

Attention 기반 한국어 의미역 결정

A Neural Attention model for Korean Semantic Role Labeling

저자 (Authors)	박광현, 나승훈 Kwang-Hyeon Park, Seung-Hoon Na
출처 (Source)	한국정보과학회 학술발표논문집 , 2017.12, 634-636(3 pages)
발행처 (Publisher)	한국정보과학회 KOREA INFORMATION SCIENCE SOCIETY
URL	http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE07322247
APA Style	박광현, 나승훈 (2017). Attention 기반 한국어 의미역 결정. 한국정보과학회 학술발표논문집, 634-636
이용정보 (Accessed)	고려대학교 163.152.3.*** 2019/07/30 17:41 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

Attention 기반 한국어 의미역 결정

박광현, 나승훈

전북대학교

khpark231@gmail.com, nash@jbnu.ac.kr

A Neural Attention model for Korean Semantic Role Labeling

Kwang-Hyeon Park, Seung-Hoon Na

Chonbuk National University

요 약

최근 Attention mechanism을 자연어처리 분야에 적용하는 연구가 많이 이루어지고 있다. Attention mechanism은 입력 문장의 단어에 중요도를 부여해서 출력을 생성할 때 어떤 단어를 더 중요하게 생각해서 출력을 생성할지 결정해주는 역할을 한다. 이러한 Attention mechanism을 이용하면 입력 시퀀스의 길이가 길어져도 정보를 잃지 않고 효율적으로 전달해 줄 수 있다. 본 논문에서는 Attention mechanism을 한국어 의미역 결정에 적용하여 Attention mechanism이 의미역 결정에 성능 향상을 가져올 수 있을지 평가해 보고자 한다.

1. 서 론

의미역(semantic role)은 서술어에 의해 기술되는 행위나 사태에 대한 명사구의 의미 역할을 나타내며 서술어에 의해 기술되는 명사구를 논항(argument)이라고 한다. 의미역은 상호참조 해결, 질의 응답, 정보 추출과 같은 시스템에 사용되어 최근 많은 연구가 이루어지고 있다[1-4,10]. 의미역은 같은 문장이라 하더라도 서술어에 따라 달라지게 된다. 예를 들면 문장 “그는 1일 자신이 맡고 있던 헌법위원회 위원장직을 사임했다.”에서 서술어 맡다의 논항은 (자신이, ARG0), (위원장직을, ARG1)이 되고, 서술어 사임했다의 논항은 (그는, ARG0), (1일, ARG-M-TMP), (위원장직을, ARG1)이 된다. 따라서 본 논문에서는 Bi-LSTM을 통해 얻어진 단어 표상에서 서술어에 해당하는 단어 표상과 각 의미역 임베딩을 결합해 Attention mechanism을 적용하여 서술어에 대한 각 단어들이 가질 수 있는 의미역을 효율적으로 예측할 수 있는 모델을 제안하고자 한다.

2. 관련 연구

최근 딥러닝 모델을 이용한 연구가 활발히 이루어지고 있다. LSTM기반 방법에 출력 노드간의 의존성을 모델링 하는 CRF를 결합한 방식인 LSTM CRF를 이용해 개체명 인식에서 우수한 성능을 보여주었고[16,17], Stack LSTM, Sequence to sequence, SyntaxNet, Deep Biaffine Attention 모델을 이용한 한국어 의존 파싱[5-8], Input Attention을 LSTM-CNN에 적용한 모델을 이용한 Relation Classification[9] 등이 있다. 또한 딥러닝 모델을 의미역 결정에 적용하는 연구도 많이 이루어지고 있다. [1]에서는 RNN에서 문제가 되는 Batch Normalization의 문제점을 해결한 Layer Normalization을 Bi-LSTM CRF에 적용한 모델을 사용하였다. [2]에서는 LSTM을 여러 개 쌓아 올려 각 층마다 Bi-LSTM을 사용하여 양 방향

정보를 모두 활용한 모델을 사용하였다. [3,14]는 LSTM에 비선형 변환(non-linear transform)을 거치지 않은 입력정보를 LSTM 유닛이 사용할 수 있게 LSTM 유닛에 새로운 게이트가 추가하여 비선형 변환을 거치지 않은 정보를 선택적으로 볼 수 있게 설계된 모델을 사용하였다. [4]는 구문 분석 정보인 서술어와 논항 사이의 dependency path라는 자질을 추가하여 dependency path가 의미역 결정에 성능 향상을 가져올 수 있음을 보였다. [10]은 서술어와 멀리 떨어져 있는 논항의 경우 잘 잡아내지 못하는 LSTM을 보완해주기 위해 그래프의 모든 이웃 단어에 대한 관련정보를 인코딩 해주고 ‘teleportation’기능을 가진 Graph Convolutional Network(GCN)을 추가하여 성능이 향상됨을 보였다. [15]는 dependency한 단어 사이에 추가적인 connection이 있고 서로 다른 유형의 dependency 관계에 대해 학습 가능한 가중치를 도입한 Syntax Aware LSTM 모델을 이용해 성능이 향상됨을 보였다. [18]은 Relation Classification과 Semantic role labeling 두 태스크를 동시에 학습하는 multi-task learning을 이용해 의미역 결정에 중요한 개선이 이루어짐을 보여주었다. Attention mechanism을 자연어처리분야에 적용하는 연구도 많이 이루어지고 있다[8,11,12,13]. Attention mechanism은 입력 문장의 단어에 중요도를 부여해 출력을 생성할 때 어떤 단어를 더 중요하게 생각해서 출력을 생성할지 결정해주는 역할을 한다. [12, 13]에서는 Attention mechanism을 문장 번역에 적용하였는데 인코더에서 문장 정보를 고정된 크기의 벡터로 인코딩 하고 디코더에서 출력 문장을 생성할 때 고정 크기 벡터가 성능을 저하 시키는 것을 방지하기 위해 인코더의 어떤 단어에 Attention 하여 생성할지에 대한 계산이 이루어진다.

