_Victim	37.50	48.44	48.44	51.04	49.94
PER_Others	80.68	81.53	80.59	80.44	80.53
PER_Accused	65.08	55.55	56.10	56.88	56.00
<b>Average</b>	<b>53.20</b>	<b>51.48</b>	<b>51.55</b>	<b>51.06</b>	<b>50.89</b>

## **Configuration#2: WV-P-LC + E-W-N10 + T-V-LC + GA**

**Corpus Preprocessing:** None

**Word Representation:** Word vectors trained on corpus initializing with pretrained vector

**Entity Context Window:** Left 10 words, Right 10 words

**Entity Representation Type:** Context words and words in Entity

**Type Representation:** Learned vector by replacing entities by tag where words are initialized with pretrained vector

**Ranking mechanism:** Goup average agglomerative among Entity word vector and Type vector

### **Results:**

<b>Role</b>	<b>P@1</b>	<b>P@2</b>	<b>P@3</b>	<b>P@4</b>	<b>P@5</b>
LOC_Event	68.78 %	62.17 %	57.85 %	56.48 %	54.60 %
LOC_Accused	28.12 %	20.31 %	19.27 %	17.97 %	19.69 %
LOC_Victim	16.67 %	16.67 %	19.44 %	19.44 %	17.78 %
LOC_Others	59.50 %	59.25 %	58.58 %	58.33 %	58.88 %
ORG_Accused	67.89 %	65.60 %	59.63 %	57.11 %	54.08 %
ORG_Victim	40.00 %	36.67 %	34.44 %	37.22 %	39.22 %
ORG_Others	76.00 %	77.50 %	78.92 %	77.46 %	77.78 %
PER_Victim	59.37 %	51.56 %	53.64 %	52.60 %	52.45 %
PER_Others	85.79 %	83.52 %	81.91 %	81.15 %	81.10 %
PER_Accused	50.79 %	51.59 %	56.08 %	58.07 %	58.86 %
<b>Average</b>	<b>55.29 %</b>	<b>52.48 %</b>	<b>51.98 %</b>	<b>51.58 %</b>	<b>51.44 %</b>

### **Configuration#3.1: WV-P-LC + E-W-N10 + T-SW30-TV-LC + GA**

**Corpus Preprocessing:** None

**Word Representation:** Word vectors trained on corpus initializing with pretrained vector

**Entity Context Window:** Left 10 words, Right 10 words

**Entity Representation Type:** Context words and words in Entity

**Type Representation:** 30 Similar words to the learned type vector by replacing entities by tag where words are initialized with pretrained vector

**Ranking mechanism:** Group average agglomerative among Entity word vector and Type vector

#### **Results:**

<b>Role</b>	<b>P@1</b>	<b>P@2</b>	<b>P@3</b>	<b>P@4</b>	<b>P@5</b>
LOC_Event					
LOC_Accused					
LOC_Victim					
LOC_Others					
ORG_Accused					
ORG_Victim					
ORG_Others					
PER_Victim					
PER_Others					
PER_Accused					
<b>Average</b>					

### **Configuration#3.2 WV-P + E-W-N10 + T-SW30-TV-LC + GA**

**Word Representation:** Word vectors pretrained. If word not present initialize with 0 vector

#### **Results:**

<b>Role</b>	<b>P@1</b>	<b>P@2</b>	<b>P@3</b>	<b>P@4</b>	<b>P@5</b>
LOC_Event	53.97	51.85	52.20	52.64	

LOC_Accused	12.5	17.19	18.23	17.19	
LOC_Victim	16.67	16.67	16.67	15.28	
LOC_Others	60.5	62.00	62.42	62.08	
ORG_Accused	51.37	50.00	48.01	48.63	
ORG_Victim	26.67	40.00	41.11	38.89	
ORG_Others	78.00	78.00	78.25	78.08	
PER_Victim	53.12	50.00	50.52	49.48	
PER_Others	85.23	83.24	83.43	82.00	
PER_Accused	44.44	50.00	51.32	51.32	
<b>Average</b>	<b>48.25</b>	<b>49.89</b>	<b>50.22</b>	<b>49.56</b>	

### **Configuration#3.3 WV-LC + E-W-N10 + T-SW30-TV-LC + GA**

**Word Representation:** Word vectors trained on corpus using 1-hot vector.  
If word not present initialize with 0 vector

#### **Results:**

<b>Role</b>	<b>P@1</b>	<b>P@2</b>	<b>P@3</b>	<b>P@4</b>	<b>P@5</b>
LOC_Event					
LOC_Accused					
LOC_Victim					
LOC_Others					
ORG_Accused					
ORG_Victim					
ORG_Others					
PER_Victim					
PER_Others					
PER_Accused					
<b>Average</b>					