**[Jae Hyung Ko](https://www.facebook.com/lookatitasitis?fref=gs&hc_ref=ARRCynwoMvrmUBcHBJJFSSeYTxHM445YrrMJJJ4pBADQg0V7n2_csiwF9p-q1EWoa3c&dti=255834461424286&hc_location=group)**님이 [링크](https://www.facebook.com/groups/TensorFlowKR/permalink/531360767204986/)를 공유했습니다.

[9월 12일](https://www.facebook.com/groups/TensorFlowKR/permalink/531360767204986/)

지난 주말에 풀 타임으로 강화학습을 공부해 보았는데요.

몇 개월 전에 김성훈 교수님 강좌 듣고 미루어 두었다가 맘 먹고 달려보았습니다.

그 경험이 도움이 될까 하여 제 나름대로 자료 보는 순서를 정리해서 글 올려봅니다.

1. 역시나 ‘모두를 위한 RL’ 강좌를 먼저 듣는 게 최선인 듯 합니다. 저는 기억을 되살리기 위해 Lecture7:DQN을 다시 들었습니다.  
[https://www.youtube.com/playlist…](https://www.youtube.com/playlist?list=PLlMkM4tgfjnKsCWav-Z2F-MMFRx-2gMGG)

2. 그 다음에 Naver Engineering에 올라와 있는 곽동현 님의 'Introduction of Deep Reinforcement Learning'을 들었습니다. 전반적인 설명이 매우 도움이 많이 됩니다.

여기서 살사와 큐러닝 차이에 대한 부분이 나오는데, 이 부분이 바로 이해가 잘 되지 않으면 이웅원 님의 ‘파이썬과 케라스로 배우는 강화학습’([http://www.yes24.com/24/goods/44136413](https://l.facebook.com/l.php?u=http%3A%2F%2Fwww.yes24.com%2F24%2Fgoods%2F44136413&h=ATO53yI2AteJkndOP2AcXsJD7IKJqBdukrT2CKROES_AqPU7QFOZkXVgrmpngx1WxlKPTaGmwltFR3Fo9mSfDL71EbBdhsZ7F0S2OSdy10L8lkyoxGuk4ECM3cZuL_Wrp5P1-ZhdcvBtdwCrnQV2M_k4GLNEMBe0TuGU7ssf97Wp9FLHeDwhhZTaeBDhLFyw6UDQc3XmjpNhr-DmKI-dsBfTCbuQOtLJN7k5Vl_B_ja1jB1QXeiBuPfn5JT7DrR5BivQVPBOt-ig45ZoLROLNA)) 책에 보면 쉽게 잘 설명이 되어 있습니다.

3. 여기까지 듣고 DQN에 대한 PR12 논문 읽기 동영상을 들으면서 정리를 했습니다.

PR-005: Playing Atari with Deep Reinforcement Learning (NIPS 2013 Deep Learning Workshop)

[https://www.youtube.com/watch…](https://l.facebook.com/l.php?u=https%3A%2F%2Fwww.youtube.com%2Fwatch%3Fv%3DV7_cNTfm2i8%26index%3D6%26t%3D362s%26list%3DPLlMkM4tgfjnJhhd4wn5aj8fVTYJwIpWkS&h=ATOifPR0H4zO8tzQil6eduy875mPntAGbAzU6iFH4jP47kj-5pxm6KGRyAYP9ApKeLxpbALc4xx3rYiXn7BQ3nkZxPVUPPcPnUThL-N9mNcAHch7QyuRC_Kvop4XhwTe9o7dTUNLCFsWKiLvB6eJ7AM_QPBrHHZIKgGGtfR_aLrPNPFDh20huidB4QIA2O5DtACI_lmRblnKNcsfueEWHEzZFsDKHkgHsu_wKHfP3QHd6RS49hlgb0YWJqbRDxkEvQvhiTytgwq7oYe3J6ZbfQ)