3. Neural Attention 모델 기반 의미역 결정

의미역은 같은 문장이라 하더라도 서술어에 따라 달라지게 된다. 따라서 본 논문에서 제안하는 모델은 Bi-LSTM을 통해 얻어진 단어 표상에서 서술어에 해당하는 단어 표상과 각 의미역 임베딩을 결합해 Attention mechanism을 적용해 각 의미역이 어떤 단어에 assign 될지에 대한 Attention score를 계산함으로써 해당 서술어에 대한 각 단어들이 가질 수 있는 의미역을 효율적으로 예측 할 수 있게 해준다.

다음 식(1)은 형태소와 음절 단위 문자로부터 얻어진 character임베딩과 additional features를 결합해 얻어지는 단어 표상을 나타낸다.

$$\begin{aligned} \mathbf{X} &= \{\mathbf{x}_1, \dots, \mathbf{x}_n\} \\ \mathbf{x}_i &= [\text{emb}(\text{char}); \text{emb}(\text{ejtag}); \text{emb}(\text{loc})] \\ \mathbf{h} &= \text{Bi-LSTM}(\mathbf{X}) \end{aligned} \quad (1)$$

식(1)에서 $\text{emb}(\ast)$ 는 \ast 의 임베딩을 가져오는 함수이고, char는 character, ejtag는 어절태그, loc는 서술어와 현재 어절사이의 위치 관계, \mathbf{h} 는 단어에 대한 표상을 나타낸다.

$$\begin{aligned} \mathbf{Q} &= [\mathbf{SR}; \mathbf{h}_j] \\ \mathbf{K} &= \mathbf{h} \\ \text{score} &= \mathbf{Q} \mathbf{W}_a \mathbf{K}^T \\ \alpha &= \text{softmax}(\text{score}) \end{aligned} \quad (2)$$

식(2)는 서술어와 의미역 그리고 단어 사이의 Attention score를 계산하는 식으로 j 는 서술어에 대한 위치를 나타내며 SR은 각 의미역에 대한 임베딩을 의미하고 α 는 각 의미역이 어떤 단어에 assign될지에 대한 score가 된다. 논항의 인식 및 분류에서 명사구의 끝 단어에만 의미역을 부여하기 때문에 대부분의 단어는 None label을 갖게 된다. 따라서 root 임베딩을 따로 두고 해당 의미역이 root를 가리키는 경우 해당 의미역은 문장에 없는 의미역으로 처리하였다. 다음 그림 1은 본 논문에서 제안하는 Neural Attention 모델 구조를 도식화 하여 보여준다.

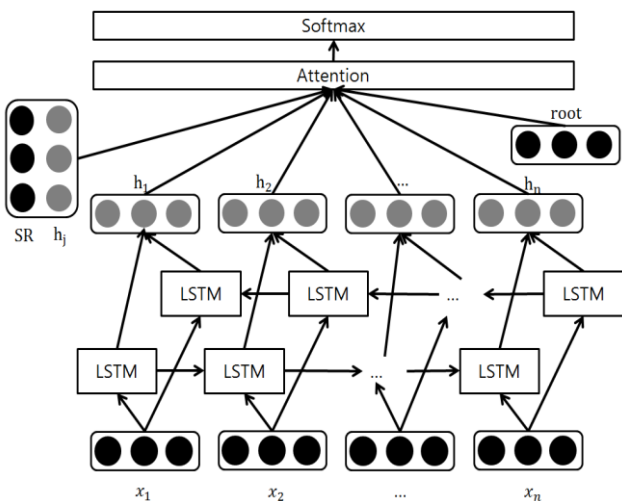


Figure 1. 제안 Neural Attention 모델

3.1 입력

본 논문에서는 입력으로 문자 LSTM기반 단어 표상(representation)과 additional features를 사용하였다. 문자 LSTM기반 단어 표상은 형태소-태그 문자와 음절-태그

문자를 각각 Bi-LSTM을 적용하였다. Additional features는 어절태그, 서술어의 어휘 및 품사정보, 서술어와 현재 어절 사이의 위치 정보, head의 위치가 사용되었다. 다음 표 1은 additional features의 예시를 보여준다.

Table 1. Additional features

Feature	Example
어절태그	남학생/NNG 이/JKS → NNG~JKS
서술어와 현재 어절 사이의 위치 정보	문장 “자신이 저지른 행동의 중대성을 인식하지”에서 서술어 “인식하지”와 현재 어절 “자신이”의 위치 정보는 PREV 4가 된다.

3.2 출력

문자들의 임베딩 벡터(character embedding vector)와 자질 임베딩 벡터(feature embedding vector)를 결합하여 Bi-LSTM을 적용해 단어 표상(word representation)을 얻어내고, 서술어, 의미역, 단어 사이의 Attention을 구하여 softmax를 적용해 각 의미역이 어떤 단어에 assign될지에 대한 출력을 생성 하였다.

