Table 4: Comparison results with state-of-the-art methods on multi-paragraph video grounding in SynopGround. Table 6: Effect of the local-level modeling, global-level model-Method mIoU IoU@0.3 IoU@0.5 IoU@0.7 ing, a Human 85.1 97.3 92.5 85.0

8.3

47.2

52.7

67.9

3.2

28.7

29.3

46.7

0.8

12.8

10.5

21.8

Random

DepNet [5]

PRVG [63]

LGMR (Ours)

7.3

30.7

34.7

44.4

ng, and cross-level reasoning in the iterative query decod						ry decoder.
Local	Global	Cross	mIoU	IoU@0.3	IoU@0.5	IoU@0.7
			ı			

53.1

66.6

67.9

32.4

44.2

46.7

13.5

19.0

21.8

34.5

42.8

44.4