TÉCNICO LISBOA

Analysis of Transformer Behaviour in Reinforcement Learning

Miguel Silva Amaral

Thesis to obtain the Master of Science Degree in

Electrical and Computer Engineering

Supervisor: Prof. Arlindo Manuel Limede de Oliveira

Examination Committee

Chairperson: Prof. Pedro Filipe Zeferino Aidos Tomás
Supervisor: Prof. Arlindo Manuel Limede de Oliveira
Members of the Committee: Prof. Alexandre José Malheiro Bernardino
Prof. Francisco António Chaves Saraiva de Melo

December 2022

Acknowledgments

S_} §J ^ °9j§ §n §\9jc §\0 q}nT0 n} §\9§ 9een°0L §_ °n}c §n @0Hnh0 qn _@e0J q}nTu 8}e_jLn me_®0_}9 °\n qe9H0L _ §}« § _j h0 9jL \0eq0L h0 °_§\ _ ®9e«9@e0 _jq«§ _j §\0 h9c_jX nT §_ §\0 _u

~0HnjLJ ^ °9j§ §n §\9jc h´ T9h_e´ Tn} §\0_} en®0 9jL «qqn}§ n®0} §\0 ´09} J °_§\n«§ §\0h ^ °n«eL \9®0 j0®0} \9L 9j nqqn}§«j_§´ §n ¼n«}_ \ 9 ^ L_L _j §\0 e9 § »®0 ´09} u

d9 §e´J §n h´ T}_0jL J §\9jc Tn} 9een°_jX h0 §n X}n° 9enjX _L0 ´n«} 0e®0 u ^ H9j§ ennc @9Hc 9§ §_ an«}j0´J °_§\n«§ T0Oe_jX q}n«L nT °\9§ °0 9HHnhqe_ \0Lu ^ cjn° §\9§ ´n«} T«§«}0 °_ee @0 @}_X\§u ^§ °n«eL @0 9 L_ 0}®_H0 §n ´n«} T}_0jL _q §n jn§ h0j§_nj ´n«} j9h0 8j§Gj_nJ ?09§}_¶ puJ ?09§}_¶ |uJ ?0}j9}LnJ G9§9}_j9J K_nXnJ [0j}_x«0 9jL ¦nh¾ u

^ LOHe9}O §\9§ §_ LnH« h0j§ _ 9j n}_X_j9e °n}c nT h´ n° j 9«§\n} _q 9jL §\9§ _§ T«e»e 9ee §\0 }0x«_}0h0j§ nT §\0 GnL0 nT GnjL«H§ 9jL WnnL p}9H§_H0 nT §\0 ª j_®0} _L9L0 L0 d_ @n9u

_®

Abstract

KOH_ _nj ¦}9j Tn}h0} rK¦s EoF _j§}nL«HO 0x«0jHO hnLOe_jX 9 9 h0§\nL Tn} 9H_O®_jX LO _}OL n«§Hnh0 _j }O_jTn}HOh0j§ e09}j_jX r|ds q}n@e0h J °_eO 9®n_L_jX Hnh hnj _ «0 §\9§ H9« 0 L_®O}XOjHO _j §_ O§§_jX E©Fu ~§9§O J 9H§_nj 9jL T«§«}O }O§«}j 9}O 9hqeOL T}nh 9 L9§9 O§ _j n}LO} §n HnjL_§_nj \n° §\0 hnLOe °_ee qe9j 9 qOH_»H §}9aOH§n}´J °_eO h9 c_jX T«§«}O q9}§ nT §\0 0x«0jHO §n h9cO 09H\ §_h0 §0q 9§§0jL nje´ §n _§ q9 § _j 9j 9«§n}}OX}O _®O T9 _nju ~nh0 nT §\0 K¦¦ @n§§e0j0Hc 9}O _L0j§_»OL _j §_ °n}cJ 9jL 9}H_§OH§«}9e H\9jXO 9}O _hqe0h0j§OL _j n}LO} §n §9@_e_¶0 \n° §_ 9eXn}_§\h q0}Tn}h @O§°OOj O9H\ Oq_ nLO nT §\0 0j®_}njh0j§u ?OH9« 0 K¦ _ e_h_§OL @´ §\0 L_®O} _§´ 9jL x«9e_§´ nT _§ L9§9 O§J @n§\ 9j n¸]qne_H´ 9jL 9j nj]qne_H´ §}9j Tn}h0} 9eXn}_§\h °O}0 LO®OenqOL 9 q}nnT]nT]HnjHOq§ nT \n° O²qen}9§_nj °n«eL 9qqe´ _j 0x«0jHO hnLOe_jX 9eXn}_§\h u S_j9ee´J 9 _T§ T}nh « _jX §n´ q}n@e0h §n °n}c_jX °_§\ }O9e]°n}eL »j9jH_9e h9}cO§ °9 _hqe0h0j§OL §n \n° §\9§ |d H9j 9e n@0 « 0L §n ne®O q}n@e0h §\9§ nHH«} _j §\0 }O9e °n}eLu N²qO}_h0j§ L0hnj §}9§0 §\9§ §\O}O _ }nnh Tn} _hq}n®0h0j§ _j H«}}Oj§ 0x«0jHO hnLOe_jX ne«§_nj J 9jL §\9§ §}9j Tn}h0} \9®O 9 _Xj_»H9j§ 9hn«j§ nT «j§9qqOL qn§0j§_9e °\0j §\O´ 9}O « 0L _jL0q0jL0j§e´ Tn} 0²qen}9§_nju

Keywords

KOOq |0_jTn}HOh0j§ d09}j_jX¡ ¦}9j Tn}h0}¡ ~0x«0jHO gnLOe_jX¡ S_j9jH_9e g9}cO§¡ g9}cn® KOH_ _nj p}nHO 0¡

®_

Resumo

KOH_ _nj ¦}9j Tn}h0} rK¦s EoF _j§}nL«¶ hnLOe9čûn LO Ox«Ljh_9 Hnhn«h9 h0§nLnenX_9 q9}9 9eH9jč9}n }O «e§9Ln LO Oa9Ln 0h q}n@e0h9 LO 9q}0jL_¶9X0h qn} }OTn}čnJ 0jx«9j§n O®_§9 q}n@e0h9 Hnh«j LO §9 ¾}O9J x«0 §_q_H9h0j§0 H9« 9h L_®O}XLjH_9 j9 9q}0jL_¶9X0h E©Fu N §9Ln J 9č80 0 }OHnhq0j 9 T«§«}9 ûn 9hn §}9Ln LO «h *L9§9 O§* LO Tn}h9 9 HnjL_H_nj9} 9 Tn}h9 Hnhn«h hnLOen qe9j0_9 «h9 HO}§9 §}9aO§G}_9J h9 H9}9jLn OHč80 T«§«}9 LO Tn}h9 9 x«0 H9L9 _j§O}®9en §0hqn}9e 9q0j9 Hnj _LO}0n 0« q9 9Ln LO Tn}h9 9«§n]}OX}0 _®9u i0 §0 §}9@9e\n ûn _LOj§_»H9Ln 9eX«j q}n@e0h9 _j0}0j§0 9n K¦0h«L9jč9 9}x«_§0§«}9_ Tn}9h _hqe0h0j§9L9 LO Tn}h9 9 0 §9@_e_¶9} 9 Tn}h9 LO Hnhn0 §0 hnLOen _j§O}9XO 9 H9L9 _j §Îjh_9 Ln q}n@e0h9u ¦9j§n«h 9eXn}_§hn *njJqne_H´* Hnhn *n¸Jqne_H´* Tn}9h LO 0j®ne®_Ln LO Tn}h9 9 q}n®9} x«0 0²qen}9čûn qnLO 0} 9qe_H9L9 9 hnLOe9čûn LO Ox«LjH_9 J }0 ne®0jLn 9e_h_§9čûn LO x«9e_L9LO 0 L_®O} _L9LO x«0 n « n LO «h L9§9 O§ H9« 9u S_j9ehOj§OJ Tn_ OTO§«9L9 «h9 §}9j _čûn LO q}n@e0h9 0²0hqe_»H9§_®n q9}9 h0}H9Ln »j9jHO_}n LO h9j0_}9 9 LOhnj §}9} x«0 9 9qe_H9@_e_L9LO LO §0 hnLOen 9 q}n@e0h9 }O9_ Ĺqn ∂®Oe 9§}9®Ĺ LO 9eXn}_§hn LO 9q}0jL_¶9X0h qn} }OTn}čn0x«0 n « n LO *§}9j Tn}h0}* Ĺ«h q9 n _Xj_»H9§_®n j9 }0 ne«čûn 0»H_0j§0 LO §0 q}n@e0h9u

Palavras Chave

8q}0jL_¶9X0h qn} |OTn}čn p}nT«jL9¡ ¦}9j Tn}h0}¡ gnLOe9čûn LO ~0x«LjH_9¡ g0}H9Ln S_j9jHO_}n¡ g9}cn® KOH_ _nj p}nHO 0¡

®__

Contents

List of Figures

Vuoo 8j´§}9L_jX |0 «e§ Tn} ¦}9j Tn}h0} ppm u u u u u u u u u u u u u u u u u u u u u u u u u u u V¢

8uo Gnj§0²§ d0jX§\I g09j |0°9}L 9jL ~§9jL9}L KO®_9§_nj 9H}n [nqq0} K9§9 0§ u u u u u UP

8u© ¦9}X0§ gnL0e_jX 9H}n [nqq0} K9§9 0§ u u u u u u u u u u u u u u u u u u u u u u u u u u u Uk

8u¨ pn _§_nj9e Nh@0LL_jXI g09j |0°9}L 9jL ~§9jL9}L KO®_9§_nj 9H}n [nqq0} K9§9 0§ u u £·

8uV W9§_jXI g09j |0°9}L 9jL ~§9jL9}L KO®_9§_nj 9H}n [nqq0} K9§9 0§ u u u u u u u u u £o

8uU Gnj§0²§ d0jX§\I g09j |0°9}L 9jL ~§9jL9}L KO®_9§_nj 9H}n ¯9ec0}©L K9§9 0§ u u u u £©

8u£ ¦9}X0§ gnL0e_jX 9H}n ¯9ec0} K9§9 0§ u u u u u u u u u u u u u u u u u u u u u u u u u u u £¨

8u¢ pn _§_nj9e Nh@0LL_jXI g09j |0°9}L 9jL ~§9jL9}L KO®_9§_nj 9H}n ¯9ec0}©L K9§9 0§ u £V

8uP W9§_jXI g09j |0°9}L 9jL ~§9jL9}L KO®_9§_nj 9H}n ¯9ec0}©L K9§9 0§ u u u u u u u u u £U

List of Tables

1

Introduction

8«§nh9§_nj \9 0h0}X0L 9 nj0 nT §\0 T9 §0 § X}n°_jX _jL« §}_0 _j }0H0j§ ´09} J _jH0 _§ 0j9@e0 @« _] j0 0 §n _jH}09 0 §\0_} 0¹H_0jH´J H«§ Hn § 9jL _hq}n®0 §\0_} 0}®_H0u ^j§0ee_X0j§ 9X0j§ j00L §n _j§0}9H§ °_§\ §\0_} 0j®_}njh0j§ nj 9 L9_e´ @9 _ 9jL °_§\ }0_jTn}H0h0j§ e09}j_jX §\0´ H9j _hq}n®0 Hnj§_j«n« e´J 9een°_jX Tn} 9 _T§ _j 0®0}´nj0¸ }n«§_j0T}nh q}nL«H§_nj H0j9}_n §n _hqe0H\n}0 e_c0X}nH0}´ \nqq_jXu

1.1 Motivation

~«§§nj 9jL ?9}§nj E¨F L}9° 9 q9}9ee0e @0§°00j }0_jTn}H0h0j§ e09}j_jX 9jL §\0 °9´ e_®_jX @0_jX e09}j §\}n«X\ q0} _ §0j§ n@ 0}®9§_nj 9jL _j§0}9H§_nj °_§\ §\0_} «}}n«jL_jX 0j®_}njh0j§u ¦_ n@ 0}®9§_nj] _j§0}9H§_nj q9_} nq0j §\0 Lnn} §n j0° °9´ nT qnjL0}9§_nj °\0j0®0} 9 j0° H\9ee0jX0 9}_0 _j °_H\ q9 § 0²q0}_0jH0 H9j @0 9qqe_0Lu |0_jTn}H0h0j§ e09}j_jX H9j @0 «hh0L «q 9 §\0 q}nH0 nT e09}j_jX H9« 9e_§´ 9jL §\0 h9qq_jX @0§°00j 9H§_nj 9jL Hnj 0x«0jH0 u

8e§\n«X\ §_ He9_h _ Hn}}0H§J §\0}0 9}0 _Xj_»H9j§ L_¸0}0jH0 @0§°00j \n° 9 e_®_jX @0_jX °n«eL e09}j 9jL \n° 9 h9H_j0 °n«eL Ln §\0 9h0u S_} § nT 9eeJ §\0 9hn«j§ nT _jTn}h9§_nj }0x«_}0L Tn} 9 h9H_j0 §n h9§H\ §\0 °9´ 9 \«h9j q0}Tn}h nj 9 q0H_»H §9 c 0²H00L §\0 j«h@0} nT q0}H0q§_nj 9 \«h9j « 0

Tn} O²§}9qne9§_nj 9jL LOH_ _nj]h9c_jXJ h9c_jX §\O §}9_j_jX nT 9j´ }O_jTn}HO h Oj§ eO9}j_jX hnLOe 9 L9§9] _j§Oj _®O q}nHO rnT§Oj }Ox«_}_jX h_ee_nj nT Oj®_}nj h Oj§ §Oq s °_H\ 9e n h9cO _§ 9 §_h O]Hnj «h_jX q}nHO §\9§ }Ox«_}O 9 e9}XO 9 h n«j§ nT Hn h q«§9§_nj 9e qn°O}J °\O}O9 9 \«h9j H9j eO9}j T}n h 9 h9ee j«h@O} nT _j§O}9H§_nj 9jL \OO} _j§«_§_nju

~OHnjL nT 9eeJ jn ´ §O h Hn h O Hen O §n §\O 9@_e_§´ nT XOjO}9e_¶9§_nj §\9§ 9 \«h9j \9 J °_§\n«§ \9®_jX H\9jXO _h qeO h Oj§OL _j§n _§ §}«H§«}OJ @OH9« O 9ee ´ §O h e9Hc §\O H9q9@_e_§´ §n XOjO}9e_¶O HnjHOq§ 9jL §\O 9@_e_§´ §n eO9}j jO° §9 c _j 9 9 h qeO]O¹H_Oj§ h9jjO} EVFu

¦\O e9Hc nT 9 h qeO O¹H_OjH´ 9e n \9 h qO} §\O °n}c nT }O O9}H\O} 9jL h9cO _§ }9§\O} L_¹H«e§ §n HnjL«H§ O²qO}_h Oj§ 9jL Hn}}OH§ h_ §9cO J O qOH_9ee´ X_®Oj \n° L_¹H«e§ _§_ §n »jL @«X _j }O_jTn}HO h Oj§ eO9}j_jX }O O9}H\ 9jL L«O §n h9j´ q}n@eO h _j\O}Oj§ §n h9H_jO eO9}j_jX EUFJ «H\ 9 }9jLn h OOL_jXJ §nH\9 §_H_§´ n} _h qe´ jO§°n}c H\n_HOu

d9 §e´J en° 9 h qeO O¹H_OjH´ h9´ q}n®_LO ne«§_nj §n §\O h9an}_§´ nT §n´ q}n@eO h §\9§ 9}O « «9ee´ §}O9§OL _j §_ 9}O9J «H\ 9 _h«e9§_nj 9jL X9 h O u ^j q_§O nT §\9§J §\O}O _ §_ee 9 _Xj_»H9j§ X9q _j \n° §_ °n«eL §}9j e9§O §n }O9e]°n}eL q}n@eO h «H\ 9 _jL« §}´ 9«§n h9§_njJ \O9e§\H9}OJ }n@n§_H 9jL §\O }9q_Le´ X}n°_jX »OeL nT 9«§njnhn« L}_®_jXJ °_H\ \9 @OOj O §_h9§OL §n \9®O 9 h9}cO§ ®9e«O nT n®O} © §}_ee_nj Lnee9} @´ ©·¨·u

^j \n}§J §\O 9@_e_§´ §n ne®O }O_jTn}HO h Oj§ eO9}j_jX q}n@eO h _j 9 \n}§ 9 h n«j§ nT §_h OJ Hn h @_jOL °_§\ _§ O®O}]_jH}O9 _jX jOOL _j }O9e]°n}eL q}n@eO h J «XXO § §\9§ 9 h qeO O¹H_OjH´ Hn«eL 9 _§ }O_jTn}HO h Oj§ eO9}j_jX _j h9c_jX §\O jO²§ @_X eO9q _§ }Ox«_}O u