4. 실험

4.1 실험 셋팅

본 논문에서는 의미역 결정 평가를 위해 Korean Propbank의 Newswire 말뭉치만을 사용하였고, Tree 구조의 말뭉치를 변환하는 도중 오류가 있는 문장은 제외하였다. 전체 문장 23659 문장 중 20110 문장은 학습데이터로, 1183 문장은 개발셋, 2366 문장은 평가셋으로 사용하였다.

실험에 사용한 음절과 형태소 문자는 300차원 임베딩 벡터를 사용하였고 형태소 문자는 glove를 통해 사전학습된 임베딩 벡터를 사용하였으며 은닉층은 100차원으로 하였다. additional features는 50차원 임베딩 벡터를 사용하였고, dropout은 0.5의 값을 적용하였다.

실험에 사용된 의미역의 수는 None label ‘O’를 포함하여 27개로 다음 표 2는 사용된 의미역 태그를 나타낸다.

Table 2. 의미역 태그

labels	ARG0, ARG1, ARG2, ARG3, ARG5, ARG4, ARG6, ARG7, ARG8, ARG9, ARG10, ARG11, ARG12, ARG13, ARG14, ARG15, ARG16, ARG17, ARG18, ARG19, ARG20, ARG21, ARG22, ARG23, ARG24, ARG25, ARG26, ARG27, ARG28, ARG29, ARG30, ARG31, ARG32, ARG33, ARG34, ARG35, ARG36, ARG37, ARG38, ARG39, ARG40, ARG41, ARG42, ARG43, ARG44, ARG45, ARG46, ARG47, ARG48, ARG49, ARG50, ARG51, ARG52, ARG53, ARG54, ARG55, ARG56, ARG57, ARG58, ARG59, ARG60, ARG61, ARG62, ARG63, ARG64, ARG65, ARG66, ARG67, ARG68, ARG69, ARG70, ARG71, ARG72, ARG73, ARG74, ARG75, ARG76, ARG77, ARG78, ARG79, ARG80, ARG81, ARG82, ARG83, ARG84, ARG85, ARG86, ARG87, ARG88, ARG89, ARG90, ARG91, ARG92, ARG93, ARG94, ARG95, ARG96, ARG97, ARG98, ARG99, ARG100, ARG101, ARG102, ARG103, ARG104, ARG105, ARG106, ARG107, ARG108, ARG109, ARG110, ARG111, ARG112, ARG113, ARG114, ARG115, ARG116, ARG117, ARG118, ARG119, ARG120, ARG121, ARG122, ARG123, ARG124, ARG125, ARG126, ARG127, ARG128, ARG129, ARG130, ARG131, ARG132, ARG133, ARG134, ARG135, ARG136, ARG137, ARG138, ARG139, ARG140, ARG141, ARG142, ARG143, ARG144, ARG145, ARG146, ARG147, ARG148, ARG149, ARG150, ARG151, ARG152, ARG153, ARG154, ARG155, ARG156, ARG157, ARG158, ARG159, ARG160, ARG161, ARG162, ARG163, ARG164, ARG165, ARG166, ARG167, ARG168, ARG169, ARG170, ARG171, ARG172, ARG173, ARG174, ARG175, ARG176, ARG177, ARG178, ARG179, ARG180, ARG181, ARG182, ARG183, ARG184, ARG185, ARG186, ARG187, ARG188, ARG189, ARG190, ARG191, ARG192, ARG193, ARG194, ARG195, ARG196, ARG197, ARG198, ARG199, ARG200, ARG201, ARG202, ARG203, ARG204, ARG205, ARG206, ARG207, ARG208, ARG209, ARG210, ARG211, ARG212, ARG213, ARG214, ARG215, ARG216, ARG217, ARG218, ARG219, ARG220, ARG221, ARG222, ARG223, ARG224, ARG225, ARG226, ARG227, ARG228, ARG229, ARG230, ARG231, ARG232, ARG233, ARG234, ARG235, ARG236, ARG237, ARG238, ARG239, ARG240, ARG241, ARG242, ARG243, ARG244, ARG245, ARG246, ARG247, ARG248, ARG249, ARG250, ARG251, ARG252, ARG253, ARG254, ARG255, ARG256, ARG257, ARG258, ARG259, ARG260, ARG261, ARG262, ARG263, ARG264, ARG265, ARG266, ARG267, ARG268, ARG269, ARG270, ARG271, ARG272, ARG273, ARG274, ARG275, ARG276, ARG277, ARG278, ARG279, ARG280, ARG281, ARG282, ARG283, ARG284, ARG285, ARG286, ARG287, ARG288, ARG289, ARG290, ARG291, ARG292, ARG293, ARG294, ARG295, ARG296, ARG297, ARG298, ARG299, ARG300, ARG301, ARG302, ARG303, ARG304, ARG305, ARG306, ARG307, ARG308, ARG309, ARG310, ARG311, ARG312, ARG313, ARG314, ARG315, ARG316, ARG317, ARG318, ARG319, ARG320, ARG321, ARG322, ARG323, ARG324, ARG325, ARG326, ARG327, ARG328, ARG329, ARG330, ARG331, ARG332, ARG333, ARG334, ARG335, ARG336, ARG337, ARG338, ARG339, ARG340, ARG341, ARG342, ARG343, ARG344, ARG345, ARG346, ARG347, ARG348, ARG349, ARG350, ARG351, ARG352, ARG353, ARG354, ARG355, ARG356, ARG357, ARG358, ARG359, ARG360, ARG361, ARG362, ARG363, ARG364, ARG365