4. 글을 읽으면서 다시 정리를 위해 카카오 AI 리포트를 읽었습니다.  
[카카오AI리포트]알파고를 탄생시킨 강화학습의 비밀 1편입니다.  
[https://brunch.co.kr/@kakao-it/73](https://l.facebook.com/l.php?u=https%3A%2F%2Fbrunch.co.kr%2F%40kakao-it%2F73&h=ATMN3BWyvIEJVQp3zty7xpr-G0cVS_00-Vr4jyFXoHWixE_Ft8DbtX3krAc-iIpKeIsl5v1cjaI_JzUZNguJffU72TF1_8Hih5jvgYNfiZVRBLlZ69DoU12E6AjA5j0nS7Lzd9F7MwShkGdltj7GWvK1n200SMEQ4uuX4jr1NscYKtMTeLnWYFTulaT6lNaW602JVahr9qT26SYNdHdZH90eQauWxQbyygA0L2LgC5NQcRqrt3gf8D4EmgK3tvExeyf1Tbvo7Tq6DyBrKB8QoA)

5. 이제 한 걸음 더 나아가서 Naver Engineering에 올라와 있는 이웅원 님의 'RLCode와 A3C 쉽고 깊게 이해하기' 강좌를 들었습니다.  
[http://tv.naver.com/v/2022271](https://l.facebook.com/l.php?u=http%3A%2F%2Ftv.naver.com%2Fv%2F2022271&h=ATPyUp-sKm3aqA7Lk-QCBK7uLlqCuOM-0pSiYd2zY_W4V8uD_uAPHiazyyD7OI6CMCHswCmm7eQ31CF2P9jkcuwrmPiM5ytea_8Baa7edVYEzRGzqNV6VAIN5wz2Ff7UNXsqCHmwFp5mwCPsqC-lBqAgrlXuJSf67IPwkhZitVLuOVsdhjxAo4RGwyeZyzrH3ioyYH-Xpt9DF7k3qHOVVWS3YxItm8yg2iHs67A9QyTZy4qUHUTfdCwZSQn9RFRUcNssCPLAyxRB6RiW9JwiGw)

6. 위 5번과 관련되어 있는 PR12에서 모두연 김승일 소장님이 리뷰하신 동영상을 들었습니다.

PR-019: Continuous Control with Deep Reinforcement Learning

[https://www.youtube.com/watch…](https://l.facebook.com/l.php?u=https%3A%2F%2Fwww.youtube.com%2Fwatch%3Fv%3Dh2WSVBAC1t4%26list%3DPLlMkM4tgfjnJhhd4wn5aj8fVTYJwIpWkS%26index%3D20%26t%3D17s&h=ATNvkWyXUifNrwsoTq3KNwCTbMNpOuUqSvnNHoLBuPte4rWdtaFp87E_VNamfvLo47aQqyN2kVLUDMQqDu4dvcwwH7MeSza2u6_Lyz_yrN1KA0ZgaOZXKh0q54V94rWcKU-P4IS99WhoeEfJnmXZc4loP6y9oYllZXWr2M6zP2WD5yBiwu4IaU6QIFHZ16sC57G1djgc9RYz35cVgYsXoO2f4tsJCKcdVTtFCAOIuTPIFnOuAlFrZmKrsb3ubDo2MeUj0M8bpjDWwVG5djfN7A)

7. 이제 카카오 AI 리포트 2편을 읽어 보았습니다.

[카카오AI리포트]알파고를 만든 강화학습의 비밀 2편입니다.  
[https://brunch.co.kr/@kakao-it/102](https://l.facebook.com/l.php?u=https%3A%2F%2Fbrunch.co.kr%2F%40kakao-it%2F102&h=ATPWHNROcC9ux1DxXIlKqhv-ECrlE8MhY3cLlCJSvbeae3SYjjeH-BPozBrAaj3iyjSKql7dwKo5zI-no3y5n8qhyE0ubzWEUIXrZvNlmnL90uxQnyiBsyvv78jEW2RkhlBUFBXJeC1D2WUbWOViDGhIqv_45XJG-X-PZ0KC_IAuvy8HB9pqCy6pk80RtkvigYbqNFSPnAw07lxPWg-waXlqbSbeM_uRqVPAjIcbKjmo8A0DmcEdLaZgh1Sgc-k79J2caIZS8LSv2gzmpIhjxQ)

2편에는 알파고 얘기가 많이 나오면서 몬테카를로 트리탐색에 대한 얘기도 많습니다. 보기에 어려움을 느끼면, 아래 2개 강좌를 듣고 난 이후에 보면 훨씬 나을 듯 합니다.

