

Errata

The scores in Table 2-6 are incorrect. We apologize for the error and correct it.

error

	Web			News		
	case analysis	zero anaphora resolution	coreference resolution	case analysis	zero anaphora resolution	coreference resolution
S&K (2018)	89.2	58.1	68.5	89.5	35.6	54.1
Kurita et al. (2018)	92.0	58.4 ¹	-	-	-	-
CAModel (ours)	88.4 (± 0.2)	70.5 (± 0.2)	78.2 (± 0.4)	85.8 (± 0.3)	56.8 (± 0.9)	68.1 (± 0.6)

Table 2: Our CAModel performance (F-measure) compared with Shibata and Kurohashi (2018) and Kurita et al. (2018).

correct

	Web			News		
	case analysis	zero anaphora resolution	coreference resolution	case analysis	zero anaphora resolution	coreference resolution
S&K (2018)	89.2	58.1	68.5	89.5	35.6	54.1
Kurita et al. (2018)	92.0	58.4	-	-	-	-
CAModel (ours)	88.3 (± 0.5)	70.3 (± 0.7)	77.0 (± 0.6)	85.9 (± 0.5)	56.7 (± 0.5)	57.6 (± 0.8)

Table 2: Our CAModel performance (F-measure) compared with Shibata and Kurohashi (2018) and Kurita et al. (2018).

error

	VPA (Web)	VPA (News)
CAModel (VPA)	76.91 (± 0.36)	68.70 (± 0.70)
CAModel (VPA + CR)	76.88 (± 0.20)	68.39 (± 0.31)
CAModel (VPA + NPA)	77.38 (± 0.17)	69.55 (± 0.29)
CAModel (VPA + BAR)	76.42 (± 0.47)	69.34 (± 0.39)
CAModel (VPA + NPA + BAR)	77.41 (± 0.25)	69.76 (± 0.38)
CAModel (VPA + NPA + BAR + CR)	77.34 (± 0.13)	69.37 (± 0.54)
CorefCAModel (CR \rightarrow VPA + NPA + BAR + CR)	76.88 (± 0.24)	69.04 (± 0.20)

Table 3: Performance (F-measure) of verbal predicate-argument structure analysis (**VPA**) evaluated on the test set of the Web and News corpora.

	NPA (Web)	NPA (News)
CAModel (NPA)	58.82 (± 1.98)	54.52 (± 0.86)
CAModel (NPA + CR)	59.00 (± 1.23)	53.39 (± 1.06)
CAModel (NPA + VPA)	61.99 (± 1.11)	56.27 (± 0.44)
CAModel (NPA + BAR)	58.64 (± 2.01)	54.15 (± 1.18)
CAModel (NPA + VPA + BAR)	61.16 (± 0.73)	56.33 (± 0.30)
CAModel (NPA + VPA + BAR + CR)	61.65 (± 0.35)	55.78 (± 0.67)
CorefCAModel (CR \rightarrow NPA + VPA + BAR + CR)	60.47 (± 0.23)	54.87 (± 0.24)

Table 4: Performance of nominal predicate-argument structure analysis (**NPA**) evaluated on the test set of the Web and News corpora.

	BAR (Web)	BAR (News)
CAModel (BAR)	61.53 (± 0.32)	55.77 (± 0.61)
CAModel (BAR + CR)	60.82 (± 0.89)	54.86 (± 0.08)
CAModel (BAR + VPA)	62.03 (± 1.19)	55.76 (± 0.73)
CAModel (BAR + NPA)	60.98 (± 1.40)	54.91 (± 0.97)
CAModel (BAR + VPA + NPA)	61.10 (± 0.83)	55.00 (± 0.96)
CAModel (BAR + VPA + NPA + CR)	59.60 (± 1.15)	54.30 (± 0.40)
CorefCAModel (CR \rightarrow BAR + VPA + NPA + CR)	59.92 (± 0.49)	53.11 (± 0.86)

Table 5: Performance of bridging anaphora resolution (**BAR**) evaluated on the test set of the Web and News corpora.