, ARG366, ARG367, ARG368, ARG369, ARG370, ARG371, ARG372, ARG373, ARG374, ARG375, ARG376, ARG377, ARG378, ARG379, ARG380, ARG381, ARG382, ARG383, ARG384, ARG385, ARG386, ARG387, ARG388, ARG389, ARG390, ARG391, ARG392, ARG393, ARG394, ARG395, ARG396, ARG397, ARG398, ARG399, ARG400, ARG401, ARG402, ARG403, ARG404, ARG405, ARG406, ARG407, ARG408, ARG409, ARG410, ARG411, ARG412, ARG413, ARG414, ARG415, ARG416, ARG417, ARG418, ARG419, ARG420, ARG421, ARG422, ARG423, ARG424, ARG425, ARG426, ARG427, ARG428, ARG429, ARG430, ARG431, ARG432, ARG433, ARG434, ARG435, ARG436, ARG437, ARG438, ARG439, ARG440, ARG441, ARG442, ARG443, ARG444, ARG445, ARG446, ARG447, ARG448, ARG449, ARG450, ARG451, ARG452, ARG453, ARG454, ARG455, ARG456, ARG457, ARG458, ARG459, ARG460, ARG461, ARG462, ARG463, ARG464, ARG465, ARG466, ARG467, ARG468, ARG469, ARG470, ARG471, ARG472, ARG473, ARG474, ARG475, ARG476, ARG477, ARG478, ARG479, ARG480, ARG481, ARG482, ARG483, ARG484, ARG485, ARG486, ARG487, ARG488, ARG489, ARG490, ARG491, ARG492, ARG493, ARG494, ARG495, ARG496, ARG497, ARG498, ARG499, ARG500, ARG501, ARG502, ARG503, ARG504, ARG505, ARG506, ARG507, ARG508, ARG509, ARG510, ARG511, ARG512, ARG513, ARG514, ARG515, ARG516, ARG517, ARG518, ARG519, ARG520, ARG521, ARG522, ARG523, ARG524, ARG525, ARG526, ARG527, ARG528, ARG529, ARG530, ARG531, ARG532, ARG533, ARG534, ARG535, ARG536, ARG537, ARG538, ARG539, ARG540, ARG541, ARG542, ARG543, ARG544, ARG545, ARG546, ARG547, ARG548, ARG549, ARG550, ARG551, ARG552, ARG553, ARG554, ARG555, ARG556, ARG557, ARG558, ARG559, ARG560, ARG561, ARG562, ARG563, ARG564, ARG565, ARG566, ARG567, ARG568, ARG569, ARG570, ARG571, ARG572, ARG573, ARG574, ARG575, ARG576, ARG577, ARG578, ARG579, ARG580, ARG581, ARG582, ARG583, ARG584, ARG585, ARG586, ARG587, ARG588, ARG589, ARG590, ARG591, ARG592, ARG593, ARG594, ARG595, ARG596, ARG597, ARG598, ARG599, ARG600, ARG601, ARG602, ARG603, ARG604, ARG605, ARG606, ARG607, ARG608, ARG609, ARG610, ARG611, ARG612, ARG613, ARG614, ARG615, ARG616, ARG617, ARG618, ARG619, ARG620, ARG621, ARG622, ARG623, ARG624, ARG625, ARG626, ARG627, ARG628, ARG629, ARG630, ARG631, ARG632, ARG633, ARG634, ARG635, ARG636, ARG637, ARG638, ARG639, ARG640, ARG641, ARG642, ARG643, ARG644, ARG645, ARG646, ARG647, ARG648, ARG649, ARG650, ARG651, ARG652, ARG653, ARG654, ARG655, ARG656, ARG657, ARG658, ARG659, ARG660, ARG661, ARG662, ARG663, ARG664, ARG665, ARG666, ARG667, ARG668, ARG669, ARG670, ARG671, ARG672, ARG673, ARG674, ARG675, ARG676, ARG677, ARG678, ARG679, ARG680, ARG681, ARG682, ARG683, ARG684, ARG685, ARG686, ARG687, ARG688, ARG689, ARG690, ARG691, ARG692, ARG693, ARG694, ARG695, ARG696, ARG697, ARG698, ARG699, ARG700, ARG701, ARG702, ARG703, ARG704, ARG705, ARG706, ARG707, ARG708, ARG709, ARG710, ARG711, ARG712, ARG713, ARG714, ARG715, ARG716, ARG717, ARG718, ARG719, ARG720, ARG721, ARG722, ARG723, ARG724, ARG725, ARG726, ARG727, ARG728, ARG729, ARG730, ARG731, ARG732, ARG733, ARG734, ARG735, ARG736, ARG737, ARG738, ARG739, ARG740, ARG741, ARG742, ARG743, ARG744, ARG745, ARG746, ARG747, ARG748, ARG749, ARG750, ARG751, ARG752, ARG753, ARG754, ARG755, ARG756, ARG757, ARG758, ARG759, ARG760, ARG761, ARG762, ARG763, ARG764, ARG765, ARG766, ARG767, ARG768, ARG769, ARG770, ARG771, ARG772, ARG773, ARG774, ARG775, ARG776, ARG777, ARG778, ARG779, ARG780, ARG781, ARG782, ARG783, ARG784, ARG785, ARG786, ARG787, ARG788, ARG789, ARG790, ARG791, ARG792, ARG793, ARG794, ARG795, ARG796, ARG797, ARG798, ARG799, ARG800, ARG801, ARG802, ARG803, ARG804, ARG805, ARG806, ARG807, ARG808, ARG809, ARG810, ARG811, ARG812, ARG813, ARG814, ARG815, ARG816, ARG817, ARG818, ARG819, ARG820, ARG821, ARG822, ARG823, ARG824, ARG825, ARG826, ARG827, ARG828, ARG829, ARG830, ARG831, ARG832, ARG833, ARG834, ARG835, ARG836, ARG837, ARG838, ARG839, ARG840, ARG841, ARG842, ARG843, ARG844, ARG845, ARG846, ARG847, ARG848, ARG849, ARG850, ARG851, ARG852, ARG853, ARG854, ARG855, ARG856, ARG857, ARG858, ARG859, ARG860, ARG861, ARG862, ARG863, ARG864, ARG865, ARG866, ARG867, ARG868, ARG869, ARG870, ARG871, ARG872, ARG873, ARG874, ARG875, ARG876, ARG877, ARG878, ARG879, ARG880, ARG881, ARG882, ARG883, ARG884, ARG885, ARG886, ARG887, ARG888, ARG889, ARG890, ARG891, ARG892, ARG893, ARG894, ARG895, ARG896, ARG897, ARG898, ARG899, ARG900, ARG901, ARG902, ARG903, ARG904, ARG905, ARG906, ARG907, ARG908, ARG909, ARG910, ARG911, ARG912, ARG913, ARG914, ARG915, ARG916, ARG917, ARG918, ARG919, ARG920, ARG921, ARG922, ARG923, ARG924, ARG925, ARG926, ARG927, ARG928, ARG929, ARG930, ARG931, ARG932, ARG933, ARG934, ARG935, ARG936, ARG937, ARG938, ARG939, ARG940, ARG941, ARG942, ARG943, ARG944, ARG945, ARG946, ARG947, ARG948, ARG949, ARG950, ARG951, ARG952, ARG953, ARG954, ARG955, ARG956, ARG957, ARG958, ARG959, ARG960, ARG961, ARG962, ARG963, ARG964, ARG965, ARG966, ARG967, ARG968, ARG969, ARG970, ARG971, ARG972, ARG973, ARG974, ARG975, ARG976, ARG977, ARG978, ARG979, ARG980, ARG981, ARG982, ARG983, ARG984, ARG985, ARG986, ARG987, ARG988, ARG989, ARG990, ARG991, ARG992, ARG993, ARG994, ARG995, ARG996, ARG997, ARG998, ARG999, ARG1000, ARG1001, ARG1002, ARG1003, ARG1004, ARG1005, ARG1006, ARG1007, ARG1008, ARG1009, ARG1010, ARG1011, ARG1012, ARG1013, ARG1014, ARG1015, ARG1016, ARG1017, ARG1018, ARG1019, ARG1020, ARG1021, ARG1022, ARG1023, ARG1024, ARG1025, ARG1026, ARG1027, ARG1028, ARG1029, ARG1030, ARG1031, ARG1032, ARG1033, ARG1034, ARG1035, ARG1036, ARG1037, ARG1038, ARG1039, ARG1040, ARG1041, ARG1042, ARG1043, ARG1044, ARG1045, ARG1046, ARG1047, ARG1048, ARG1049, ARG1050, ARG1051, ARG1052, ARG1053, ARG1054, ARG1055, ARG1056, ARG1057, ARG1058, ARG1059, ARG1060, ARG1061, ARG1062, ARG1063, ARG1064, ARG1065, ARG1066, ARG1067, ARG1068, ARG1069, ARG1070, ARG1071, ARG1072, ARG1073, ARG1074, ARG1075, ARG1076, ARG1077, ARG1078, ARG1079, ARG1080, ARG1081, ARG1082, ARG1083, ARG1084, ARG1085, ARG1086, ARG1087, ARG1088, ARG1089, ARG1090, ARG1091, ARG1092, ARG1093, ARG1094, ARG1095, ARG1096, ARG1097, ARG1098, ARG1099, ARG1100, ARG1101, ARG1102, ARG1103, ARG1104, ARG1105, ARG1106, ARG1107, ARG1108, ARG1109, ARG1110, ARG1111, ARG1112, ARG1113, ARG1114, ARG1115, ARG1116, ARG1117, ARG1118, ARG1119, ARG1120, ARG1121, ARG1122, ARG1123, ARG1124, ARG1125, ARG1126, ARG1127, ARG1128, ARG1129, ARG1130, ARG1131, ARG1132, ARG1133, ARG1134, ARG1135, ARG1136, ARG1137, ARG1138, ARG1139, ARG1140, ARG1141, ARG1142, ARG1143, ARG1144, ARG1145, ARG1146, ARG1147, ARG1148, ARG1149, ARG1150, ARG1151, ARG1152, ARG1153, ARG1154, ARG1155, ARG1156, ARG1157, ARG1158, ARG1159, ARG1160