1.2 Objectives

¦\O q}_h9}´ Xn9e nT §_ °n}c _ §n _j®O §_X9§O §\O H9q9@_e_§_O 9jL qn§Oj§_9e nT 9 ®9}_O§´ nT §}9j Tn} h O} 9}H_§OH§«}O Tn} ne®_jX }O_jTn}HO h Oj§ eO9}j_jX q}n@eO h °_§\ 9 e_h_§OL j«h@O} nT 9 h qeO J 9 °Oee 9 §n 9 0 \n° O9H\ 9}H_§OH§«}O 9jL _h qeO h Oj§9§_nj H9j _h q9H§ qO}Tn} h 9j HO _j e_h_§OL L9§9 Oj®_}nj h Oj§ «H\ 9 §\O KOOq g_jL Gnj§}ne ~«_§O E£F n} §\O 8§9}_ o··c E¢Fu

g n}O qOH_»H9ee´J §_ °n}c °_ee 9e n O²qen}O §\O H9q9@_e_§_O nT h nLOe]T}OO }O_jTn}HO h Oj§ eO9}j_jX 9}H_§OH§«}O °_§\n«§ qe9j j_jX 9eXn}_§\h J «§_e_¶_jX 9}H_§OH§«}O _h_e9} §n §\n O « OL _j e9}XO q}O] §}9_jOL h nLOe §\9§ 9}O »jO]§«jOL §n Ln° j §}O9 h §9 c EPJkF 9jL \n° L_¸O}Oj§ 9§§Oj§_nj h OH\9j_ h «H\ 9 h «e§_]\O9L 9§§Oj§_njJ « OL _j §\O §}9j Tn} h O} 9}H_§OH§«}O Eo·F °_ee qO}Tn} h °\Oj 9qqe_OL §n §_ §´qO nT q}n@eO hu

^j 9LL_§_njJ }O O9}H\ °_ee @O HnjL«H§OL §n _j®O §_X9§O §\O qn§Oj§_9e Tn} 9L®9j HO h Oj§ _j §\O H«}}Oj§ §9§O nT §\O 9}§u

¦\O §}9j Tn} h O} 9}H_§OH§«}O §n @O _j®O §_X9§OL 9e n \9®O 9j 9«§n]}OX}O _®O j9§«}O 9 njO nT §\O_} §9jL]n«§ TO9§«}O J °\O}O H9« 9e_§´ _ }Oq}nL«HOLJ h_h_Hc_jX §\O HnjL_§_nj nT }O_jTn}HO h Oj§ eO9}j_jX q}n@eO h 9jL q}n®_L_jX 9 °9´ §n _h h OL_9§Oe´ O²q9jL nj H«}}Oj§ °n}c Lnj O _j §\O n°_jO]|d «@ O§ nT

q}n@e0h EooFu ¦\0}0 °_ee 9e n @0 9§§0hq§ §n }0 09}H\ 9jL LO®0enq nje_j0 |d ne«§_nj °_§\ 9§§0j§_nj hnLOeu ¦_ }0 09}H\ \9 §\0 9LL0L L_¹H«e§´ nT _§ §_ee @0_jX _j 9 }9§\0} «j0²qen}0L 9}09u

1.3 Contributions

¦_ §\0 _ Tneen° }0H0j§Hnj§}_@«§_nj @´G\0j 0§u 9eEoF§\9§_j§}nL«H0§}9j Tn}h0} §n§\0}0_jTn}H0h0j§ e09}j_jX «@ 0§ nT L00q e09}j_jX q}n@e0h u ¦\0Tneen°_jXHnj§}_@«§_nj 9}0 q}0 0j§0LI

- ą ^L0j§_»H9§_nj nT 0x«0jH0 hnLOe_jX @n§§e0j0Hc _j §\0 K0H_ _nj ¦}9j Tn}h0} 9}H_§0H§«}0
- ą N®9e«9§_nj nT L_¸0}0j§ §}9j Tn}h0} _hqe0h0j§9§_nj _j n°_j0 }0_jTn}H0h0j§ e09}j_jX
- ą 8qqe_H9§_nj nT §}9j Tn}h0} nj hnLOe]T}00 nje_j0 }0_jTn}H0h0j§ e09}j_jX 9eXn}_§\h
- ą N®9e«9§_nj nT 0x«0jH0 hnLOe_jX _j §}9L_jX 0j®_}nj h0j§

1.4 Structure of the Document

¦\0 »} §H\9q§0} nT §_ §\0 _ _ §\0 _j§}nL«H§_njJ°\0}0§\0 hn§_®9§_njJ n@aOH§_®0 J 9jL §\0 §}«H§«}0 nT §\0 LnH«h0j§ 9}0 q}0 0j§0Lu

¦\0 }0h9_jL0} nT §_ LnH«h0j§_ §}«H§«}0L 9 Tneen° I

G\9q§0} © q}0 0j§ 9 e_§0}9§«}0 }0®_0° nj ros §}9j Tn}h0} 9}H_§0H§«}0 J r©s }0_jTn}H0h0j§ e09}j_jX h0§\nL 9jL r¨s §\0 §9§0]nT]§\0]9}§ Hnh@_j9§_nj nT @n§\ 9}09 9jL 9L_H« _nj nj _§ qn§0j§_9e _j §_ @nL´nT °n}cu

G\9q§0} ¨0²qe9_j §\0 q}n@e0h §n@0§9HceOL 9jL §\0§\0n}0§_H9e ne«§_nj nj ros K0H_ _nj ¦}9j Tn}h0} ®9e_L9§_nj 9jL r©s nje_j0 }0_jTn}H0h0j§ e09}j_jX §}9j Tn}h0} _hqe0h0j§9§_nj u

G\9q§0} V q}0 0j§ }0 «e§ nj ros @9 0 9}H_§0H§«}0 ®9e_L9§_njJ r©s _hqe0h0j§0L 9}H_§0H§«}9e H\9jX0 J r¨s §}9j Tn}h0} w]e09}j_jX 9jL rVs §}9j Tn}h0} qne_H´nq§_h_¶9§_nju

G\9q§0} U L_H« 0 §\0 }0 «e§ §\9§ °0}0n@§9_j0L _j §_ °n}c 9jL T«}§\0} °n}c §n @0 LO®0enq0Lu

G\9q§0} £HnjHe«L0 §_ °n}c @´L0H}_@_jX _§ h9_j 9H_0®0h0j§ 9jL @´q}nqn _jXT«§«}0 _hq}n®0] h0j§ §nH«}}0j§ °n}cu

2

Literature Review

¦_ H\9q§0} °_ee q}0 0j§ 9 e_§0}9§«}0 }O®_0° nj ros §}9j Tn}h0} 9}H_§OH§«}0 9jL r©s }0_jTn}H0h0j§ e09}j_jX h0§\nL 9 °0ee 9 9 L0§0}h_j9§_nj nT §\0_} _hq9H§ 9jL }0eO®9jH0 nj T«§«}0 |d }0 09}H\u ¦\0}0 °_ee 9e n @0 9 r¨s n®0}®_0° nT §\0 §9§0 nT §\0 9}§ 9jL \n° _§ °_ee _hq9H§ §\0 q}nqn 0L °n}cu

¦}9j Tn}h0} 9}H_§OH§«}0 °_ee @0 §\0 h9_j TnH« nT §_ @nL´ nT °n}c 9jL Tn} §\9§ }09 nj 9j 0²qe9] j9§_nj nT \n° §\0´ T«jH§_nj °_ee @0 q}0 0j§0Lu 8T§0} «jL0} §9jL_jX §\0_} @9 0 9}H_§OH§«}0J 9 0§ nT §9§0 nT §\0 9}§ _hqe0h0j§9§_nj °_ee @0 q}0 0j§0L 9HHn}L_jX §n §\0_} L_¸0}0j§ _hqn}§9jH0 _j §\0 9}09 nT L00q e09}j_jX 9jL nj §_ @nL´ nT °n}cu

|0_jTn}H0h0j§ e09}j_jX _ §\0 9}09 °\0}0 §\0 §«L_0L 9eXn}_§\h _j §_ °n}c 9}0 _j 0}§0L 9jL Tn} §\9§ }09 nj 9 @}_0T _j§}nL«H§_nj §n §\0 «@a0H§ _ q}0 0j§0LJ @0Tn}0 hn®_jX nj§n 9j n¸]qne_H´ 9eXn}_§\h _j w]e09}j_jX 9jL 9j nj]qne_H´ 9eXn}_§\h _j q}n²_h9e qne_H´ nq§_h_¶9§_nju

S_j9ee´ §\0 9qqe_H9§_nj nT §}9j Tn}h0} hnL0e _j }0_jTn}H0h0j§ e09}j_jX °_ee @0 9j9e´¶0L 9 9 °\ne0J _j °\9§ °_ee @0 §\0 §9§0 nT §\0 9}§ q}0 0j§9§_nju

2.1 Transformer Architectures

mj0 nT §\0 _ «0 °_§\ G i i 9jL n§\0} §´q0 nT j0«}9e j0§°n}c _ §\9§ §\0´ \9®0 9 »²0L j«h@0} nT _jq«§ 9jL n«§q«§ J 9 °0ee 9 9 »²0L j«h@0} nT h9qq_jX §0q u ¦_ h9c0 _§L_¹H«e§ §n q}nH0 0x«0jH0 e_c0 0j§0jH0 9jL §\9§ _ °\´ 9}H_§0H§«}0 e_c0 }0H«}}0j§ j0«}9e j0§°n}c Eo©FJ d~¦g Eo¨FJ 9jL W|ª EoVF \9®0 @00j « 0L _j q}n@e0h e_c0 j9§«}9e e9jX«9X0 q}nH0 _jXu

¦}9j Tn}h0} 9}0 9 hnL0e §\9§ « 0 0eT]9§§0j§_njJ 9 h0H\9j_ h nT Hn}}0e9§_nj @0§°00j qn _§_nj _j 9 0x«0jH0J §n @0§§0} }0q}0 0j§ q9§§0}j u ¦_ 9}H_§0H§«}0 h9c0 0²§0j _®0 « 0 nT q9}9eeOe q}nH0 _jXJ }0L«H_jX §\0 Hnhqe0²_§´ nT _§ §}9_j_jX §_h0 _Xj_»H9j§e´u ¦\0 »} § §}9j Tn}h0} hnL0e Eo·F 9H_0®0L §9§0] nT]§\0]9}§ }0 «e§ _j §}9j e9§_nj §9 c J9jL nh0 nT _§ L0}_®9§_nj \9®0 \n° j X}09§ q}nh_0 _j 9}09 «H\ 9 ®_ _nj EoUF 9jL ®_L0n q}nH0 _jX Eo£Fu

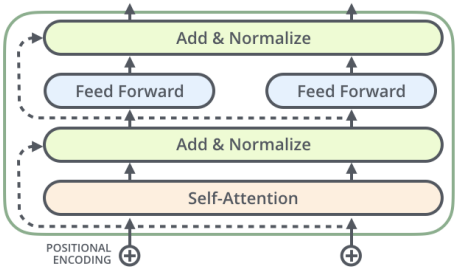
¦\0 §}9j Tn}h0} h9´ @0 qe_§ _j§n §°n q9}§ J §\0 0jHnL0} 9jL LOHnL0} §9Hc \n° j _j »X«}0 ©uo 9jL ©u© }0 q0H§_®0e´u ?n§\ 9}0 h9L0 «q nT O®0}9e _L0j§_H9e e9´0} u ¦\0 0jHnL0} §9Hc }0H0_®0 §\0 _jq«§ nT §\0 ´ §0h 9T§0} _§ \9 @00j 0h@0LL0L _j 9 ®0H§n} °_§\ §\0 qn _§_nj9e 0jHnL_jX nT 09H\ q9}§ nT §\0 0x«0jH0J °_H\ _ 9j 9}}9´ §\9§ Tneen° 9 q9§§0}j §\9§ °_ee @0 e09}j0L @´ §\0 hnL0e _j n}L0} §n cjn° §\0 qn _§_nj nT 09H\ nT _§ §nc0j n}_jq«§ u

¦_ 0H§_nj \9 §\0 q«}qn 0 nT 0²qe9_j_jX \n° §_ hnL0e @0H9h0 nj0 nT §\0 hn § qnq«e9} L00q e09}j_jX 9}H_§0H§«}0 J @´ L0 H}_@_jX §\0 hn§_®9§_nj @0_jL_§ L0®0enqh0j§ 9jL 0j«h0}9§_jX _§ @0j0»§ u

2.1.1 Base Transformer

©uouou8 NjHnL0}

¦\0 0jHnL0} e9´0} ¦q}_h9}´ Xn9e _ §n 0jHnL0 n«}H0 0x«0jH0 9jL Hnj®0}§ §\0h _j§n §9§0 ®0H§n} J 9een°_jX }0q}0 0j§9§_nj §n @0 e09}j0L 9jL L9§9 L_h0j _nj §n @0 }0L«H0Lu S_X«}0 ©uo L0q_H§ 9j 0jHnL0} 9 9j 0²9hqe0u



S_X«}0 ©uoI NjHnL0} 9}H_§0H§«}0 r|0q}_j§0L T}nh Eo¢Fs

N9H\ 0jHnL0}_j §\0 §9Hc _ h9L0 «q nT §°n «@]e9´0} u ^j §\0 »} § nj0J §\0 _jq«§ °_ee q9 §\}n«X\ 9 0eT]9§§0j§_nj e9´0}J °_H\ 9een° §\0 0jHnL0} §n ennc 9§ n§\0} §nc0j °_e0 0jHnL_jX 09H\ q0H_»H §nc0j u ¦\0 n«§q«§ nT §_ »} § «@]e9´0} °_ee @0 T0L §n 9 T00LTn}°9}L j0«}9e j0§°n}c §\9§ _ §\0 9h0 _j 9ee

qn _§_nj u Sneen°_jX §\9§J _j °\9§_ cjn°j 9 9 }0 _L«9e HnjjOH§_njJ §\0 n«§q«§ nT O9H\ «@]e9´0} °_ee @0 «h h0L °_§\ _§ _jq«§ §n 9een° Tn} LOOq0} j0§°n}c °_§\n«§ ®9j_ _jX X}9L_0j§ EoPFu

~0eT]9§§0j§_nj \9 @00j }0»j0L ®_9 9 §OH\j_x«0 cjn°j 9 h«e§_]\09L 9§§0j§_njJ °_H\ °_ee _hq}n®0 §\0 hnL0e _j §°n °9´ I

ou ?´ 0²q9jL_jX §\0 9@_e_§´ nT §\0 hnL0e §n TnH« nj L_¸0}0j§ qn _§_nj

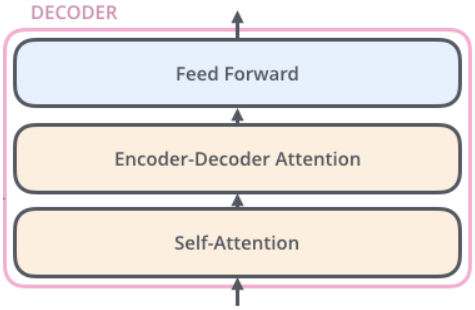
©u ?´ q}n®_L_jX j«h0}n« }0q}0 0j§9§_nj q9H0 §n §\0 9§§0j§_nj e9´0}J 09H\ §nc0j °_ee @0 }0q}0 0j§0L _j q9}9ee0e h9j´ §_h0 J 9een°_jX h«e§_qe0 q9§§0}j §n @0 L_ Hn®0}0Lu

N 0j§_9ee´J °\9§ \9qq0j _j »X«}0 ©uo °_ee \9qq0j h«e§_qe0 §_h0 _j q9}9ee0e 9jL §\0}0 °_ee @0 h«e§_qe0 n«§q«§ Hnh_jX T}nh §\0 9§§0j§_nj e9´0}u 8ee §\0 n«§q«§ °_ee @0 HnjH9§0j9§0L 9jL h«e§_qe_0L @´ 9 °0_X\§ h9§}_² §\9§ °_ee 0}®0 9 n«§q«§u

©uouou? KOHnL0}

¦\0 L0HnL0} 9}H_§0H§«}0 L0 _Xj _ x«_§0 _h_e9} TO9§«}_jX 9 §_}L e9´0} §\9§ q0}Tn}h h«e§_]\09L 9§§0j§_nj n®0} §\0 n«§q«§ nT §\0 0jHnL0} §9Hcu ¦\0 9§§0j§_nj _j §\0 »} § L0HnL0} «@]e9´0} h9´ @0 h9 cOLJ Tn}H_jX §\0 9}}9´ §n L_ h_ T«}§\0} _jq«§ §\9j §\0 H«}}0j§ nj0u ¦_ h9c0 0j 0 @0H9« 0 Hn}}0e9§_nj \n«eL nje´ @0 Lnj0 °_§\ q}_n} cjn°e0LX0J 0 §9@e_ _jX 9j _hqn}§9j§ }0e9§_nj _q nT H9« 9e_§´u

S_X«}0 ©u© L_ qe9´ §\0 L_¸0}0jH0 T}nh 9j 0jHnL0} e9´0} §n §\0 L0HnL0} e9´0}u



S_X«}0 ©u©I KOHnL0} 9}H_§0H§«}0 r|0q}_j§0L T}nh Eo¢Fs

¦_ §9Hc Hnj§9_j 9ee nT §\0 9§§}_@«§0 nT §\0 0jHnL0} §9Hc «H\9 qn _§_nj9e 0jHnL_jX 9jL }0 _L«9e HnjjOH§_nj u

©uououG 8§§0j§_nj g OH\9j_ h

¦\0 9§§0j§_nj T«jH§_nj _ Hnhq«§0L @´ Hnj §}«H§_jX §\}00 }0q}0 0j§9§_nj nT §\0 _jq«§ §\9§ 9}0 H9ee0L §\0 w«0}´J b0´J 9jL -9e«0 ®0H§n} rQJ KJ 9jL Vsu ¦\0 0 9}0 H}09§0L @´ h«e§_qe´_jX §\0 0h@0LL0L _jq«§ @´ §\}00 h9§}_H0 §\9§ 9}0 §}9_j0LJ Tn} 0²9hqe0 §\}00 e_j09} e9´0} u