	CR (Web)	CR (News)
CAModel (CR)	79.13 (± 0.63)	60.11 (± 0.81)
CAModel (CR + VPA)	78.82 (± 0.65)	59.42 (± 0.42)
CAModel (CR + NPA)	78.31 (± 0.31)	59.27 (± 0.85)
CAModel (CR + BAR)	78.53 (± 0.39)	58.92 (± 0.71)
CAModel (CR + VPA + NPA)	78.59 (± 0.72)	59.24 (± 0.51)
CAModel (CR + VPA + NPA + BAR)	78.20 (± 0.36)	58.42 (± 0.48)
CorefCAModel (CR \rightarrow VPA + NPA + BAR + CR)	80.87 (± 0.25)	59.61 (± 0.67)

Table 6: Performance (F-measure) of coreference resolution (**CR**) evaluated on the test set of the Web and News corpora.

correct

	VPA (Web)	VPA (News)
CAModel (VPA)	76.91 (± 0.36)	68.70 (± 0.70)
CAModel (VPA + CR)	77.04 (± 0.25)	68.77 (± 0.24)
CAModel (VPA + NPA)	77.38 (± 0.17)	69.55 (± 0.29)
CAModel (VPA + BAR)	76.42 (± 0.47)	69.34 (± 0.39)
CAModel (VPA + NPA + BAR)	77.41 (± 0.25)	69.76 (± 0.38)
CAModel (VPA + NPA + BAR + CR)	77.15 (± 0.62)	69.37 (± 0.26)
CorefCAModel (CR \rightarrow VPA + NPA + BAR + CR)	77.56 (± 0.52)	69.27 (± 0.25)

Table 3: Performance (F-measure) of verbal predicate-argument structure analysis (**VPA**) evaluated on the test set of the Web and News corpora.

	NPA (Web)	NPA (News)
CAModel (NPA)	58.82 (± 1.98)	54.52 (± 0.86)
CAModel (NPA + CR)	59.58 (± 0.56)	53.15 (± 0.94)
CAModel (NPA + VPA)	61.99 (± 1.11)	56.27 (± 0.44)
CAModel (NPA + BAR)	58.64 (± 2.01)	54.15 (± 1.18)
CAModel (NPA + VPA + BAR)	61.16 (± 0.73)	56.33 (± 0.30)
CAModel (NPA + VPA + BAR + CR)	61.82 (± 0.80)	55.77 (± 0.52)
CorefCAModel (CR \rightarrow NPA + VPA + BAR + CR)	61.45 (± 0.50)	55.43 (± 0.95)

Table 4: Performance of nominal predicate-argument structure analysis (**NPA**) evaluated on the test set of the Web and News corpora.

	BAR (Web)	BAR (News)
CAModel (BAR)	61.53 (± 0.32)	55.77 (± 0.61)
CAModel (BAR + CR)	60.19 (± 0.80)	54.00 (± 0.95)
CAModel (BAR + VPA)	62.03 (± 1.19)	55.76 (± 0.73)
CAModel (BAR + NPA)	60.98 (± 1.40)	54.91 (± 0.97)
CAModel (BAR + VPA + NPA)	61.10 (± 0.83)	55.00 (± 0.96)
CAModel (BAR + VPA + NPA + CR)	60.39 (± 1.12)	53.93 (± 1.26)
CorefCAModel (CR \rightarrow BAR + VPA + NPA + CR)	60.14 (± 0.92)	54.32 (± 1.00)

Table 5: Performance of bridging anaphora resolution (**BAR**) evaluated on the test set of the Web and News corpora.

	CR (Web)	CR (News)
CAModel (CR)	77.81 (± 0.62)	59.94 (± 0.75)
CAModel (CR + VPA)	77.36 (± 0.49)	58.89 (± 0.48)
CAModel (CR + NPA)	77.30 (± 0.33)	58.40 (± 0.58)
CAModel (CR + BAR)	77.88 (± 0.39)	58.13 (± 0.77)
CAModel (CR + VPA + NPA)	77.46 (± 0.84)	58.74 (± 0.50)
CAModel (CR + VPA + NPA + BAR)	77.02 (± 0.60)	57.55 (± 0.76)
CorefCAModel (CR \rightarrow VPA + NPA + BAR + CR)	76.56 (± 0.58)	57.58 (± 1.28)

Table 6: Performance (F-measure) of coreference resolution (**CR**) evaluated on the test set of the Web and News corpora.