, ARG1161, ARG1162, ARG1163, ARG1164, ARG1165, ARG1166, ARG1167, ARG1168, ARG1169, ARG1170, ARG1171, ARG1172, ARG1173, ARG1174, ARG1175, ARG1176, ARG1177, ARG1178, ARG1179, ARG1180, ARG1181, ARG1182, ARG1183, ARG1184, ARG1185, ARG1186, ARG1187, ARG1188, ARG1189, ARG1190, ARG1191, ARG1192, ARG1193, ARG1194, ARG1195, ARG1196, ARG1197, ARG1198, ARG1199, ARG1200, ARG1201, ARG1202, ARG1203, ARG1204, ARG1205, ARG1206, ARG1207, ARG1208, ARG1209, ARG1210, ARG1211, ARG1212, ARG1213, ARG1214, ARG1215, ARG1216, ARG1217, ARG1218, ARG1219, ARG1220, ARG1221, ARG1222, ARG1223, ARG1224, ARG1225, ARG1226, ARG1227, ARG1228, ARG1229, ARG1230, ARG1231, ARG1232, ARG1233, ARG1234, ARG1235, ARG1236, ARG1237, ARG1238, ARG1239, ARG1240, ARG1241, ARG1242, ARG1243, ARG1244, ARG1245, ARG1246, ARG1247, ARG1248, ARG1249, ARG1250, ARG1251, ARG1252, ARG1253, ARG1254, ARG1255, ARG1256, ARG1257, ARG1258, ARG1259, ARG1260, ARG1261, ARG1262, ARG1263, ARG1264, ARG1265, ARG1266, ARG1267, ARG1268, ARG1269, ARG1270, ARG1271, ARG1272, ARG1273, ARG1274, ARG1275, ARG1276, ARG1277, ARG1278, ARG1279, ARG1280, ARG1281, ARG1282, ARG1283, ARG1284, ARG1285, ARG1286, ARG1287, ARG1288, ARG1289, ARG1290, ARG1291, ARG1292, ARG1293, ARG1294, ARG1295, ARG1296, ARG1297, ARG1298, ARG1299, ARG1300, ARG1301, ARG1302, ARG1303, ARG1304, ARG1305, ARG1306, ARG1307, ARG1308, ARG1309, ARG1310, ARG1311, ARG1312, ARG1313, ARG1314, ARG1315, ARG1316, ARG1317, ARG1318, ARG1319, ARG1320, ARG1321, ARG1322, ARG1323, ARG1324, ARG1325, ARG1326, ARG1327, ARG1328, ARG1329, ARG1330, ARG1331, ARG1332, ARG1333, ARG1334, ARG1335, ARG1336, ARG1337, ARG1338, ARG1339, ARG1340, ARG1341, ARG1342, ARG1343, ARG1344, ARG1345, ARG1346, ARG1347, ARG1348, ARG1349, ARG1350, ARG1351, ARG1352, ARG1353, ARG1354, ARG1355, ARG1356, ARG1357, ARG1358, ARG1359, ARG1360, ARG1361, ARG1362, ARG1363, ARG1364, ARG1365, ARG1366, ARG1367, ARG1368, ARG1369, ARG1370, ARG1371, ARG1372, ARG1373, ARG1374, ARG1375, ARG1376, ARG1377, ARG1378, ARG1379, ARG1380, ARG1381, ARG1382, ARG1383, ARG1384, ARG1385, ARG1386, ARG1387, ARG1388, ARG1389, ARG1390, ARG1391, ARG1392, ARG1393, ARG1394, ARG1395, ARG1396, ARG1397, ARG1398, ARG1399, ARG1400, ARG1401, ARG1402, ARG1403, ARG1404, ARG1405, ARG1406, ARG1407, ARG1408, ARG1409, ARG1410, ARG1411, ARG1412, ARG1413, ARG1414, ARG1415, ARG1416, ARG1417, ARG1418, ARG1419, ARG1420, ARG1421, ARG1422, ARG1423, ARG1424, ARG1425, ARG1426, ARG1427, ARG1428, ARG1429, ARG1430, ARG1431, ARG1432, ARG1433, ARG1434, ARG1435, ARG1436, ARG1437, ARG1438, ARG1439, ARG1440, ARG1441, ARG1442, ARG1443, ARG1444, ARG1445, ARG1446, ARG1447, ARG1448, ARG1449, ARG1450, ARG1451, ARG1452, ARG1453, ARG1454, ARG1455, ARG1456, ARG1457, ARG1458, ARG1459, ARG1460, ARG1461, ARG1462, ARG1463, ARG1464, ARG1465, ARG1466, ARG1467, ARG1468, ARG1469, ARG1470, ARG1471, ARG1472, ARG1473, ARG1474, ARG1475, ARG1476, ARG1477, ARG1478, ARG1479, ARG1480, ARG1481, ARG1482, ARG1483, ARG1484, ARG1485, ARG1486, ARG1487, ARG1488, ARG1489, ARG1490, ARG1491, ARG1492, ARG1493, ARG1494, ARG1495, ARG1496, ARG1497, ARG1498, ARG1499, ARG1500, ARG1501, ARG1502, ARG1503, ARG1504, ARG1505, ARG1506, ARG1507, ARG1508, ARG1509, ARG1510, ARG1511, ARG1512, ARG1513, ARG1514, ARG1515, ARG1516, ARG1517, ARG1518, ARG1519, ARG1520, ARG1521, ARG1522, ARG1523, ARG1524, ARG1525, ARG1526, ARG1527, ARG1528, ARG1529, ARG1530, ARG1531, ARG1532, ARG1533, ARG1534, ARG1535, ARG1536, ARG1537, ARG1538, ARG1539, ARG1540, ARG1541, ARG1542, ARG1543, ARG1544, ARG1545, ARG1546, ARG1547, ARG1548, ARG1549, ARG1550, ARG1551, ARG1552, ARG1553, ARG1554, ARG1555, ARG1556, ARG1557, ARG1558, ARG1559, ARG1560, ARG1561, ARG1562, ARG1563, ARG1564, ARG1565, ARG1566, ARG1567, ARG1568, ARG1569, ARG1570, ARG1571, ARG1572, ARG1573, ARG1574, ARG1575, ARG1576, ARG1577, ARG1578, ARG1579, ARG1580, ARG1581, ARG1582, ARG1583, ARG1584, ARG1585, ARG1586, ARG1587, ARG1588, ARG1589, ARG1590, ARG1591, ARG1592, ARG1593, ARG1594, ARG1595, ARG1596, ARG1597, ARG1598, ARG1599, ARG1600, ARG1601, ARG1602, ARG1603, ARG1604, ARG1605, ARG1606, ARG1607, ARG1608, ARG1609, ARG1610, ARG1611, ARG1612, ARG1613, ARG1614, ARG1615, ARG1616, ARG1617, ARG1618, ARG1619, ARG1620, ARG1621, ARG1622, ARG1623, ARG1624, ARG1625, ARG1626, ARG1627, ARG1628, ARG1629, ARG1630, ARG1631, ARG1632, ARG1633, ARG1634, ARG1635, ARG1636, ARG1637, ARG1638, ARG1639, ARG1640, ARG1641, ARG1642, ARG1643, ARG1644, ARG1645, ARG1646, ARG1647, ARG1648, ARG1649, ARG1650, ARG1651, ARG1652, ARG1653, ARG1654, ARG1655, ARG1656, ARG1657, ARG1658, ARG1659, ARG1660, ARG1661, ARG1662, ARG1663, ARG1664, ARG1665, ARG1666, ARG1667, ARG1668, ARG1669, ARG1670, ARG1671, ARG1672, ARG1673, ARG1674, ARG1675, ARG1676, ARG1677, ARG1678, ARG1679, ARG1680, ARG1681, ARG1682, ARG1683, ARG1684, ARG1685, ARG1686, ARG1687, ARG1688, ARG1689, ARG1690, ARG1691, ARG1692, ARG1693, ARG1694, ARG1695, ARG1696, ARG1697, ARG1698, ARG1699, ARG1700, ARG1701, ARG1702, ARG1703, ARG1704, ARG1705, ARG1706, ARG1707, ARG1708, ARG1709, ARG1710, ARG1711, ARG1712, ARG1713, ARG1714, ARG1715, ARG1716, ARG1717, ARG1718, ARG1719, ARG1720, ARG1721, ARG1722, ARG1723, ARG1724, ARG1725, ARG1726, ARG1727, ARG1728, ARG1729, ARG1730, ARG1731, ARG1732, ARG1733, ARG1734, ARG1735, ARG1736, ARG1737, ARG1738, ARG1739, ARG1740, ARG1741, ARG1742, ARG1743, ARG1744, ARG1745, ARG1746, ARG1747, ARG1748, ARG1749, ARG1750, ARG1751, ARG1752, ARG1753, ARG1754, ARG1755, ARG1756, ARG1757, ARG1758, ARG1759, ARG1760, ARG1761, ARG1762, ARG1763, ARG1764, ARG1765, ARG1766, ARG1767, ARG1768, ARG1769, ARG1770, ARG1771, ARG1772, ARG1773, ARG1774, ARG1775, ARG1776, ARG1777, ARG1778, ARG1779, ARG1780, ARG1781, ARG1782, ARG1783, ARG1784, ARG1785, ARG1786, ARG1787, ARG1788, ARG1789, ARG1790, ARG1791, ARG1792, ARG1793, ARG1794, ARG1795, ARG1796, ARG1797, ARG1798, ARG1799, ARG1800, ARG1801, ARG1802, ARG1803, ARG1804, ARG1805, ARG1806, ARG1807, ARG1808, ARG1809, ARG1810, ARG1811, ARG1812, ARG1813, ARG1814, ARG1815, ARG1816, ARG1817, ARG1818, ARG1819, ARG1820, ARG1821, ARG1822, ARG1823, ARG1824, ARG1825, ARG1826, ARG1827, ARG1828, ARG1829, ARG1830, ARG1831, ARG1832, ARG1833, ARG1834, ARG1835, ARG1836, ARG1837, ARG1838, ARG1839, ARG1840, ARG1841, ARG1842, ARG1843, ARG1844, ARG1845, ARG1846, ARG1847, ARG1848, ARG1849, ARG1850, ARG1851, ARG1852, ARG1853, ARG1854, ARG1855, ARG1856, ARG1857, ARG1858, ARG1859, ARG1860, ARG1861, ARG1862, ARG1863, ARG1864, ARG1865, ARG1866, ARG1867, ARG1868, ARG1869, ARG1870, ARG1871, ARG1872, ARG1873, ARG1874, ARG1875, ARG1876, ARG1877, ARG1878, ARG1879, ARG1880, ARG1881, ARG1882, ARG1883
--------	--