¦ \0 9§§0j§_nj T«jH§_nj °_ee @0 H9eH«e9§0L _j §\0 Tneen°_jX °9´J °\0}0 d_k _ §\0 L_h0j _nj nT §\0 c0´ ®OH§n}

$$Attention(Q; K; V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V:$$ r©uos

8 q}0®_n« e´ _jL_H9§0LJ h«e§_]\09L 9§§0j§_nj 9een° Tn} h«e§_qe0 }0q}0 0j§9§_nj «@] q9H0 u ¦ _ _ 9H_0®0L @´ §\0 hnL0e \9®_jX h9j´ 9§§0j§_nj \09L _j 09H\ e9´0} §n °_ee }0q}0 0j§ 09H\ «@] q9H0u g«e§_] \09L 9§§0j§_nj _ 9 0§ nT 9§§0j§_nj Hnhq«§9§_nj HnjH9§0j9§0L 9jL h«e§_qe_0L @´ 9jn§\0} e09}j0L °0_X\§ h9§}_²u

g«e§_]\09L 9§§0j§_nj 0}®0 0®0}9e q«}qn 0 §\}n«X\n«§ §\0 9}H_§0H§«}0 L0q0jL_jX °\0}0 §\0 QJ KJ 9jL V ®OH§n} n}_X_j9§0 9jL °\0}0 §\0´ 9}0 q}nH0 0Lu ¦ \0 n}_X_j9e 9}§_HeO 0®0j §9§0 §\9§ 9§§0j§_nj H9j }0q}0 0j§ @0\9®_n} e_c0 ´j§9H§_H 9jL 0h9j§_H §}«H§«}0 _j 0j§0jH0 u

©uououK ¦ \0 T«ee 9}H_§0H§«}0

¦ \0 T«jH§_nj_jX nT §\0 §}9j Tn}h0} 9}H_§0H§«}0J 9 q}0 0j§0L _j »X«}0 ©u¨J @0X_j °_§\ §\0 q}nH0 _jX nT §\0 _jq«§ 0x«0jH0 _j §\0 0jHnL0} 9jL L0HnL0} §9Hc u N9H\ 0jHnL0}¸ n«§q«§ _ «qqe_0L 9 §\0 9§§0j§_nj ®0H§n} K 9jL V §\9§ °_ee @0 « 0L _j 0jHnL0}]L0HnL0} 9§§0j§_nj §\9§ °_ee \0eq §\0 L0HnL0} TnH« nj §\0 9qq}nq}_9§0 }0X_nj u

¦ _ q}nH0 _ }0q09§0L «j§_e 9 qOH_»H ´h@ne _Xj9e §\9§ §\0 0jHnL0} \9 Hnhqe0§0L _§ n«§q«§J 9§ °_H\ qn_j§ _§ _ q9 0L §n §\0 @n§§nh L0HnL0}J °_H\ 9XX}0X9§0 §\0 }0 «e§ 9jL 0h@0L _§ n°j qn _§_nj9e 0jHnL_jXu

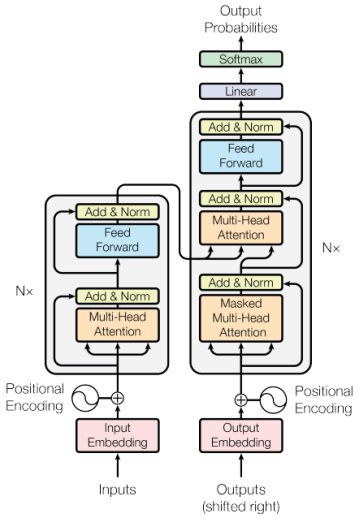
¦ \0 ´ §0h HnjHe«L0 @´ §9c_jX _j §\0 9}}9´ n«§q«§ nT §\0 e9 § L0HnL0} 9jL §}9j e9§_jX _§ nj0 §nc0j 9§ 9 §_h0 _j§n §\0 L0 _}0L §´q0 nT _jTn}h9§_nju ¦ _ _ 9HHnhqe_ \0L @´ q9 _jX §\0 9}}9´ §\}n«X\ 9 T«ee´ Hnjj0H§0L e9´0} 9jL 9 ~nT§h9² 9H§_®9§_nj T«jH§_nj §\9§ °_ee Hnj®0}§ 09H\ qn _@e0 §nc0j _j§n 9 q}n@9@_e_§´ 9jL 0eOH§_jX §\0 n«§q«§ @9 0L nj §\0 _X\0 § q}n@9@_e_§´u

2.1.2 Vision Transformer

^§ °9 jn§ enjX @0Tn}0 §\0 «j}09e_¶0L qn§0j§_9e nT §\0 §}9j Tn}h0} 9}H_§0H§«}0 °9 §}0§H\0L _j§n n§\0} L00q e09}j_jX »0eL «H\ 9 Hnhq«§0} ®_ _nju ¦ \0 -_ _nj ¦}9j Tn}h0} r-_¦s EoUF q}n®0L §n @0 9 Hnhq0§_§n} §n G i i 9}H_§0H§«}0 °_e0 « _jX e0 Hnhq«§9§_nj9e }0 n«}H0 u

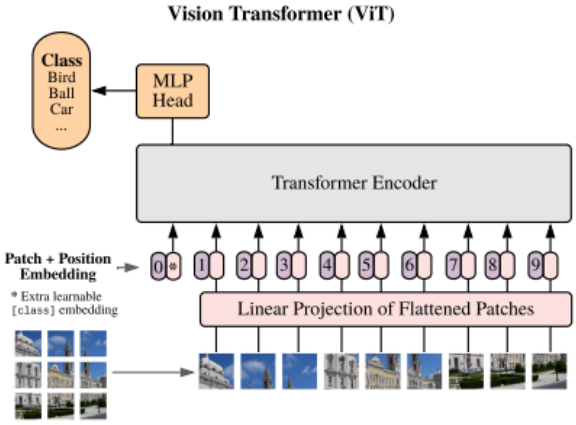
8 h9an} q}n@e0h °_§\ §}9j Tn}h0} _j ®_ _nj _ §\0 e9Hc nT _jL«H§_®0 @_9 0 Tn«jL _j n§\0} 9}H_§0H§«}0 J °_H\ °n«eL T9H_e_§9§0 Tn} §}9_j_jX °_§\ en° 9hn«j§ nT L9§9u ¦ \0 « 9X0 nT 9 ®9 § 9hn«j§ nT L9§9 §n §}9_j §\0 j0§°n}c 9jL §\0j »j0]§«j0 _§ §n 9 h9ee0} §9 c h_§_X9§0 §\0 9@ 0jH0 nT hnL0e 9 «hq§_nj u

-_¦ L_¸0} T}nh §\0 n}_X_j9e §}9j Tn}h0} L0 _Xj _j §\9§ _j §09L nT °n}L J §\0 _jq«§ _ ¼9§§0j0L q9§H\0 nT 9j _h9X0 §nX0§\0} °_§\ _§ qn _§_nj9e 0h@0LL_jXu K0 q_§0 _§ XnnL }0 «e§ °_§\ e9}X0 9hn«j§ nT L9§9J °n}c j00L §n @0 Lnj0 _j n}L0} §n }0L«H0 n®0}»§§_jX _j h9ee0} L9§9 0§ u



S_X«}O ©u¨I ?9 0 ¦}9j Tn}h0} 9}H_§OH§«}O r|Oq}_j§OL T}nh Eo·Fs

S_X«}O ©uV \n° 9 }Oq}O Oj§9§_nj nT §\O -_¦ 9}H_§OH§«}Ou



S_X«}O ©uVI -_¦ 9}H_§OH§«}O r|Oq}_j§OL T}nh EoUFs

^jL«H§_®O @_9 0 H9j 9e n @O _j§}nL«HOL _j §\O O9}e´ §9XO nT §\O 9}H_§OH§«}OJ 9 }OHOj§ 9L®9jHO _j§}nL«HOL @´ Gnj-_¦ EokF LOhnj §}9§Ou

~°_j ¦}9j Tn}h0} E©·F_j§}nL«HO _O}9}H_H9e TO9§«}O h9q §n }OL«HO §\O _¶O nT TO9§«}O T}nh njOe9´O} §n 9jn§\O}J 9jL _T§OL °_jLn° 9§§Oj§_nj §n }OL«HO }O §}_H§_nj _j 9§§Oj§_nj H9eH«e9§_nj T}nh q9§H\O §n §\O _h9XO 9 9 °\neOu

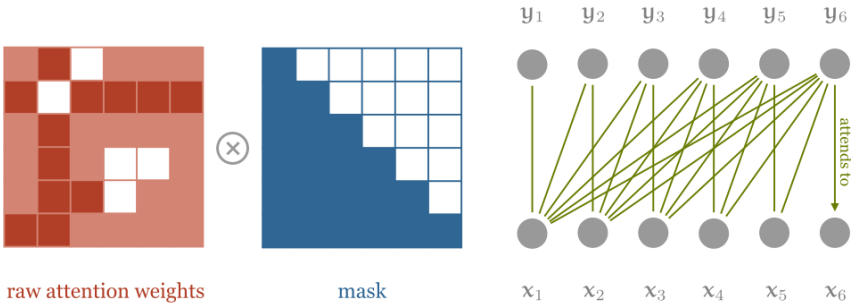
2.1.3 GPT-2

¦ \0 Wp ¦]© rW0j0}9§_®0 p}0] ¦ }9_j_jXs 9}H_§0H§«}0 E©oF _j jn®9§0 @´ @}_jX_jX 9 §9§0]nT]§\0]9}§ ne«§_nj §n j9§«}9e e9jX«9X0 q}nH0 _jX §\9§ _ @9 0L nj 9 §}9j Tn}h0} hnL0e @«_e§ « _jX nje´ 9 L0HnL0} §9Hc §\9§ Ln0 jn§ T09§«}0 0jHnL0}]L0HnL0} 9§§0j§_nju

¦ \0 q}_h9}´ L_ §_jH§_nj °9 §\0 « 0 nT 9«§n]}0X}0 _®0 h9 c_jX §n q}n®_L0 §\0 0j 0 nT H9« 9e_§´J ° _H\ nje´ h9q 9 }0e9§_nj _q @0§°00j §\0 §nc0j 9§ §\0 H«}}0j§ hnh0j§ 9jL _§ q}_n} 0x«0jH0u

¦ _ _ 9H_0®0L @´ 9qqe´_jX 9 h9 c Hnhqn 0L @´ 9 _X\0} §}_9jX«e9} h9§}_² °_§\ §\0 ®9e«0 *1* §n §\0 9§§0j§_nj h9§}_²J §\9§ h9c0 §\0 ®9e«0 9@n®0 §\0 L_9Xnj9e nT §\0 h9§}_² · 9T§0} §\0 9qqe_H9§_nj nT §\0 ~nT§h9² 9H§_®9§_nj T«jH§_nju

-_ «9ee´ §_ _ }0q}0 0j§0L @´ \9®_jX 09H\ §nc0j nT §\0 hnL0e nje´ ennc 9§ q}0®_n« §nc0j J 9 H9j @0 00j _j »X«}0 ©uUu



S_X«}0 ©uUI G9« 9e 8§§0j§_nj r | 0q}_j§0L T}nh E©©Fs

~_jH0 §_ hnL0e °9 q}0]§}9_j0L nj \«X0 9hn«j§ nT L9§9 §n q}0L_H§ §\0 Tneen°_jX §nc0jJ_§ °9 H9q9@e0 nT ne®_jXn§\0}§9 c @´ @0_jX»j0]§«j0L §n §\0hJ 9een°_jX Tn} hn}0 9hqe0 0¹H_0j§_hqe0h0j§9§_nj «H\ 9 §\0 H«}}0j§ §9§0]nT]§\0]9}§ 9qqe_H9§_nj nT §}9j Tn}h0} _j }0_jTn}H0h0j§ e09}j_jX 9jL nj _§ e9jX«9X0 hnL0e_jX q}0L0H0 n}J ?N | ¦ EkFu

S_j0]§«j_jX_ Lnj0@´ §}9_j_jX 9 He9 _»0} nj §\0e9 § _LL0j §9§0 nT 9 hnL0e §\9§ « 0 e9}X0 0x«0jH0 nT §0²§ _j n}L0} §n e09}j enjX]}9jX0 L0q0jL0jH_0 u

2.1.4 BERT

Gnj§}9}´ §n Wp ¦ J ?N | ¦ EkF r?_L_}0H§_nj9e NjHnL0} | 0q}0 0j§9§_nj T}nh ¦ }9j Tn}h0} sJ ° _H\ _ jn° « 0L _j WnnXe0¸ 09}H\ 0jX_j0J _ 9j 0jHnL0} nje´ hnL0e §\9§ _ 9e n q}0]§}9_j0L _j e9}X0 9hn«j§ nT L9§9 9jL »j0]§«j0L §n qOH_»H §9 c u ^§ _ q}0]§}9_j0L @´ §}´_jX §n »jL °\9§ h9 c0L §nc0j nj 9 H0}§9_j 0x«0jH0 9}0 9jL §\0 q9}9h0§0} 9}0 9®0L Tn} T«}§\0} « 0u

~_jH0 _h_e9} q9}§ nT 9 0x«0jH0 H9j \9®0 L_¸0}0j§ h09j_jX 9HHn}L_jX §n §\0_} Hnj§0²§J ?N | ¦ « 0 9 @_L_}0H§_nj9e d~ ¦ g Tn} qn _§_nj9e 0h@0LL_jX _j §09L nT »²0L rT}n¶0j n} §}9_j9@e0s 0h@0LL_jX J ° _H\ h09j _§ Ln0 jn§ a« § HnjL_§_nj nj §\0 q9 § e_c0 Wp ¦]©J @«§ 9e n nj §\0 T«§«}0u 8j 0²9hqe0 nT §_ _

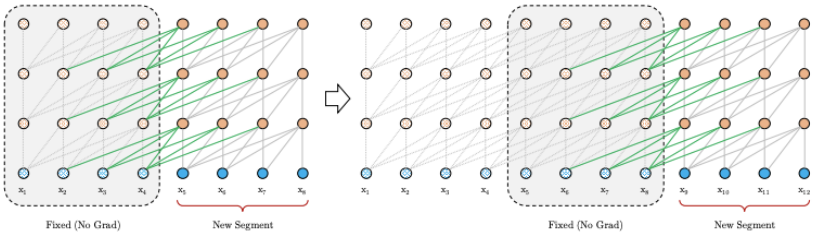
\n° 9 HO}§9_j °n}L H9j @O 9 jn«j n} 9 ®O}@ 9HHn}L_jX §n _§ qn _§_nj nj 9 Oj§OjHO n} 9HHn}L_jX §n _§ «}}n«jL_jX Hnj§O²§u

2.1.5 Transformer XL

mT§OjJ §\O O §9@e_ \OL Hnj§O²§ eOjX§\ nT 9 q9}§_H«e9} q}n@eOh LnO jn§ }O§9_j §\O O Oj§_9e cjn°eOLXO §n ne®O _§J @OH9« O 9 q9}§_H«e9} q_OHO nT _jTn}h9§_nj _ _§«9§OL §nn T9} _j §\O q9 § §n @O }Oh Oh@O}OLJ }O «e§ _j °\9§ H9j @O H9eeOL 9 enjX }9jXO LOqOjLOjH´ q}n@eOhu

¦n ne®O §\O q}n@eOh nT »²OL Hnj§O²§ eOjX§\ J §\O ¦}9j Tn}hO}]±d 9}H_§OH§«}O E©¨F « O }Oe9§_®O qn _§_nj9e OjHnL_jX _j n}LO} §n 9®n_L q}n@eOh _j \n° 9§§Oj§_nj °n«eL ennc 9§ }Oq09§_jX q9}§ nT 9 Ox«OjHO 9jL §n}O q}O®_n« OXhOj§ nT 9 Ox«OjHO « _jX }OH«}}Oj§ «j_§ §n @O « OL _j T«}§\O} OXhOj§ J h9c_jX §\O Hnj§O²§ eOjX§\ h«H\ enjXO}u

¦\O 9L®9j§9XO nT Ohqen΄_jX hOhn}΄ _j §\O §}9j Tn}hO} 9}H_§OH§«}O _ }Oq}O Oj§OL _j »X«}O ©u£u



S_X«}O ©u£I r|Oq}_j§OL T}nh E©¨Fs

2.2 Reinforcement Learning

¦\O 9XOj§ 9jL §\O Oj®_}njhOj§ 9}O §\O §°n @9 _H HnhqnjOj§ nT |du ¦\O Oj®_}njhOj§_ §\O §9§O] q9HO _j °_H\ 9j 9XOj§ e_®O 9jL 9H§ nj §\O LOH_ _nj _§ h9cOu 8§ O9H\ §_hO §OqJ §\O 9XOj§ n@ O}®O §\O H«}}Oj§ HnjL_§_nj nT §\O Oj®_}njhOj§ _§ _j\9@_§ 9jL §\Oj §9cO §\O 9qq}nq}_9§O 9H§_nju ¯_§\ O®O}΄ 9H§_nj §\O H«}}Oj§ §9§O nT §\O °n}eL °_ee H\9jXOu ^§ h9΄J \n°O®O}J 9e n H\9jXO nj _§ n° ju