(1) 충북대 이건명 교수님의 '탐색과 최적화1' 부분만 일단 듣습니다.  
[http://www.kocw.net/home/cview.do?lid=79a36e94d86a2ddc](https://l.facebook.com/l.php?u=http%3A%2F%2Fwww.kocw.net%2Fhome%2Fcview.do%3Flid%3D79a36e94d86a2ddc&h=ATP1ca7RnWYOtqHFNC9JJsmQHhYhIaYzlbh7Vmu7EXHH_8hbSov-cbMwoVKtXqaTTSWtFtN6ZzcR6IKYSITlen7jLTL4xQqomNFwMfdDf24HEyLHkZETbT8_Ag34oJ-1paNMs7fwsRYBZpQ5zz7RsRG6U0eLyPjsIylVmaihenrgOXqAyL7E1VfUSmMS1y7dnx8CJJdjETa2QKWCF6uQuC5sjJRDuCmtGxTrWraYKKkxpIoembotfWys0vksz9L0P_mr6-LbEnJxIT8NTKQKng)

(2) 그 다음에는 유튜브에서 '알파고는 어떻게 바둑을 두나 (충북대학교 소프트웨어학과 이건명 교수)' 영상을 듣습니다.

[https://www.youtube.com/watch…](https://l.facebook.com/l.php?u=https%3A%2F%2Fwww.youtube.com%2Fwatch%3Fv%3D94IX97APSLs%26t%3D2759s%26list%3DPL5_lXmbN9r0v4ze6pp2a0UEN69FOSnIIz%26index%3D4&h=ATPjVd7mtFFzIAI-wbwrWALGCDdLCrNHRjfaaS6IGhmr2vAMxj6qmLK4n_XxMvLtd_axmaQIgefZW4pnvTwO0UMGOdgXLNdcfkdxyAo0R1vsQgLN9XtlBB8klgFTPmKdJyNCEFXYSlxZZT1aOxbxv8yB8ok2YHgDvs19l7vn76HQ-U1vk6coTJK_FYjOa-bx4kBvZbXME9SkVhBKX40gETQejkafkjtFlDziBT1WZ7RxT2B_vYrxY8JlqvtDOAJGKah9u8iJoNuCb1WQbnTzBw)

8. 그리고 마지막으로 Naver Engineering에 올라와 있는 김태훈 님의 '알아두면 쓸데있는 신기한 강화학습'을 들었습니다.  
[http://tv.naver.com/v/2051482](https://l.facebook.com/l.php?u=http%3A%2F%2Ftv.naver.com%2Fv%2F2051482&h=ATNXQbQp5_s9TaOTv1_8buNat83uZI8Arii8Wk19Bz_7hjr6GRbLk4AOjmVpo2zXEqiW2s--A-hQ6qslAvtOMY0g3IaeBpWpga65s8VQqlZplKA-bGPUXA_lg_8mTq4SQQn2X9mLvSBlWWxFoCkKnF7AHsVh0K7SfCdvvozVoif7ErJ72DC1QCGsnNMo3jcHvhAf4EBLJWmpUVGakuXAsnKtcmZeV0_skSFnlBdr22em1EVAQHA0RYjQjjAMr342TLTtkQqCF7ER-dlIGL_kJw)

아무래도 다음 주말에 한번 더 돌아야 할 것 같습니다.

역강화학습 내용인 것 같은데, 아래 내용은 주말에 다시 복습하면서 들어야 할 듯 합니다.

PR-029- Apprenticeship Learning via Inverse Reinforcement Learning

< 주말에 한번 전체 Full로 돌아보기에 좋지 않을까 하여 올려 봅니다. 전부 다 우리말입니다. >

< 부작용 있을 수 있습니다. 두통이 심하게 올 수 있습니다. >