4.1 실험 결과

Attention이 의미역 결정에 도움을 주는지 알아보기 위해 Bi-LSTM모델과 비교 하였다. 다음 표 3는 Bi-LSTM 모델, Bi-LSTM CRF 모델, LN Bi-LSTM CRF 모델, Attention 모델의 결과를 보여준다. 표 3에서 제시하고 있는 성능은 의미역 결정의 서술어 인식 및 분류(PIC)와 논항 인식 및 분류(AIC)중 논항 인식 및 분류에 대한 성능으로 Precision과 Recall의 조화평균으로 F1을 사용 하였고 실험 결과는 파라미터를 동일하게 하여 5번 실험한 결과의 평균으로 제시하였다.

Table 3. 실험 결과 (AIC, F1)

	Dev	Test
Bi-LSTM Model	78.77%	76.92%
Bi-LSTM CRF Model	80.14%	77.57%
LN Bi-LSTM CRF Model[1]	80.56%	78.10%
Attention Model	78.78%	76.57%

LN: Layer Normalization

Attention 기반 model이 Precision은 82.93%로 높게 나오는 반면 Recall이 71.76%로 많이 떨어지는 결과가 나타나 1st stage에서 root임베딩을 제외하고 모든 의미역을 Attention mechanism을 통해 단어에 assign하고 2nd stage에서 해당 단어가 의미역을 갖는지 갖지 않는지에 대해 binary detect를 하여 filtering을 하는 2nd stage Attention model을 이용해 binary detect의 성능을 100%로 가정하였을 때의 실험을 진행하였다. 다음 표 4는 binary detect의 성능을 100%로 가정하였을 때 2nd stage Attention model의 결과를 나타낸다.

Table 4. 2nd stage model 실험 결과 (AIC, F1)

	Dev	Test
2 nd stage Model	90.19%	88.24%

5. 결론

본 논문에서는 Attention 기반 모델을 이용해 Attention mechanism이 의미역 결정 성능 향상에 도움이 되는지 실험 하였다. 실험 결과, 제안 방법을 이용한 Attention 기반 모델이 기존 연구 결과보다 Precision은 높게 나타났지만 Recall이 많이 떨어져 F1 성능이 더 떨어지는 결과가 나타났다.

향후 연구로는 표4에 제시 된 2nd stage Attention model을 2nd stage에서 정답을 사용하지 않고 post filtering을 하였을 때의 성능을 평가 하고자 한다.

6. 참고문헌

- [1] 박광현, 나승훈, Layer Normalized LSTM CRF를 이용한 한국어 의미역 결정, 2017, HCLT
- [2] 배장성, 이창기, Stacked Bidirectional LSTM-CRFs를 이용한 한국어 의미역 결정, 2017, KIISE
- [3] 배장성, 이창기, 김현기, Highway BiLSTM-CRFs 모델을 이용한 한국어 의미역 결정, 2017, HCLT

- [4] Michael Roth, Mirella Lapata, Neural Semantic Role Labeling with Dependency Path Embeddings, 2016, ACL
- [5] 나승훈, 이건일, 신종훈, 김강일, 순환 컨트롤러를 이용한 Stack LSTM기반 한국어 의존 파싱, 2016, KCC
- [6] 민진우, 나승훈, Sequence to Sequence 모델을 이용한 한국어 의존 파싱, 2016, KCC
- [7] 민진우, 나승훈, 전이 기반 순환 유닛을 이용한 SyntaxNet 기반 한국어 의존 파싱, 2017, KCC
- [8] 나승훈, 이건일, 신종훈, 김강일, Deep Biaffine Attention 을 이용한 한국어 의존 파싱, 2017, KCC
- [9] 민진우, 나승훈, Input Attention 기반 LSTM-CNN 모델을 이용한 Relation Classification, 2017, KCC
- [10] Diego Marcheggiani, Ivan Titov, Encoding Sentences with Graph Convolutional Networks for Semantic Role Labeling, 2017, EMNLP
- [11] 박천음, 황현선, 이창기, 김현기, 멀티 레이어 포인터 네트워크를 이용한 한국어 의존 구문 분석, 2017, HCLT
- [12] Minh-Thang Luong, Hieu Pham, Christopher D.Manning, Effective Approaches to Attention-based Neural Machine Translation, 2015, EMNLP
- [13] Dzmitry Bahdanau, Kyunghyun Cho, Yoshua Bengio, Neural Machine Translation by Jointly Learning to Align and Translate, 2015, ICLR
- [14] Luheng He, Kenton Lee, Mike Lewis, Luke Zettlemoyer, Deep Semantic Role Labeling: What Works and What's Next, 2017, ACL
- [15] Feng Qian, Lei Sha, Baobao Chang, Lu-chen Liu, Ming Zhang, Syntax Aware LSTM Model for Chinese Semantic Role Labeling, 2017, EMNLP
- [16] 나승훈, 민진우, 문자 기반 LSTM CRF를 이용한 개체명 인식, 2016, KCC
- [17] 민진우, 오효정, 나승훈, 식품 도메인 개체명 인식을 위한 문자 기반 LSTM CRF, 2016, KCC
- [18] Jiang Guo, Wanxiang Che, Haifeng Wang, Ting Liu, Jun Xu, A Unified Architecture for Semantic Role Labeling and Relation Classification, 2016, COLING