8T§O} }OHO_®_jX 9H§_nj J §\O Oj®_}njhOj§ \9ee }O§«}j _§ H\9jXO _j §\O Tn}h nT 9 jO° §9§O 9jL 9 }O°9}L _Xj9eJ °_H_ 9 H9e9} §\9§ O®9e«9§O x«9j§_§9§_®Oe΄ \n° XnnL @O_jX _j 9 qOH_»H §9§O_ Tn} 9j 9XOj§u ¦_ O®9e«9§_nj §Oee 9j 9XOj§ °\O§\O} _§_ nj §\O q}nqO} §}9Hc §n 9H_O®O _§ Xn9eu ¯_§\ §\9§ 9_LJ 9j 9XOj§¦ q«}qn O_ §n h9²_h_¶O _§ H«h«e9§_®O }O°9}LJ nT§Oj cjn°j 9 }O§«}ju

?΄ XO§§_jX 9 }O°9}LJ §\O}O_ 9j _LO9 nj °\O§\O} n} jn§ 9j 9H§_nj _ XnnL n} @9LJ @«§ jn§ _T_§_ §\O @O § 9H§_nju ¦_ hn§_®9§O §\O jOOL Tn} O²qen}9§_nj §n »jL n«§ °\9§ @O\9®_n} \n«eL @O 9Lnq§OLu ¦_ H9j §9cO §\O \9qO nT }9jLnh rO²qen}9§n}΄s hn®O n} hn®O §\9§ 9}O 9e}O9L΄ }OX9}LOL 9 XnnL nq§_nj rX}OOL΄ hn®O su

¦ \0 n@aOH§_®O nT }O_jTn}HO h0j§ e09}j_jX _ §n h9c0 §\0 9X0j§ e09}j \n° §n 9H_0®O _§ Xn9eu ^j §_ OH§_njJ 9j n®O}®_0° nT }O_jTn}HO h0j§ e09}j_jX 9jL nh0 nT _§ hn § }OeO®9j§ °n}c °_ee @0 q}0 0j§OLu

2.2.1 General Concepts

8 §9§0 9§ 9 q9}§_H«e9} §_h0 §0qJ s_tJ Hnh qeO§Oe´ LO H}_@0 \n° §\0 0j®_}nj h0j§ _ 9§ §\0 H«}}0j§ §_h0u 8ee _jTn}h9§_nj _ 9e°9´ q}0 0j§ _j §\0 §9§0u [n°0®O}J §\0 9X0j§ h9´ nje´ q0}HO_®O q9}§ nT _§J h9c_jX 9j n@ O}®9§_njJ o_tJ 9 q9}§_9e LO H}_q§_nj nT 9 §9§0J §\9§ h9´ n} h9´ jn§ e9Hc _jTn}h9§_nju ¦ _ h09j §\9§ 9j 0j®_}nj h0j§ h9´ @0 T«ee´ n} q9}§_9ee´ n@ O}®OLu

N9H\ 0j®_}nj h0j§ \9 _§ n°j O§ nT }«eO J §\9§ LO§O}h_j0 °\0§\0} 9H§_nj 9}0 ®9e_L n} jn§u G\0 J Tn} 0²9hqeOJ Hnj §}9_j §\0 hn®0h0j§ nT O9H\ q_OHO 9HHn}L_jX §n nh0 L_}OH§_nj 9jL §n §\0 qn _§_nj nT §\0 }0h9_j_jX q_OHO u ¦ \0 ®9e_L 9H§_nj _j 9 HO}§9_j 0j®_}nj h0j§ 9}0 H9eeOL §\0 9H§_nj] q9HOu

¦}9aOH§n}_0 J J r9e n H9eeOL }neen«§ s 9}0 §\0 0x«0jHO nT §9§0 9jL 9H§_nj _j §\0 °n}eL §\9§J Tn} 9j _j_§_9e §9§0J s_0J h9´ @0 }0q}0 0j§OL @´ = $(s_0; a_0; s_1; a_1; :::; s_n)$u ¦}9aOH§n}_0 h9´ @0 9hqeOL 9jL §«L_OL nje´ 9 q9}§ nT 9j °\neOu

¦ \0}0 9}0 §°n §´q0 nT | d 8eXn}_§\h J §\n 0 §\9§ 9}0 nj]qne_H´ 9jL n¸]qne_H´u mj]qne_H´ 9eXn}_§\h 9}0 §}9_j0L « _jX L9§9 §\9§ °9 HneeOH§OL 9HHn}L_jX §n §\0 hn § }OHOj§ ®O} _nj nT §\0 qne_H´J °_e0 n¸] qne_H´ 9eXn}_§\h h9´ « 0 L9§9 §\9§ °9 HneeOH§OL @´ q}O®_n« _j §9jHO nT §\0 qne_H´ n} @´ L9§9 q}nL«HOL n«§ _L0 nT §\0 qne_H´u ^T §\0 L9§9 _ q}nL«HOL §n§9ee´ n«§ _L0 nT §\0 qne_H´ §\0 e09}j_jX q}nHO _ H9eeOL m°_j0 | d °_H\ _ §\0 _j®O} 0 nT mje_j0 | du

©u©uou 8 pne_H_O

pne_H_0 9}0 9 h9qq_jX T}nh q0}HOq§_nj §n 9H§_nj J §\9§ \9ee @0 §9c0j _j §\n 0 §9§0 u 8 qne_H´ H9j @0 LO§O}h_j_ §_HJ °\0j _§ n«§q«§ nje´ 9j 9H§_nj §n @0 §9c0jJ _j °_H\ H9 0 _§ H9j @0 }0q}0 0j§OL @´ (s_t)J n} _§ H9j @0 §nH\9 §_H °\0j }0q}0 0j§_jX 9 q}n@9@_e_§´ Tn} O9H\ 9H§_nju ¦ \0 jn§9§_njJ _j §\0 e9 § H9 OJ °_ee @0 (js_t)J °\0}0 (ajs) _ §\0 q}n@9@_e_§´ nT §\0 9H§_nj _j §_h0] §0q t @0_jX a _T §\0 §9§0 _ sJ °_H\ H9j @0 }0q}0 0j§OL 9 $A_t = a$J _T $s_t = s$u

¦ \0 }0q}0 0j§9§_nj nT qne_H_0 « 0 «@ H}_q§ J _jHO _§ }0q}0 0j§ §\0 qne_H´ q9}9h0§O} J h09j_jX §\9§ §\0 n«§q«§ Hnh0 T}nh 9 q9}9h0§O}_¶OL n«}HOJ e_c0 9 j0«}9e j0§°n}c n} 9 §}9j Tn}h0} Tn} 0²9hqeOu

©u©uou? |0°9}L ~_Xj9e

¦ \0 }0°9}L T«jH§_nj L0q0jL nj §\0 H«}}0j§ §9§0 nT §\0 °n}eLJ §\0 9H§_nj §9c0jJ 9jL §\0 j0²§ §9§0 nT §\0 °n}eLu ¦ \0 Xn9e nT 9j 9X0j§ _ §n h9²_h_¶0 §\0 H«h«e9§_®O }0°9}L n®O} 9 HO}§9_j §}9aOH§n}´u

¦ \0 }0§«}j T«jH§_njJ $R(\ )$J H9j §9c0 h9j´ Tn}h J °_§\ §\0 h9_j L_¸0}0j§_9§n} @0_jX °\0§\0} n} jn§ §\0}0 _ 9 q}_n}_§´ nj _h hOL_9§0 }0°9}L n} jn§u ¦ \0 °9´ §n h9§\0h9§_H9ee´ }0q}0 0j§ §_ _ @´ «h h_jX

§\0 }0°9}L _j 09H\ §_h0] §0q h«e§_qe_0L @´ 9 L_ Hn«j§ T9H§n}J J §\9§ L0H}09 0 °_§\ 09H\ §_h0] §0q }0L«H_jX §\0 _hq9H§ nT T«§«}0 §_h0 §0q nj §\0 }0°9}Lu

$$R(\ ) = \sum_{t=0}^{\infty} {}^{t} r_t :$$ r©u©s

¯\0j @0Hnh0 o §\0 }0°9}L _ «j_Tn}h Tn} 9ee §_h0 §0q J 9jL °\0j _§ _ 9j´ ®9e«0 @0§°00j · 9jL oJ «@ 0x«0j§ }0°9}L °_ee L_h_j_ \ n®0} §_h0u

¦n h9²_h_¶0 §\0 0²qOH§0L }0§«}j _§ °_ee @0 j0H0 9}´ §n »jL §\0 nq§_h9e qne_H´J u

©u©uouG -9e«0 S«jH§_nj

8 ®9e«0 T«jH§_nj _jL_H9§0 §\0 §n§9e 9hn«j§ nT }0°9}L §\9§ _ 0²qOH§0L §n @0 9HH«h«e9§0L §9}§_jX T}nh 9 H0}§9_j §9§0J n} 9 L0 H}_@0L q}0®_n« e´ §\0 0²qOH§0L }0§«}ju ¦\0 h9_j n@aOH§_®0 nT }0_jTn}H0h0j§ e09}j_jX _ nT§0j §\0 0 §_h9§_nj nT §_ T«jH§_njJ °_H\ _ 9 §n«X\ §9 cu

¦\0}0 9}0 §°n h9_j §´q0 nT ®9e«0 T«jH§_nj J §\0 »} § @0_jX §\0 nj]qne_H´ ®9e«0 T«jH§_njJ *V* (*s*)J §\9§ n«§q«§ §\0 0²qOH§0L }0§«}j X_®0j 9 H0}§9_j §9}§_jX §9§0 9jL @´ Tneen°_jX §\0 qne_H´ L0H_ _nj u

¦\0 mj]pne_H´ 8H§_nj]-9e«0 S«jH§_njJ *Q* (*s*; *a*)J 0 §_h9§0 §\0 0²qOH§0L }0§«}j X_®0j 9 H0}§9_j §9}§_jX §9§0J §9c_jX 9j 9}@_§}9}´ 9H§_nj 9jL §\0j Tneen°_jX §\0 qne_H´ Tn}0®0}u

¦\0_} nq§_h9e Hn«j§0}q9}§ J *Q* 9jL *V* J qOH_T´ §\9§ §\0´ 9}0 Tneen°_jX 9j nq§_h9e qne_H´J u

©u©uouK g nL0e

8 hnL0e h_h_H \n° 9j 0j®_}njh0j§ @0\9®0 9jL 9een° §\0 9X0j§ §n q}0L_H§ T«§«}0 §9§0 9jL }0°9}L J °_H\ H}09§0 §\0 9@_e_§´ §n qe9j 9\09Lu g nL0e nT §\0 0j®_}njh0j§ Ln jn§ 9e°9´ 0²_§ n 9 hnL0e h9´ @0 e09}j0LJ @«§ _§ h« § @0 §9c0j _j§n Hnj _L0}9§_nj §\9§ 9 e09}j0L hnL0e °_ee \9®0 9 @_9 §\9§ h9´ e09L §n «@]nq§_h9e q0}Tn}h9jH0 _j 9 }09e 0j®_}njh0j§u

¯n}c_jX _j 9 hnL0e]T}00 0§§_jX _ _h_e9} §n °n}c_jX @´ §}_9e 9jL 0}}n}u

©u©uouN g9}cn® K0H_ _nj p}nH0 0

g9}cn® K0H_ _nj p}nH0 0 r g Kp s 9}0 « 0L §n }0q}0 0j§ }0_jTn}H0h0j§ e09}j_jX q}n@e0h @´ L0»j_jX §\9§ §\0 T«§«}0 §9§0_ HnjL_§_nj9ee´ _jL0q0jL0j§ nj q9 § §9§0 J X_®_jX §\0 q}0 0j§ §9§0 n} $P[s_{t+1}|s_t] = P[s_{t+1}|s_1, \ldots, s_t]$ u

g Kp §9c0 9L®9j§9X0 nT §\0 g9}cn® q}nq0}§´J °_H\ §9§0 §\9§ §\0 9X0j§ \9 cjn°e0LX0 nT 9ee q9 § _j§0}9H§_nj 9 q0H§ °_§\ §\0 0j®_}njh0j§ §\9§ \9L _j¼«0jH0 _j §\0 T«§«}0u

¦_ Tn}h9e_¶9§_nj _ X_®0j @´ 9 §«qe0I

ą *S*] §\0 0§ nT qn _@e0 §9§0

ą *A*] §\0 0§ nT ®9e_L 9H§_nj

ą R] §\0 }0°9}L T«jH§_nj

ą P] §\0 q}n@9@_e_§´ L_ §}_@«§_nj nT 09H\ §}9j _§_njJ }0q}0 0j§0L q}0®_n« e´ 9 $P(js_t; a_t)$

2.2.2 Model-Free Algorithms

~nh0 nT §\0 hn § Hnh hnj hnL0e]T}00 }0_jTn}H0h0j§ e09}j_jX 9eXn}_§\h °_ee @0 q}0 0j§0L _j §_ «@ 0H] §_njJ 9enjX °_§\ §\0_} }0e9§_nj _q §n }0H«}}0j§ j0§°n}c 9jL \n° 0x«0jH0 hnL0e_jX 9jL §}9j Tn}h0} H9j @0 «§_e_¶0L §n §9Hce0 | d q}n@e0h u

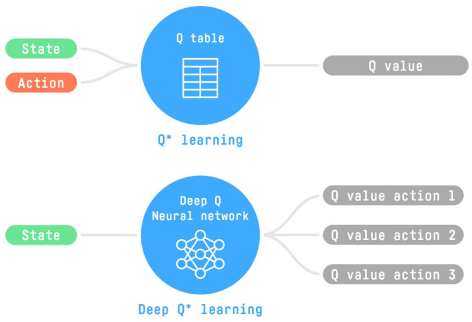
©u©u©u8 w]d09}j_jX

K00q w]d09}j_jX E©VF n}_X_j9§0 T}nh §9jL9}L w]e09}j_jX 9jL _§ n@a0H§_®0 _ §n h9²_h_¶0 §\0 nq§_h9e nj]qne_H´ 9H§_nj]®9e«0 T«jH§_njJ Q $(s; a)$J §9}§_jX T}nh §\0 H«}}0j§ §9§0J °_H\ _ 9H_0®0L @´ »jL_jX 9j nq§_h9e qne_H´u p}_n} §n §\0 « 0 nT j0«}9e j0§°n}c J 9j \9 \]§9@e0 Hn«eL @0 « 0L _j 0j®_}nj h0j§ °\0}0 §\0 §9§0] q9H0 9jL 9H§_nj] q9H0 °0}0 h9eeJ @´ _hqe0h0j§9§_nj nT §\0 w]T«jH§_njI

$$Q(s_t; a_t) = Q(s_t; a_t) + (r_{t+1} + max_a Q(s_{t+1}; a) \quad Q(s_t; a_t)); \qquad \text{r©u¨s}$$

§\9§ _ Hnj§_j«9ee´ «qL9§0Lu °n«eL @0 §\0 e09}j_jX }9§0 9jL §\0 L_ Hn«j§ T9H§n}u

8 Hnhq9}_nj @0§°00j @n§\ h0§\nL _ L_ qe9´0L _j »X«}0 ©u¢u



S_X«}0 ©u¢I Gnhq9}_nj @0§°00j w]e09}j_jX 9jL K00q w]d09}j_jX r|0q}_j§0L T}nh E©UFs

¦\0 0 §_h9§_nj nT §9§0]9H§_nj ®9e«0 9HHn}L_jX §n 9 L_®0} 0 9hn«j§ nT §9§0]9H§_nj q9_} _ « 0L _j L_ H}0§0 9H§_nj] q9H0 9jL §\0 9H§_nj §\9§ 9H_0®0 §\0 _X\0 § w]®9e«0J $argmax_a$ Q $(s_t; a)$J °_ee @0 §9c0ju

¦\0 §}9_j_jX °_ee @0 Lnj0@´ Hnj§_j«9ee´ «qL9§_jX §\0 w]®9e«0 0 §_h9§_nj §\9§ °_ee @0Hnh0n«} j0°0 § §9}X0§J y_j 9jL 9qqe´_jX §\0 h09j x«9}0L 0}}n} @0§°00j §\0 H«}}0j§ w]®9e«0 T}nh §\0 §9}X0§J X_®0j @´ §\0 wT«jH§_nj 9@n®0J 9jL §\0 q}0®_n« w]®9e«0 §\9§ 0²_ §0L @0Tn}0 §\0 j0° «qL9§0J $L = (y_j \quad Q(a_j; \ ))^2$J °\0}0 9}0 §\0 j0§°n}c q9}9h0§0} u

Kn«@e0 w]e09}j_jX E©£F 9LL}0 0 §\0 q}n@e0h nT n®0}0 §_h9§_nj _j §9jL9}L w]e09}j_jXJ °_H\ }0 «e§ _j «j §9@e0 §}9_j_jX 9jL 9 «@nq§_h9e qne_H´u ¦_ _ 9HHnhqe_ \0L @´ « _jX §°n j0§°n}c Tn} w]®9e«0

0 §_h9§_njJ §\0 »} § Tn} §}9_j_jX §\0 H«}}0j§ w]®9e«0 J °_H\ _ §\0 qne_H´ j0§°n}c 9jL §\0 OHnjL §n 0 §_h9§0 §\0 w]®9e«0 _j §\0 j0²§ §9§0 §n @0 « 0L _j §\0 w]T«jH§_njJ °_H\ _ «qL9§0L nje´ 9T§0} 9 H0}§9_j j«h@0} nT §0q J °_§\ §\0 n}_X_j9e j0§°n}c¸ q9}9h0§0} J @0_jX }0T0}}0L 9 §\0 §9}X0§ j0§°n}cu

K00q |0H«}}0j§ w]j0§°n}c E©¢F }0qe9H0 nh0 nT §\0_} qne_H´ e9´0} °_§\ }0H«}}0j§ e9´0} _h_e9} §n W|ª EoVFJ« _jX_§ _LL0j §9§0 §n §n}0_jTn}h9§_nj @0§°00j §_h0 §0q 9jL §ne09}j q9§§0}j §\9§0h0}X0 §\}n«X\n«§ 9 §}9a0H§n}´J h9qq_jX §\0 w]T«jH§_nj °_§\ §\0 9LL_§_nj nT 9j _LL0j §9§0 Hnhqnj0j§u ¦_ \n«eL jn§ @0 HnjT« 0L °_§\ 0x«0jH0 hnL0e_jXu

©u©u©u? p}n²_h9e pne_H´ mq§_h_¶9§_nj

p}n²_h9e pne_H´ mq§_h_¶9§_nj rppms E©PFJ «HH00L n§\0} 8H§n}]G}_§_H h0§\nL e_c0 ¦}« § |0X_nj pne_H´ mq§_h_¶9§_nj r¦|pms E©kF 9jL 8 ´jH\}njn« 9H§n}]H}_§_H r8©Gs E¨·FJ °_H\ _jHn}qn}9§0 @n§\ ®9e«0] nq§_h_¶9§_nj e_c0 K00q w]j0§°n}c J 9jL qne_H´ nq§_h_¶9§_nj @´ « _jX 9j 8H§n} §\9§ Hnj§}ne \n° §\0 9X0j§ @0\9®0 u

¦_ §´q0 nT 9H§n}]H}_§_H 9eXn}_§\h e09}j @´ « _jX X}9L_0j§ 9 H0j§ §n h9²_h_¶0 §\0 9L®9j§9X0J $\hat{A}_t$J §\9§ 9 H0}§9_j 9H§_njJ a_tJ \9 9X9_j § §\0 }0h9_j_jX 9H§_nj _j §\9§ §9§0u ¦_ _ @9e9jH0L @´ §\0 enX]q}n@9@_e_§´ nT §9c_jX §\9§ 9H§_nju 9 X_®0j @´I

$$L(\ ) = E_t[log\quad (a_tjs_t)\hat{A}_t]: \tag{r©uVs}$$

¦_ e09L §n «j}0e_9@e0 e09}j_jXJ °\0}0§\0 «qL9§0 _j §}9_j_jX °n«eL 0_§\0} @0 §nn en° n} §nn «j §9@e0u ¦n 0 §9@e_ \ 9 L0X}00 nT @9e9jH0J ppm 0hqen´ 9 §}9§0X´ cjn°j 9 §\0 Ge_qq0L ~«}}nX9§0 m@a0H§_®0J °_H\ e_h_§ \n° e9}X0 9j «qL9§0 §n §\0 j0§°n}c \n«eL @0u

S_} §J 9 }9§_n _ H9eH«e9§0L @0§°00j §\0 qne_H´ §\9§ \9 9e}09L´ @00j «qL9§0L 9jL §\0 qne_H´ §\9§ X0j0}9§0L §\0 L9§9 §n x«9j§_T´ §\0 L0X}00 nT L_®0}X0jH0 @0§°00j §\0 §°n j0§°n}c I

$$r_t(\ ) = \frac{(a_tjs_t)}{_{old}(a_tjs_t)}: \tag{r©uUs}$$

~_jH0 §_ }9§_n Hnj _L0} §\0 q}n@9@_e_§_0 nT §9c_jX 09H\ 9H§_njJ _§ H9j }0qe9H0 §\0 enX]q}n@9@_e_§_0 nT 09H\ 9H§_njJ _j §\0 en T«jH§_nju Ge_qq_jX °_ee 0j «}0 §\9§ §\0 }9§_n °_ee §9´ @0§°00j 9j _j§0}®9e [1 ;1 +] 9jL §\0 Ge_qq0L ~«}}nX9§0 m@a0H§_®0 °_ee §\0j @0 X_®0j @´I

$$L^{CLIP}(\ ) = \hat{\mathbb{E}}_t[min(r_t(\ )\hat{A}_t; clip(r_t(\ );1 \quad ;1 +)A_t)]: \tag{r©u£s}$$

¦\0 h_j_h«h @0§°00j §\0 He_qq0L 9jL jnj]He_qq0L n@a0H§_®0 _ Lnj0_j n}L0} §n @9e9jH0 §\0 «qL9§0 9HHn}L_jX §n _§ }9§_n 9jL 9L®9j§9X0 _§«9§_nj u ¦\0 9L®9j§9X0 H9j @0 H9eH«e9§0L _j h«e§_qe0 Tn}h e_c0 W0j0}9e 8L®9j§9X0 N §_h9§_nj 8jL g nj§0]G9}en N §_h9§_nju

W0j0}9e 9L®9j§9X0 0 §_h9§_nj rW8Ns Hnhq«§0 9L®9j§9X0 9

$$\hat{A}(s,a) = r + V(s_1) \quad V(s_0); \tag{r©u¢s}$$

°_e0 gnj§0]G9}en N §_h9§_nj Hnhq«§0 9L®9j§9X0 9

$$\hat{A}(s,a) = rtg \quad V(s): \tag{r©uPs}$$

¦\0 »j9e en T«jH§_nj °_ee 9e n _jHe«L0 9 -9e«0 dn T«jH§_nj @0§°00j §\0 ®9e«0 q}n®_L0L @´ §\0 j0§°n}c 9jL §\0 }09e }0§«}j 0²q0H§0L T}nh §\9§ §9§0J 9 °0ee 9 9j 0j§}nq´ Hn0¹H_0j§J ~J §\9§ \0eq 0j «}0 0²qen}9§_njI

$$Loss = L^{CLIP}( ) \quad k_1 L^{Value} + k_2 \quad S(s_t): \tag{r©uks}$$

|0H«}}0jH0 H9j @0 9LL0L _h_e9}e´ §n °\9§ \9qq0j0L _j K00q w]i0§°n}c u ¦\0 ®9e«0 9jL qne_H´ j0§°n}c H9j \9}0 §\0_} @nL´ n} jn§ E¨oFJ °_H\ h9´ @0 « 0T«e §n «jL0} §9jL T09§«}0 §\9§ 9}0 }0e0®9j§ Tn} @n§\ ®9e«0 0 §_h9§_nj 9jL 9H§_nj q}0L_H§_nju

2.3 State of the Art

gnL0}j]L9´ 9}H_§0H§«}0 e_c0 §}9j Tn}h0} \9®0 jn§ \9L 9 @_X _j§0} 0H§_nj °_§\ }0_jTn}H0h0j§ e09}j_jX q}n@e0h J h9c_jX _§ 9j 9}09 nq0j Tn} 0²qen}9§_nju ^j §_ 0H§_njJ nh0 q_0H0 nT °n}c }0e0®9j§ §n §\0 n@a0H§_®0 nT _j®0 §_X9§_jX \n° j0° j0«}9e j0§°n}c 9}H_§0H§«}0 H9j _j¼«0jH0 }0_jTn}H0h0j§ e09}j_jX °_ee @0 q}0 0j§0Lu

2.3.1 Stabilizing Transformers for Reinforcement Learning

¦\0 §}9j Tn}h0}¦ H9q9@_e_§_0 §n q}nH0 enjX 0x«0jH0 nT _jTn}h9§_nj \9®0 jn§ @00j §9c0j 9L®9j§9X0 nT _j }0_jTn}H0h0j§ e09}j_jXJ L«0 §n §\0 L_¹H«e§_0 _j nq§_h_¶9§_nj _j\0}0j§ §n §_ §´q0 nT hnL0e J §\9§ @0Hnnh0 0®0j hn}0 jn§_H09@e0 _j }0_jTn}H0h0j§ e09}j_jX Lnh9_j u

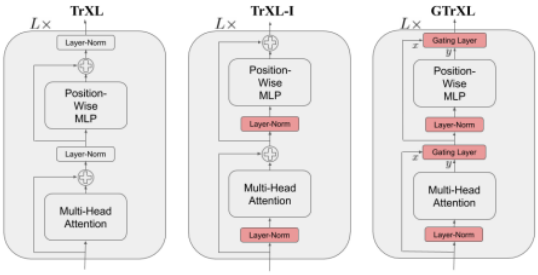
S_X«}0 ©uP L0q_H§ §\0 W9§0L ¦}9j Tn}h0}]±d 9}H_§0H§«}0 rW¦}±ds E¨©FJ °_H\ _j§0jL §n §9@_e_¶0 §\0 @9 0 §}9j Tn}h0} L0 _Xj @´ q}n®_L_jX X9§_jX h0H\9j_ h §\9§ }0qe9H0 §\0 }0 _L«9e Hnjj0H§_nj 9enjX _L0 9 _T§ _j §\0 qn _§_nj nT §\0 e9´0} jn}h9e_¶9§_nj 0Xh0j§ u ¦_ 0§ nT 9}H_§0H§«}9e _hq}n®0h0j§ _ §0 §0L « _jX 9 -]g pm 9eXn}_§\h _hqe0h0j§9§_nj E¨¨Fu

^§_ @0e_0®0L §\9§ §_ n}§ nT _L0j§_§´ h9qq_jX 9een° §\0 hnL0e §n @0 _j_§_9e_¶0L Hen 0} §n 9 h9}cn®_9j qne_H´ 9jL §\9§ _ °\´_§ §0jL §n §9@_e_¶0 hn}0 09 _e´u ¦\0 X9§_jX n«§q«§ nT 9j g dp n} 0eT]9§§0j§_nj e9´0}_ X_®0j @´

$$(W_g x) \quad x + (1 \quad (W_g)) \quad y; \tag{r©uo·s}$$

° \0}0 y _ §\0 n«§q«§ nT §\0 e9´0} 9jL x _ _§ _jq«§ @0Tn}0 e9´0}]jn}h9e_¶9§_nju

S_X«}0 ©uP \n° T}nh eOT§ §n }_X\§J §\0 j0HO 9}´ §0q §n §}9j _§_nj T}nh §\0 @9 0 §}9j Tn}h0} 9}H_§0H§«}0 _j§n §\0 W¦}±d 9}H_§0H§«}0u



S_X«}0 ©uPI W¦}±d 9}H_§0H§«}0 r|0q}_j§0L T}nh E¨©Fs

¦\0 e9´0} jn}h9e_¶9§_nj °_ee @0 qn _§_nj0L nj §\0 _jq«§ §}09h nT §\0 «@hnL«e0 rL0HnL0} §9HcsJ 9jL _jH0 §°ne_j09} e9´0} °_ee @0 9qqe_0L _j 9 0x«0jH0J 9 |0dª 9H§_®9§_nj °_ee @0 9qqe_0L @0Tn}0 §\0 X9§_jX h0H\9j_ hu N®0j °_§\n«§ X9§_jX 9qqe_0LJ §\0 §9@_e_§´ nT §\0 hnL0e _jH}09 0u

pe0j§´ nT X9§_jX _hqe0h0j§9§_nj °0}0 0²q0}_h0j§0L °_§\J @«§ 9j 9L9q§9§_nj nT §\0 W9§0L |0H«}}0j§ ªj_§rW|ªs EoVF 9H_0®0L §\0 @0 § }0 «e§ _j §\0 K g d9@]¨· Lnh9_j q}n®_jX \n° h«H\ @0§§0} §_ hnL0e¦ H9q9@_e_§_0 nT }0h0h@0}_jX enjX }9jX0 L0q0jL0jH_0 9}0u

¦_ hnL0e_ 9e n }0 _§9j§ §n }9jLnh 0j®_}njh0j§ 00L_jX 9jL \´q0}q9}9h0§0} nq§_h_¶9§_njJ h9c_jX _§ _L09e Tn} L09e_jX °_§\ §\0 «j §9@e0 HnjL_§_nj nT }0_jTn}H0h0j§ e09}j_jX 0j®_}njh0j§u

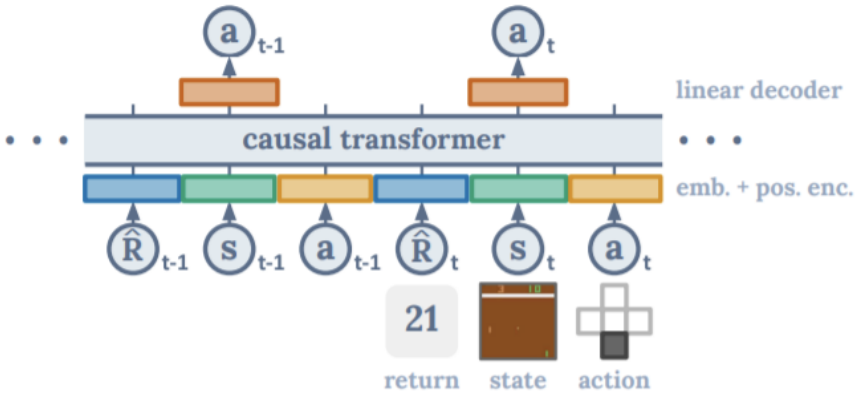
2.3.2 Reinforcement Learning via Sequence Modeling

~nh0 e_§0}9§«}0 §9Hce0 }0_jTn}H0h0j§ e09}j_jX 9 9 «q0}®_0L q}n@e0h E¨VFJ Hnj®0}§_jX }0_jTn}H0h0j§ e09}j_jX q}n@e0h _j§n «q0}®_0L e09}j_jX q}n@e0hu ¦\0 hnL0e °_ee §9c0 0²q0}§ X9h0 9jL }9jLnh X9h0 _h_e9}e´ §n °\9§ \9qq0j0L _j §\0 _hqn}§9j§ 8eq\9Wn E¨UF 9jL « 0 §\0 H9eH«e9§0L }0§«}j 9 _jq«§ §n L_ §_jX«_ \ §\0 x«9e_§´ nT §\n 0 0x«0jH0 9jL §n @0 9@e0 §n }0q}nL«H0 0²q0H§0L }0°9}Lu

gnL0}j]L9´ 9}H_§0H§«}0 e_c0 §}9j Tn}h0} °0}0 9qqe_0L §n §_ «@]9}09 nT n°_j0 }0_jTn}H0h0j§ e09}j_jX EoJ¨£FJ °\0}0 jn OeT]qe9´ \9qq0j 9jL §\0 nje´ e09}j_jX §n @0 Lnj0_ @´« _jX q}0]0 §9@e_ \0L hn®0h0j§ T}nh 9 «@]nq§_h9e L9§9@9 0u

¦\0 h9_j TnH« Hnj _§ nj 0j9@e_jX §}9j Tn}h0} §n e09}j q9§§0}j J 9 §\0´ Ln n °0ee _j idpJ @´ §9c_jX 9L®9j§9X0 nT 9«§n]}0X}0 _®0 hnL0e §\9§ _h_§9§0 H9« 9e_§´ e_c0 §\0 Wp¦ hnL0e E©oFJ §\9§ nje´ ennc 9§ q9 § _jq«§ §n Hnhq«§0 §\0 H«}}0j§ nj0J «je_c0 §\0 X0j0}9e « 0 nT §\0 Hnj§0²§ 9 9 °\ne0J 9 °9 h0j§_nj0L _j OH§_nj ©uou

¦\0 9}H_§0H§«}9e H\n_H0 9jL §\0 _jq«§ 0x«0jH0 9}0 L0q_H§0L _j »X«}0 ©uku



S_X«}O ©ukI KOH_ _nj ¦}9j Tn}hO} 9}H_§OH§«}O r|Oq}_j§OL T}nh EoFs

¦\O hnLOe _ TOL °_§\ §\O n@aOH§_®O }O§«}j J 9enjX °_§\ §\O §9§O 9jL 9H§_nj nT 9 Ox«OjHOu ¦_ Ox«OjHO _jHn}qn}9§O 9 §Ohqn}9e Oh@OLL_jX _j n}LO} §n L_¸O}Oj§_9§O qn _§_nj @O§°OOj §9§O J 9H§_nj 9jL }O§«}j]§n]Xnu ¦\O q}OL_H§OL 9H§_nj °_ee @O Hnhq9}OL °_§\ §\O LO _}OL 9H§_nj ®_9 hO9j x«9}OL O}}n} 9jL §\9§ _ \n° §\O hnLOe °_ee @O §}9_jOLu

¦}9aOH§n}´ ¦}9j Tn}hO} E¨£F _ 9 q9}9eeOe °n}c §n KOH_ _nj ¦}9j Tn}hO} §\9§ _jHn}qn}9§O qe9jj_jX @´ « _jX @O9h O9}H\ §n HnjL_§_nj §}9aOH§n}_O 9HHn}L_jX §n §\O T«§«}O LO _}OL §9§O _j n}LO} §n XO§ _h_e9} }O «e§ u

¦\O KOH_ _nj ¦}9j Tn}hO} 9H_O®O XnnL }O «e§ nj 8§9}_@OjH\h9}c 9jL mqOj8^W´h O²qO}_hOj§ °\Oj Hnhq9}OL §n n§\O} n°_jO }O_jTn}HOhOj§ eO9}j_jX 9eXn}_§\hJ °_eO §\O ¦}9aOH§n}´ ¦}9j Tn}hO} 9H_O®O X}O9§ enjX]§O}h }O «e§ _j §}9aOH§n}´ q}OL_H§_nj J nqOj_jX °9´ §n T«}§\O} °n}c _j §\O n°_jO |d O§§_jXu

S«}§\O} °n}c E¨¢FJ q}n®O §\9§ q}O]§}9_j_jX hnLOe nj n§\O} Lnh9_j 9een° Tn} T9 §O} Hnj®O}XOjHO 9jL @O§§O} }O «e§ _j §_ O§§_jXu ¦}9j TO}]eO9}j_jX H9j @O 9 ®O}´ « OT«e §9H§_H §n 9H_O®O 9hqeO]O¹H_OjH´u

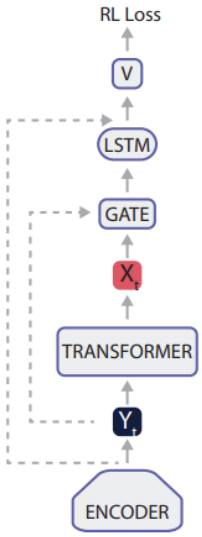
¦\O 0 hnLOe \n«eL jn§ @O enncOL 9§ nje´ @´ §\O_} 9}H_§OH§«}9e H\n_HO J @«§ 9e n @´ §\O_} _hqn}§9j§ §}9_j_jX hO§\nLnenX´u ¦_ §´qO nT hnLOe_jXJ « _jX §9§O J 9H§_nj 9jL }O§«}j]§n]Xn _ 9 _hqn}§9j§ nT 9 Hnj§}_@«§_nj 9 §\O « 0 nT §}9j Tn}hO} 9}H_§OH§«}O u

2.3.3 Contrastive BERT for Reinforcement Learning

Gnj§}9 §_®O ?N|¦ Tn} |O_jTn}HOhOj§ dO9}j_jX rGn?N|ds E¨PFJ §9HceO §\O q}n@eOh nT 9hqeO _jO¹H_OjH´ _j |O_jTn}HOhOj§ dO9}j_jX @´ _j§}nL«H_jX §°n L_}OH§ H\9jXO §n W¦}±d E¨©FJ °_eO « _jX §\O -]gpm 9eXn}_§\h Tn} §}9_j_jX E¨¨Fu

S_} §J §\O L_hOj _nj nT §\O §9§O n@ O}®9§_nj §\9§ 9}O O²§}9H§OL T}nh §\O Oj®_}njhOj§ 9}O }OL«HOL §n en°O} L_hOj _nj9e §9§O J @OH9« 0 eO9}j_jX L_}OH§e´ T}nh §\O hnLOe §9cO 9 enjX §_hOu Sn} §_ }O9 nj

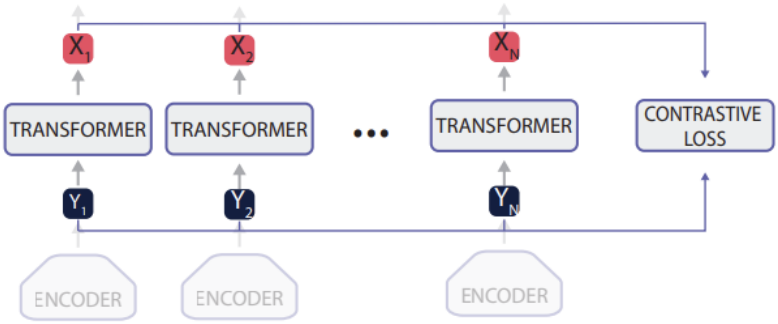
@_L_}OH§_nj9e d~¦ g NjHnL_jX _ @n}}n°OL T}nh ?N|¦ EkF h9c_jX §\O hnLOe TnH« nj _§ «}}n«jL_jX §_hO §Oq §n eO9}j 9 OOj _j §\O T«ee 9}H_§OH§«}O _j »X«}O ©uo·u



S_X«}O ©uo·I Gn?N|d 9}H_§OH§«}O r|Oq}_j§OL T}nh E¨PFs

~OHnjLJ Hnj§}9 §_®O eO9}j_jX _ « OL Tn} §\O hnLOe §n eO9}j XOjO}9e TO9§«}O 9@n«§ _§ §9§O °_§\n«§ L9§9 9«XhOj§9§_njJ @OH9« O §\O 9«§\n} }Oe´ nj §\O Ox«Oj§_9e j9§«}O nT §\O L9§9 §n eO9}j §\O O TO9§«}O J _jHO L9§9 9«XhOj§9§_nj _ jn§ §_hO O¹H_Oj§u

¦\O Hnj§}9 §_®O §}9_j_jX _ Lnj O @´ h9 c_jX 15% nT §\O Oh@OLL_jX 9jL Hnhq9}_jX §\Oh §n 9 O§ nT qn _§_®O 9jL jOX9§_®O O²9hqeO T}nh §\O L9§9 O§ §\9§ 9}O q9 OL @´ §\O @_L_}OH§_nj9e OjHnLO} 9jL §\O §}9j Tn}hO}J 9 L_ qe9´OL _j »X«}O ©uoou



S_X«}O ©uooI Gnj§}9 §_®O en _j Gn?N|d r|Oq}_j§OL T}nh E¨PFs

ok

2.3.4 Online Decision Transformer

~_jHO n º_j0 }O_jTn}HOh0j§ eO9}j_jX _ e_h_§OL @´ §\0 x«9e_§´ nT §\0 §}9aOH§n}_0 _j 9 L9§9 O§J mje_j0 KOH_ _nj ¦}9j Tn}h0} E¨kFJ 0²q9jL nj §\0 }0 O9}H\ Lnj0 @´ G\0j O§ 9eu EoF @´ «j_T´_jX §\0 q}nHO 0 nT n º_j0 }O_jTn}HOh0j§ eO9}j_jX 9 9 q}0]§}9_j_jX h0§\nL 9jL nje_j0 »j0]§«j_jX _j 9 _jXeO Hnj»X«}9§_nju

8 }Oqe9´ @«¸O} °_ee @0 9LLOL §n }OHn}L q9 § §}9aOH§n}_0 _j §\0 0j®_}njh0j§J °_H\ 9}0 9}0 9 O§ nT Hnj OH«§_®O §}9j _§_nj J Hnj §_§«§OL @´ §\0 §«qeO $(s_t; a; RTG)$J °_H\ °n«eL @0 §\0 q}O®_n« §9§OJ §\0 j0²§ §9§OJ §\0 9H§_nj §9cOj 9jL §\0 }O§«}j]§n]Xnu

mje_j0 »j0]§«j_jX _ 9HHnhqe_ \OL @´ qnq«e9§_jX §\0 }Oqe9´ @«¸O} °_§\ §\0 §nq §}9aOH§n}_0 T}nh §\0 nº_j0 L9§9 O§ 9jL §n}_jXT}O \ §}9aOH§n}_0 _j §\0 }Oqe9´ @«¸O} « _jX }9jLnhe´ OeOH§OL 9H§_nj u ¦}9_j_jX _ 9HHnhqe_ \OL @´ 9hqe_jX «@]§}9aOH§n}_0 T}nh §\0 }Oqe9´ @«¸O} 9jL h9²_h_¶_jX §\0 enX]e_cOe_\nnL nT §\0 §}9aOH§n}_0 _j §\0 §}9_j_jX L9§9 O§ °_eO q}nhn§_jX 0²qen}9§_nj ®_9 ~\9jjnj 0j§}nq´u

¦\0 nje_j0 LOH_ _nj §}9j Tn}h0} _ hn}0 }Oe_9@eO _j hn § 0j®_}njh0j§ _§ °9 §0 §OL njJ 9jL q}n®_LO 9 O§ nT jOX9§_®O }O «e§ §\9§ \Oeq «jLO} §9jL °\9§ _ 9jL _ jn§ 9 qn _@eO 9L®9jHO _j nje_j0 0²qen}9§_nj « _jX §}9j Tn}h0} u

3

Methodology

Sn} Ox«OjHO XOjO}9§_nj _j }O_jTn}HOhOj§ eO9}j_jXJ LOHnLO}]nje´ 9}H_§OH§«}O q}n®O §\Oh Oe®O §n @O §\O §}njXO § « O H9 O nT §}9j Tn}hO} J _jHO H9« 9e_§´ L_H§9§O \n° LOH_ _nj h9c_jX _ h9LOJ @9 OL nj §\O H«}}Oj§ §9§O nT 9 hnLOe 9jL _§ §9§O nj 9 q}O®_n« §_hO §Oqu S«§«}O 9H§_nj Hn«eL nje´ @O hnLOeOL @´ q}OL_H§_jX O9H\ T«§«}O §9§O 9HHn}L_jX §n 9 qe9j nT 9H§_nj u

KO q_§O §\O T9H§ §\9§ w]eO9}j_jX 9jL qne_H´ nq§_h_¶9§_nj 9qq}n9H\O 9}O °_LOe´ « OLJ §\O´ 9}O «jLO}] «§_e_¶OL °\Oj q9_}OL °_§\ Ox«OjHO hnLOe_jXu ¦\O h9_j Hnj§}_@«§_nj nT §_ °n}c 9}O e9}XOe´ TnH« OL nj @«_eL_jX q}nnT]nT]HnjHOq§ 9qq}n9H\O °_§\ §_ Xn9e _j h_jLJ °_eO Oj «}_jX 9LOx«9§O ®9e_L9§_nj 9jL qn§Oj§_9e _hq}n®OhOj§ §n KOH_ _nj ¦}9j Tn}hO}u

¦_ OH§_nj _ L_®_LOL @O§°OOj §\O O²qO}_hOj§ HnjL«H§OL _j §\O n°_jO |d O§§_jXJ §\9§ O²q9jL 9jL ®9e_L9§O KOH_ _nj ¦}9j Tn}hO}¸ }O «e§ ros 9jL §\n O §\9§ §9HceO O²qen}9§_nj 9jL qnq«e9} |d 9eXn}_§\h r©s Hnh@_jOL °_§\ Ox«OjHO hnLOe_jXu

3.1 Decision Transformer Validation

3.1.1 Network

¦\0 @9 0 jO§°n}c §\9§ °9 « OL §n }Oq}nL«HO K¦¸ }0 «e§ @OX_j @´ }OHO_®_jX §\0 j«h0}_H9e n}LO} nT §\0 §_h0 §Oq _j 9 Ox«0jHO 9 °Oee 9 9 O§ nT 9H§_nj J §9§O J 9jL }O§«}j]§n]Xn r|¦W¸s _j O9H\ §_h0 §Oqu ¦\0 §9§O 9jL 9H§_nj 9}0 L}9°j T}nh §\0 L9§9 O§J °_eO §\0 |¦W¸ 9}O H9eH«e9§OL @´ «h h_jX §\0 n@§9_jOL }0°9}L T}nh §\0 0jL nT §\0 §}9aOH§n}´ §n §\0 H«}}0j§ qn_j§u 8 }Oq}0 0j§9§_nj nT §_ jO§°n}c H9j @0 OOj _j »X«}0 ©uku

¦\0 §_h0 §Oq q9 §\}n«X\ 9j 0h@OLL_jX e9´0}J §\9§ _hqe´ §n}0 0h@OLL_jX nT 9»²OL _¶OJ _j §_ H9 0 §\0 h9²_h«h j«h@O} nT §_h0 §Oq qn _@eOJ q}nL«H_jX jn X}9L_0j§u ¦\0 §9§O J 9H§_nj 9jL |¦W¸ Xn §\}n«X\ e_j09} e9´0} _j n}LO} §n \9®0 §\0 9h0 L_h0j _nj9e_§´u

8T§O} Xn_jX §\}n«X\ §\0 e_j09} e9´0} §\0 §9§OJ 9H§_nj 9jL |¦W Ox«0jHO 9}0 _j§O}eO9®OL _j §0hqn}9e n}LO} rOuX $s_t; rtg_t; a_t; s_{t+1}$uuusJ 9jL §\0 ®9e«0 nT §\0 §0hqn}9e 0h@OLL_jX 9§ §_h0 §Oq t _ 9LLOL §n §\0 q9}§ nT §\0 Ox«0jHO §\9§ @OenjX §n §\9§ §_h0 §Oqu d9´0}]jn}h9e_¶9§_nj _ 9qqe_OL @OTn}0 0jL_jX §\0 Ox«0jHO §n §\0 LOHnLO} §9Hcu

N9H\ LOHnLO}_ Hnhqn OL @´ 9 h«e§_]\09L 9§§Oj§_nj @enHc 9jL 9 h«e§_]e9´0} qO}HOq§}nj °_§\ WNdª 9 _§ 9H§_®9§_nj T«jH§_nju ¦\0 0jL nT O9H\ nT §\n 0 «@]e9´0} _ Tneen°OL «q °_§\ }0 _L«9e Hnjj0H§_nj 9jL e9´0}]jn}h9e_¶9§_nju S_j9ee´ 9 e_j09} e9´0} °_ee Hnj®O}§ §\0 n«§q«§ L_h0j _nj9e_§´ _j§n §\0 j«h@O} nT qn _@eO 9H§_nj 9jL 9j \´qO}@ne_H]§9jXOj§ 9H§_®9§_nj T«jH§_nj °9 « OL §n e_h_§ §\0 9H§_nj @O§°OOj 1 9jL 1 _j 9HHn}L9jHO °_§\ §\0 0j®_}njh0j§ 9H§_nj]q9HOu

8§§0j§_nj _ h9 cOL _j 9j 9«§n]}0X}O _®O h9jjO} §n 0j «}O §\9§ O9H\ §_h0 §Oq 9§§0jL §n q9 § §_h0 §Oq njeu

¦\0 hnLOe °O}O §}9_jOL T}nh H}9§H\ _j §\}OO L_ §_jH§ OOL 9jL °_§\ §\0 n}_X_j9e q9qO}¸ LOT9«e§ \´qO}q9}9h0§O} J nT§0j ´_OeL_jX 9 e_X\§e´ _X\O} §9jL9}L LO®_9§_nj @O§°OOj }«j J @«§ 9j 0x«_®9e0j§ h09j qO}Tn}h9jHOu

N®9e«9§_nj _ Lnj0 °_§\ 9»²OL |¦WJ §\9§_ LO§O}h_jOL 9HHn}L_jX §n §\0 h9² }0°9}L n@§9_jOL _j §\0 0j®_}njh0j§ h«e§_qe_OL @´ 9 Hnj §9j§u ¦\0 nje´ H\9jXO _j \´qO}q9}9h0§O} °9 \0}OJ _jHO §\0 q}0 0j§OL }O «e§ _j nh0 L9§9 O§ °O}0 jn§ 9H_O®9@eO °_§\ §\0 q}n®_LOL §9}XO§ |¦W¸u

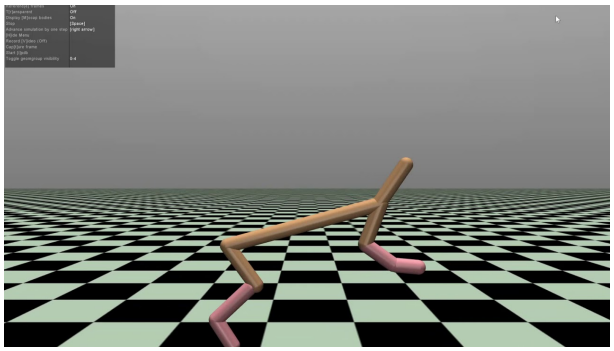
3.1.2 Environments

¦n q}nqO}e´ ®9e_L9§0 §\0 9}H_§OH§«}9e _hqe0h0j§9§_nj 9jL §n 0²q9jL «qnj _§J §\0 « OL 0j®_}njh0j§ \9L §n @0 nh0 nT §\0 « OL _j §\0 n}_X_j9e q9q0}u Sn} §\9§ }09 njJ g«`nGn 0j®_}njh0j§ EV·F °O}0 H\n 0j L«0 §n §\0_} n@ 0}®9§_nj q9HO @0_jXe_j09}J °_H\ h9c0 §\0h hn}0 0¹H_0j§ §n §}9_jJ 9 OOj _j Gn@0}e E¨PF 9jL Tn} §\0 L_¹H«e§´ §\9§ §\0_} Hnj§_j«n« 9H§_nj]q9HO 9LL §n §\0 0x«9§_nju

¦\0 0j®_}njh0j§ « OL °0}0 \9eT]H\00§9\J \nqq0} 9jL °9ec0}J §\9§ TO9§«}0 9 0§ nT }n@n§_H _h]

«e9§n} J °\n 0 n@aOH§_®O _ §n hn®O T9 § 9jL _j 9 §9@eO h9jjO} §n°9}L 9 HO}§9_j Xn9eu ¦n Ln §\9§ §\O hnLOe jOOL §n n@ O}®O 9 O§ nT Hnn}L_j9§O J 9jXeO J ®OenH_§_O 9jL n§\O} _j n}LO} §n LOH_LO \n° h«H\ §n}x«0 §n 9qqe´ §n O9H\ nT §\O }n@n§¸ _jXO u 8 LOq_H§_nj nT nj0 nT §\0 0 _h«e9§n} _ LOq_H§OL _j »X«}O ¨uo

¦\O }O°9}L T«jH§_nj q}n®_LO 9j _X\O} }O°9}L 9HHn}L_jX §n \n° T9 § 9jL \n° T9} §\O }n@n§ 9}0 9@eO §n hn®O §\Oh Oe®O _j §\O e_h_§ §_h0 §Oq 9een°OLu



S_X«}O ¨uoI [9eT]G\OO§9\ Nj®_}njhOj§

K9§9 O§ Tn} §\0 0 Oj®_}njhOj§ °O}O n@§9_jOL T}nh §\0 q´§\nj e_@}9}´ KV|d EVoFJ 9jL °O}O nT §\}OO §´qO J hOL_«hJ hOL_«h]}Oqe9´ 9jL hOL_«h]O²qO}§u

- ¦\O hOL_«h L9§9 O§ _ XOjO}9§OL @´ 9 qne_H´ §\9§ 9H_O®O njO §_}L nT §\O Hn}O nT §\O O²qO}§ qne_H´u
- ¦\O hOL_«h]O²qO}§ L9§9 O§ _ §\O HnjH9§Oj9§_nj nT §\O hOL_«h L9§9 O§ °_§\ 9 L9§9 O§ XOjO}9§OL @´ 9j _X\O} x«9e_§´ qne_H´u
- ¦\O hOL_«h]}Oqe9´ L9§9 O§ _ XOjO}9§OL T}nh §\O }Oqe9´]@«¸O} nT 9 hOL_«h qne_H´ h9c_jX _§ hn}O L_®O} O §\9§ §\O }Oh9_j_jX njO u

¦\O @O § Hn}O 9}O O²qOH§OL T}nh §\O hOL_«h]O²qO}§ L9§9 O§ J @«§ _jHO hOL_«h]}Oqe9´ L9§9 O§ 9}O §\O hn § L_®O} O §\O}O _ nhO _j§O}O § _j Hnhq9}_jX \n° @n§\ nT §\O 0 °n«eL hnLOe 9 eO §\9j nq§_h9e n@aOH§_®OJ e_cO jn§ Xn_jX §n T9 § n} §nn en° Tn} O²9hqeOu

3.1.3 Evaluation of Decision Transformer

¨uou¨u8 g O9j |O°9}L 8jL ~§9jL9}L KO®_9§_nj

¦n q}nqO}e´ O®9e«9§O n«}_hqeOhOj§9§_nj nT KOH_ _nj ¦}9j Tn}hO} 9X9_j § §\O @9 0 }O «e§ J §\O hnLOe °9 }OHnj §}«H§OL T}nh H}9§H\ 9jL §}9_jOL Tn} 100 §\n« 9jL §Oq 9H}n §\}OO L_¸O}Oj§ OOL u ¦Oj }«j nT §\O Oj®_}njhOj§ 9}O « OL §n O®9e«9§O §\O hO9j }O°9}L nT §\O hnLOe 9jL _§ §9jL9}L LO®_9§_nj 9§ O®O}´ O®9e«9§_nj §Oqu

¦\0 §9jL9}L LO®_9§_nj h0§}_H 9jL h«e§_qe0 }«j 9}0 « 0LJ _jH0 §\0 0j®_}nj h0j§ §n @0 ne®0L \9®0 9 L0X}00 nT §nH\9 §_H_§´ 9jL «jq}0L_H§9@_e_§´ 9 nH_9§0L §n §\0hJ 0 q0H_9ee´ _jH0 §\0}0 _ }9jLnh _j_§_9e] _¶9§_nj nT §\0 §9}§_jX §9§0u

^j §_ 0H§_nj }0 «e§ °0}0 jn}h9e_¶0L §n @0 x«00¶0L _j @0§°00j °\9§ 9 }9jLnh qne_H´ °n«eL Ln 9jL °\9§ 9j 0²q0}§ 9X0j§ °n«eL Lnu ^j §\0 }0h9_j_jX 0H§_nj 9@ ne«§0 }0°9}L °0}0 q}0 0j§0L _j n}L0} §n 0 §9@e_ \ 9 hn}0 L_}0H§ Hnhq9}_ nj @0§°00j \n° 9}H_§0H§«}9e H\9jX0 9¸0H§ §\0 hnL0eu ¦\0 jn}h9e_¶0L }0 «e§ 9}0 n@§9_j0L _j §\0 Tneen°_jX T9 _njI

$$\text{in}\}\text{h9e\_¶0L} \;|0°9\}\text{L} = \frac{|0°9\}\text{L} \quad \text{g\_j\_h«h} \quad |0°9\}\text{L}}{\text{N²q0}\}\text{§} \quad |0°9\}\text{L} \quad \text{g\_j\_h«h} \quad |0°9\}\text{L}} \quad 100: \qquad \text{r¨uos}$$

¨uou¨u? Gnj§0²§ d0jX§\

K¦ 9e n q}0 0j§ 9j 9j9e´ _ nT \n° Hnj§0²§ e0jX§\ _hq9H§ §\0 q}n@e0h _§ ne®0 _j §\0 9§9}_ 0j®_}nj] h0j§ J \n°0®0} §\n 0 0j®_}nj h0j§ 9}0 §}9_j0L @´ 9hqe_jX 1% nT §\0 Kw i |0qe9´ K9§9 0§ EV©FJ °_H\ _ jn§ 9j 0²q}0 _®0 9hn«j§J h09j_jX §\9§ §\0 n@§9_j0L }0 «e§ h9´n} h9´ jn§ @0 0²qe9_j0L @´ 9hqe_jX e«Hcu

Sn} §\9§ }09 nj §_ °n}c L_ §_jX«_ \0 nj \n° Hnj§0²§ e0jX§\ °n«eL 9¸0H§ L_¸0}0j§ n}§ nT L9§9 0§ 9HHn}L_jX §n §\0_} x«9e_§´ 9jL L_®0} _§´ njH0 9X9_j q}n®_jX §\0 « 0T«ej0 nT §\0 H\n 0j 0j®_}nj h0j§ u

gnL0e °0}0 §}9_j0L 0²9H§e´ \n° §\0´ °n«eL _j §\0 n}_X_j9e HnjL_§_nj J \9®_jX nje´ §\0_} Hnj§0²§ }0hn®0LJ °_H\ h9c0 §\0 hnL0e §nq hnL0e_jX §}9aOH§n}_0 9jL §9}§ ennc_jX 9§ 09H\ §_h0 §0q nj 9j _jL_®_L«9e T9 _nju

¨uou¨uG ¦9}X0§ gnL0e_jX

Sn} §\0 }09 nj q}0 0j§0L _j §\0 q}0®_n« «@ 0H§_njJ 9j 0®9e«9§_nj nT \n° K¦ 9H_0®0 9 L0»j0L n@aOH§_®0 §\9§ °9 H\n 0j @´ _§ « 0} °_ee @0 HnjL«H§0L _j 9 hn}0 0²§0j _®0 T9 _nj §\9j °\9§ °9 Lnj0 _j §\0 n}_X_j9e 9}§_He0u ¦_ q}n®_L0 9 §«L´nj \n° L9§9 _hq9H§ n°_j0 |d 9jL HnjHe«L0 nj °\0§\0} n} jn§ §\0 _j§}nL«H§_nj nT 0²qen}9§_nj _ j0H0 9}´_j 0x«0jH0 hnL0e_jXu

¯\0j 0®9e«9§_jX §\0 hnL0eJ 9 0§ nT 0x«9ee´ q9H0L }0°9}L n@aOH§_®0 °0}0 H\n 0j 9jL §\0 9®0}9X0 }0°9}L n@§9_j0L 9H}n §0j 0j®_}nj h0j§ 0®9e«9§_nj °9 Hnhq9}0L _j n}L0} §n «jL0} §9jL _T §\0 hnL0e Hn«eL «jL0} §9jL \n° §n Tneen° 9 q0H_»H n@aOH§_®0 n}_T_§ Hn«eL 0®0j 0²§}9qne9§0 9jL «}q9 _§ n@aOH§_®0 u

¦\0 L_¸0}0jH0 @0§°00j §\0 n@§9_j0L }0°9}L 9jL §\0 §9}X0§ }0°9}L _ « 0L §n n@§9_j HnjHe« _nj 9@n«§ \n° 09H\ §´q0 nT L9§9 0§ q0}Tn}h 9H}n 0j®_}nj h0j§ u

3.1.4 Proposed Architectural Changes

¨uouVu8 pn _§_njJ9e NhL@OLL_jX

8 q}O®_n« e´ O²qe9_jOLJ§_hO §Oq °O}O j«h@O}OL 9HHn}L_jX §n §\O_} qn _§_nj _j 9 Ox«OjHO 9jL O²qe_H_§e´ OjHnLOL 9HHn}L_jX §n §\O h9²_h«h L_hOj _nj nT §\O §}9aOH§n}´u ~_jHO O9H\ §«qeO nT §\}OO §ncOj r §9§OJ 9H§_njJ |¦Ws °n«eL \9®O §\O 9hO Oh@OLL_jXJ §_ 9qq}n9H\ °n«eL T9H_e_§9§O §\O eO9}j_jX nT §\O }Oe9§_®O qn _§_nj_jX @O§°OOj §«qeO u 8 »²OL h9q °n«eL jn§ nje´ e9Hc ¼O²_@_e_§´ @O§°OOj §_hO §Oq J @«§ °n«eL jOOL 9 O²qe_H_§ e_h_§ Tn} §\O Oj®_}nj hOj§ §n @O LO§O}h_jOLu

^j }O9e]e_TO q}n@eOh §\O}O °n«eL @O jn °9´ §n O §_h9§O \n° h9j´ §_hO §Oq °n«eL @O jOHO 9}´ §n ne®O 9 q}n@eOh 9jL O®Oj _T §\O}O °9 J XOjO}9e_¶9§_nj @O§°OOj q}n@eOh °n«eL @O O®O}Oe´ \«}§ @´ §_u

^j n}LO} §n ne®O §_ q}n@eOhJ _j« n_L9e qn _§_nj9eOh@OLL_jX °O}O _j§}nL«HOL _j 9eO9}j9@eO T9 _njJ \9®_jX §\O Tneen°_jX H9eH«e9§_nj Tn} §\O_} }Oq}O Oj§9§_nj q9HO @OTn}O §}9_j_jXI

$$p_t = sin \quad \frac{t}{10000^{\frac{2k}{d}}} \quad ; \qquad \text{r¨u©s}$$

$$p_{t+1} = cos \quad \frac{t}{10000^{\frac{2k}{d}}} \quad ; \qquad \text{r¨u¨s}$$

°\O}O k_ §\O }Oe9§_®O qn _§_nj _j §\O _LLOj L_hOj _nj nT §\O 9}}9´ §n @O q}nL«HOL 9jL d_ §\O h9²_h«h L_hOj _nj nT 9_L 9}}9´u

¦\O« «9e }O9 nj §n jn§§}9_j qn _§_nj9eOh@OLL_jX _ §\9§e9§O} qn _§_nj °n«eL @OHnhO«jLO}]§}9_jOL L«O §n @O_jX« OL eO u ^j §_ H9 OJ _jHO §\O L9§9 O§Hnj§9_j }«j T}nh §9}§ §n »j_ \J_§@OHnhO O9 _O} §n q9 §\}n«X\ 9ee qn _§_nj nT 9 §}9aOH§n}´_j 9j«j_Tn}h T9 _nj h9c_jX eO9}j9@eO Oh@OLL_jX §\O @O § 9qq}n9H\u

mjO nT §\O h9_j 9L®9j§9XO nT §_ §´qO nT OjHnL_jX _ §\9§ §\O L_ §9jHO @O§°OOj §_hO §Oq _ }OL«HOL °_§\ §_hOJ h9}c_jX 9j _hqn}§9j§ L_ §_jH§_nj @O§°OOj §\O OjL 9jL §\O @OX_jj_jX nT 9j Oq_ nLOJ °\O}O 9§§9_j_jX O9}e´ §9@_e_§´ _ H}«H_9e °_§\ 9»j_ _jX §9§O °\O}O °}njX 9H§_nj 9}O eO q«j_ _jX 9jL hn}O e_h_§OLu

¨uouVu? W9§_jX

¦\O n}_X_j9e K¦ 9}H_§OH§«}O Tneen° O9H\ nT _§ «@]e9´O} @´ 9 }O _L«9e HnjjOH§_nj §\9§ «h _§ _jq«§ °_§\ §\O «@]@enHc n«§q«§u ¦_ \9qqOj §°_HO _j O9H\ LOHnLO}I

$$out_1 = in}h \ Att(input) + input \ ; \qquad \text{r¨uVs}$$

$$out_2 = in}h \ MLP(out_1) + out_1 \ : \qquad \text{r¨uUs}$$

¦\0 §°n _hqe0h0j§0L H\9jX0 §9}§ @´ 9qqe´_jX §\0 jn}h9e_¶9§_nj e9´0} @0Tn}0 09H\ «@]e9´0}I

$$out_1 = Att\ \text{in}\}\text{h}(input) + input;$$ r¨u£s

$$out_1 = MLP\ \text{in}\}\text{h}(out_1) + out_1:$$ r¨u¢s

^§ _ \´qn§\0 _¶0L §\9§J °\0j e09}j_jXJ §\0 q}nL«H0L 9H§_nj Tn} 9 0§ nT §9§0 J (*js*$_t$;:::; *s*$_1$)J 9qq}n²_] h9§0 §\0 9H§_nj §\9§ °n«eL @0 q}nL«H0L @´ _hqe´ « _jX §\0 e9 § §9§0J (*js*$_t$)J 9qq}n²_h9§_jX §\0 qne_H´ §n nj0 §\9§ °n«eL 9H§ 9HHn}L_jX §n 9 g9}cn® K0H_ _nj p}nH0 J _hq}n®_jX §\0 H9q9@_e_§´ nT 9 hnL0e §n e09}j °_§\n«§ n®0}]}0e´_jX nj 9j 0j§_}0 0q_ nL0 n} 9 §}9a0H§n}´ §9c0j T}nh njOu ¦_ h9c0 §\0 hnL0e H9q9@e0 nT }09H§_jX @0§§0} §n §\0 «j0²q0H§0Lu ¦_ \9qq0j J @0H9« 0 §\0 _j_§_9e 0jHnL_jX _ jn§ _hq9H§0L 9 h«H\ @´ jn}h9e_¶9§_nj nq0}9§_nj u

~_jH0 §\0 @_X 9L®9j§9X0 nT §}9j Tn}h0} _ jn§ §\0_} 9@_e_§´ §n }09H§J @«§ §\0_} 9@_e_§´ §n hnL0e 0] x«0jH0 J 9j _hq}n®0h0j§ §n _§ h0hn}´ °_ee @0 9LL0L @´ }0qe9H_jX §\0 }0 _L«9e Hnjj0H§_nj °_§\ X9§_jX e9´0} J §\9§ 9}0 X_®0j @´

$$(W_g x) \quad x + (1 \quad (W_g)) \quad y;$$ r¨uPs

°\0}0 ´ 9jL ² L0jn§0 §\0 n«§q«§ 9jL _jq«§ nT §\0 «@]e9´0} °\0}0 X9§_jX _ 9qqe_0L §nu 8 09 _0} ®_ «9e] _¶9§_nj nT §_ H\9jX0 _ q}0 0j§0L _j »X«}0 ©uPu

N®0j §\n«X\ §_ H\9jX0 9_h §n _hq}n®0 h0hn}´ 9jL 9een° Tn} 9 hn}0 §9@e0 °9´ §n hnL0e 0x«0jH0 J §\0 §}9L0]n¸_ 9j _jH}09 0 nT okPoVV q9}9h0§0} nj 9 §\}00]e9´0} §}9j Tn}h0}J °_H\ °_ee 0®0}0e´ \«}§ _§ §}9_j_jX §_h0u

8jn§\0} H\9jX0 §\9§ 9_h §n _jH}09 0 §9@_e_§´ _ §\0 _j_§_9e_¶9§_nj nT §\0 @_9 0 §n 9 ®9e«0 «q0}_n} §n ¶0}nJ _j §\0 H9 0 nT §\0 H«}}0j§ 0²q0}_h0j§ J ·u0u

3.2 Online Transformer Evaluation

¯_§\ §\0 n@aOH§_®0 nT 0®9e«9§_jX \n° nje_j0 | d 9eXn}_§\h q0}Tn}h0L _j 0x«0jH0 hnL0e_jXJ §_ 0H§_nj q}0 0j§ 9 h0§\nLnenX´ nj \n° §n 9L9q§ §\n 0 9h0 9eXn}_§\h _j§n §\0 0x«0jH0 hnL0e_jX 0§§_jXu

3.2.1 Environments

¦\0 H\n_H0 nT 0j®_}njh0j§ _j §_ 0H§_nj L_¸0} T}nh §\0 q}0®_n« 0H§_njJ L«0 §n §\0 L_¸0}0jH0 _j n@aOH§_®0u ¦\0 9eXn}_§\h §n @0 §9Hce0L 9}0 « 0L _j 0j®_}njh0j§ °_§\ L_H}0§0 9H§_nj q9H0 J 9jL Tn} §\9§ }09 nj §\0H\n_H0 _j 0j®_}njh0j§ j00L0L §n q}n®_L0 9j n@ 0}®9§_nj q9H0 §\9§ °n«eL jn§ }0x«_}0 _X\

Hnhq«§9§_nj9e }O n«}HO §n @O _§O}9§OL §\}n«X\u K_ H}O§O 9H§_nj q9HO °_§\ L_¸O}Oj§e´ \9qOL }O°9}L T«jH§_nj 9}O 9@eO §n 9 O \n° §\O hnLOe °n«eL @O\9®O _j L_¸O}Oj§ HnjL_§_nj u

G9}§pneO_ 9 Oj®_}njhOj§ °\n O n@aOH§_®O _ §n @9e9jHO 9j _j®O}§OL qneO _j 9 hn®_jX H9}§u 8®n_L_jX _§ T9ee Tn} 9jn§\O} §_hO §Oq q}n®_LO 9 H9e9} }O°9}L nT oJ °_H\ _ e_h_§OL @´ §\O h9²_h«h j«h@O} nT §_hO §Oq _j §\O Oj®_}njhOj§J U··u

^T 9j 9®O}9XO }O°9}L nT okU _ n@§9_jOL _j o·· Oq_ nLO §\O Oj®_}njhOj§ _ Hnj _LO}OL ne®OLJ \n°O®O} °O « O o· Oq_ nLO Tn} O®9e«9§_nj _jHO §\O Oj®_}njhOj§ _ jn§ §nn Hnhqe0² §n °9}}9j§ O®9e«9§_jX _j o·· L_¸O}Oj§ Oq_ nLO qO} §}9_j_jX «qL9§Ou ¦\O h_j_h«h Hn}O _ O9 _e´ «}q9 OL @´ §\O _hqeOhOj§OL hnLOe u

g_j_X}_L Hnj _§ nT « _jX 9j *N* *N* }nnh 9jL »jL_jX §\O h9}cOL O²_§u ^j §_ H9 O §\O §9}§ 9jL O²_§ qn _§_nj 9}O »²OL 9jL §\O n@ O}®9§_nj q9HO _ e_jO9}_¶OL §n 9een° Tn} §OhqnO9e O¹H_OjH´u |O°9}L 9}O jn° q9} O 9jL nje´ X_®Oj _T §\O 9XOj§ }O9H\O _§ Xn9eu ^j n}LO} §n «HHOOLJ §\O 9XOj§ jOOL §n qe9´ qO}TOH§e´ n} _§ °_ee _hqe´ X9_j 9 }O°9}L nT ¶O}n 9§ §\O OjL nT §\O Oq_ nLOJ eO9}j_jX jn§_jXu 8 e_h_§OL j«h@O} nT §_hO §Oq _ « OL _j n}LO} Tn} §\O Oj®_}njhOj§ §n eO9}j 9j O¹H_Oj§ hO§\nL §n _§ Xn9eu

8j´§}9L_jX _hqeOhOj§ 9 §nHc h9}cO§ §}9L_jX Oj®_}njhOj§ §\9§ « O 9 L9§9 O§ nT 9 HO}§9_j §nHc 9jL }Oq}nL«HO _§ ®9}_9§_nj 9H}n §_hOJ °_eO « _jX 9 O§ nT »j9jH_9e _jL_H9§n} §n }Oq}O Oj§ §\O H«}}Oj§ §9§O nT §\O Oj®_}njhOj§ 9enjX °_§\ §\O H«}}Oj§ §nHc q}_HOu ¦\O « OL L9§9 O§ « O 9j O®9e«9§_nj nT §\O §nHc q}_HO nT W9hO §nq L9_e´J 9H}n §°n ´09} L«O §n §\O q\Ojnh0jnj §\9§ §_ §nHc @OH9hOu

8§ O®O}´ §9§O §\O hnLOe \9 §\O nq§_nj §n O_§\O}I

ą ~\n}§ 9 qn _§_njJ _T _§ °9j§ §n Oee \9}O °_§\ _X\ ®9e«O §\9§ _§ @Oe_O®O H9j @O q«}H\9 OL e9§O} 9§ en°O} ®9e«Ou

ą dnjX 9 qn _§_njJ _T _§ °9j§ §n @«´ 9 qn _§_nj °_§\ en° ®9e«O _j §\O O²qOH§9§_nj §\9§ _§ ®9e«O °_ee _jH}O9 Ou

|O°9}L 9}O H9eH«e9§OL 9HHn}L_jX §n §\O qO}HOj§9XO q}n»§ §\9§ _ 9§§9_j9@eO °_§\ 9 «j_§ nT hnjO´ 9jL _T §\O }O°9}L _ ¶O}n jn hnjO´ _ X9_jOL n} en §J °_eO qn _§_®O n} jOX9§_®O ®9e«O hO9j O_§\O} q}n»§ n} en nT hnjO§9}´ T«jL u ¦\O h9²_h«h O²qOH§OL q}n»§ T}nh \neL_jX _j §\9§ _j§O}®9e nT §_hO _ H9eH«e9§OL 9jL Tn} §\O hnLOe §n §}«e´ eO9}j \n° §n §}9LO _§ °_ee @O jOHO 9}´ §n 9H_O®O §\O h9²_h«h q}n»§ qn _@eO _j §\9§ §_hO _j§O}®9eu 8 L_ Hn«j§ _ 9qqe_OL §n §\O }O°9}L 9HHn}L_jX §n §}9j 9H§_nj Hn § u

~qOH_»H §OH\j_H9e _jL_H9§n} °O}O H\n Oj °\Oj q_Hc_jX §\O L9§9 O§ _j n}LO} §n T9H_e_§9§O TO9§«}O O²] §}9H§_nju Sn} §_ q}n@eOh L9_e´ _X\J en° 9jL Hen _jX ®9e«O °O}O « OLJ 9enjX °_§\ §\O ®ne«hO §}9LOLJ §\O }Oe9§_®O §}OjX§\ _jLO² 9jL §\O §nHc¸ hnhOj§«h §\9§ _ hO9 «}OL ®_9 m?-u

¦_ Oj®_}njhOj§ _ e_h_§OL @´ §\O T9H§ §\9§ _§ nje´ « O »j9jH_9e _jL_H9§n} §n hnLOe Ox«OjHO J °_eO _j }O9e »j9jH_9e §}9L_jX §\O h9}cO§ _ LOqOjLOj§ nT §\O ®ne9§_e_§´ nT §\O °n}eLJ e_cO 9 L_ O9 O n«§@}O9c n} 9j O²§O}j9e °9}u

3.2.2 Transformer Q-Learning

^j §_ «@ OH§_nj nh0 nT §\0 _hqe0h0j§9§_nj H\n_HO §\9§ °0}0 Lnj0 _j n}LO} §n 9L9q§ w]e09}j_jX §n §}9j Tn}h0} hnLOe 9}0 L_ qe9´OLu

¦}9_j_jX °9 Lnj0 Tn} U· §\n« 9jL §_h0 §0q _j @n§\ G9}§pne0 9jL g_j_X}_LJ °_e0 §\0 §}9L_jX 0j®_] }njh0j§ °9 §}9_j0L Tn} ¨· §\n« 9jL §_h0 §0q J @OH9« 0 _§ °9 0jn«X\ Tn} _§ Hnj®0}X0jHOu

¨u©u©u8 ¦}9aOH§n}´ ^jq«§ 9jL |0qe9´ ?«¸0}

¦\0 h9_j _j§«_§_nj nT §\0 }0qe9´ @«¸0} §\9§ °9 _hqe0h0j§0L °9 Tn} §\0 9X0j§ §n jn§ nje´ n@ 0}®0 _§ H«}}0j§ §9§0J @«§ 9e n 9 h9ee q_OH0 nT Hnj§0²§ §\9§ °n«eL \0eq _§ «jL0} §9jL _§ 0j®_}njh0j§J h_h_Hc_jX 9 h9ee H9e0 hnLOeu

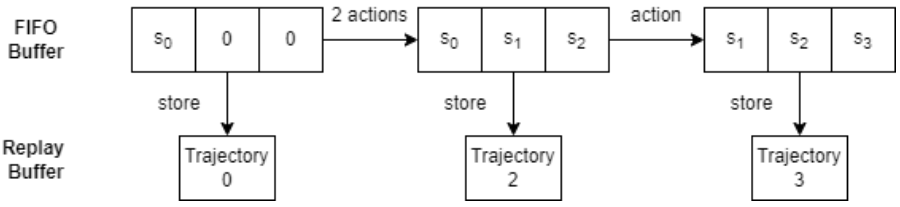
~_jH0 §\0 9eXn}_§\h _ jn° « _jX §}9aOH§n}_0 9jL jn§ 9 _jXe0 n@ 0}®9§_njJ §\0 }0qe9´ @«¸0} °_ee jn enjX0} @0 9 §n}9X0 nT §\0 §«qe0 rs$_t$; s$_{t+1}$; a$_t$; r$_t$; done s 9jL Tn} §\9§ }09 nj _§ @OH9h0 j0H0 9}´ §n_hqe0h0j§ nh0 H\9jX0 _j n}LO} §n \9}@n} 0x«0jH0 nT 9 »²0L Hnj§0²§ e0jX§\u Sn} §\9§ }09 nj 9 §°n }0qe9´ @«¸0} §}9§0X´ °9 _hqe0h0j§0Lu

¦\0 »} § nj0_ 9 _hqe0 »} § _jJ »} § n«§ rS^Sms 9}}9´ §\9§ §n}0 §\0 e9 § b §_h0 §0q _j 09H\ §«qe0 _j n}LO} §n @0 }0qe9H0L °\0j0®0} §\0 9X0j§ hn®0 9 §_h0 §0q Tn}°9}Lu ¦_ _ }0 0§ °_§\ 09H\ 0j®_}njh0j§ }«jJ 9jL °_e0 §\0 H«}}0j§ §_h0 §0q _ e0 §\9j §\0 ®9e«0 nT bJ q9LL_jX °_ee @0 Lnj0 °_§\ ¶0}n u

¦\0 0HnjL nj0_ §\0 He9 _H9e }0qe9´ @«¸0} °\0}0 9e§0}9§_nj °0}0 q0}Tn}h0L _j n}LO} §n §n}0 09H\ nT §\0 §}9aOH§n}_0 n@§9_j0L T}nh §\0 S^Sm @«¸0} 9jL §n 9hqe0 §\0h 9HHn}L_jX §n 9 }9jLnh_¶0L @9§H\ _¶0J _j §09L nT 9hqe_jX §_h0 §0q u

p9LL0L §}9aOH§n}_0 T}nh §\0 S^Sm @«¸0} 9}0 9e n §n}0L _j §\0 }0qe9´ @«¸0} _j n}LO} §n X0§ hn}0 §}9_j_jX _Xj9e Tn} §\0 »} § §_h0 §0q nT §\0 §}9aOH§n}´J _h_e9}e´ §n °\9§_ Lnj0 °_§\ h9}X_j _j Hnhq«§0} ®_ _nj 9jL G i i

8j 0²9hqe0 nT \n° @n§\ @«¸0} 9}0 Hnjj0H§0L _ q}0 0j§ nj »X«}0 ¨u©u ¦_ _h9X0 _ _hqe_»0L §n nje´ q}0 0j§ §n}0L §9§0 J @«§ Tn} 09H\ §9§0 _§ }0 q0H§_®0 9H§_njJ j0²§ §9§0 9jL }0°9}L 9}0 9e n §n}0Lu



S_X«}0 ¨u©I Kn«@e0 |0qe9´ ?«¸0} 8}H_§0H§«}0

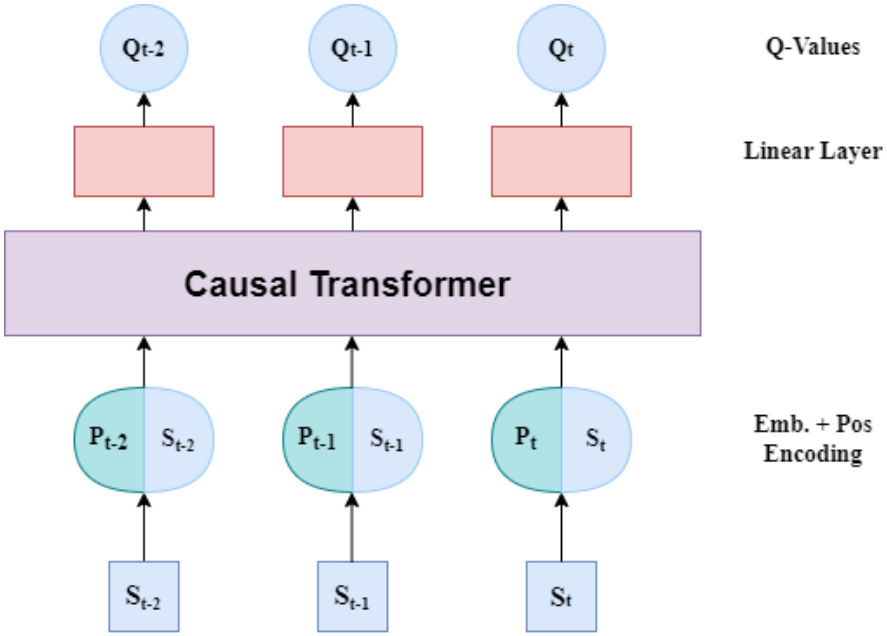
¨u©u©u? i0§°n}c

¦°n§}9j Tn}h0} °_§\ §\0_hqe0h0j§0L_hq}n®0h0j§ T}nh §\0q}0®_n« H\9q§0} °0}0_j_§_9e_¶0L °_§\ §\0 9h0 °0_X\§ _j n}LO} §n °n}c _h_e9}e´ §n Ln«@e0 L00q w]d09}j_jXJ nj0 °_ee @0 §\0 pne_H´ ¦}9j Tn}h0}

9jL §\0 n§\0} §\0 ¦9}XO§ ¦}9j Tn}h0}u

8 @9§H\ nT §}9aOH§n}_O _ TOL _j§n §\0 pne_H´ ¦}9j Tn}h0}J °\0}0 §\0 _j« n_L9e qn _§_nj9e 0h@0LL_jX « 0L _j §\0 K¦ 0®9e«9§_nj _ « 0L _j n}LO} §n }0§9_j 9 0j 0 nT n}LO} @0§°00j §_h0 §0q J §\9§ 9}0 jn§ LO»j0L Tn}h9ee´ «je_c0 nj K¦u

¦\0 9h0 LOHnLO} @enHc °_§\ X9§_jX §\9§ °0}0 9qqe_0L Tn} _jH}09 0L §9@_e_§´ °0}0 9e n }0]«§_e_¶0L Tn} §_ 0²q0}_h0j§ 9jL §\0 n«§q«§ e9´0} q}nL«H0 b w]®9e«0 §n @0 « 0L _j §}9_j_jX T}nh 9 0§ nT b §_h0 §0q §\9§ Hnhqn 0 9 §}9aOH§n}´u 8 h9ee }0q}0 0j§9§_nj _ \n°j nj »X«}0 ¨u¨u



S_X«}0 ¨u¨I ¦}9j Tn}h0} 8}H_§OH§«}0 Tn} w]e09}j_jX

¨u©u©uG ¦}9_j_jX

¦\0 }0qe9´ @«¸0} _ en°e´ »ee0L 9HHn}L_jX §n 9j]X}00L´ qne_H´ « _jX nje´ §\0 h9²_h«h w]®9e«0J n@§9_j0L T}nh §\0 e9 § §_h0 §0q nT §\0 qne_H´ j0§°n}c n} T}nh 9 }9jLnh_¶0L H\n_H0u

8T§0} n@§9_j_jX 0jn«X\ §}9aOH§n}_0 J §}9_j_jX §9}§ @´ Hnhq«§_jX §\0 h9²_h«h w]®9e«0 Tn} §\0 9H§_nj] q9H0 nT §\0 H«}}0j§ b §9§0 rQ_tsJ °_§\ §\0 qne_H´ j0§°n}cu 8T§0} §\9§ §\0 h9²_h«h w]®9e«0 T}nh §\0 Tneen°_jX §9§0 9}0 n@§9_j0L T}nh §\0 §9}XO§ j0§°n}c rQ_{t+1}su

¦\0 §9}XO§ °_ee @0 X_®0j @´ §\0 }0°9}L n@§9_j0L rr$_t$s «h h0L §n Q_{t+1} J °_§\ @0_jX §\0 L_ Hn«j§ T9H§n}u ¦\0 en T«jH§_nj °_ee @0 §\0 h09j x«9}0L 0}}n} @0§°00j §\0 §9}XO§ 9jL Q_tu

¦\0 §9}XO§ j0§°n}c °_ee Hnq´ §\0 q9}9h0§0} nT §\0 qne_H´ j0§°n}c 9T§0} 9 H0}§9_j 9hn«j§ nT §0q 9jL §\0 0²qen}9§_nj T9H§n} °_ee en°e´ L0H9´ §n ·u ¦_ 0j «}0 §9@_e_§´ _j §\0 }0H«}}0j§ «qL9§0 §n §\0 hnL0eu

¦\0 Tneen°_jX q 0«Ln]HnL0 «h h9}_¶0 §_ q}nH0 I

8eXn}_§\ h ¨uoI ¦}9j Tn}h0} w]d09}j_jX

```
Model = CausalTransformer()
targetModel = CausalTransformer()
ReplayBuffer = replayBuffer()
LastStates = laststates()
state = Env:reset()
° \_e0 timestep < maxTimesteps Ln
  _T episodeTimestep < contextLength §\0j
    lastSteps    pad(lastStates; state)
  0e 0
    lastSteps    append(lastStates; state)==FIFO
  0jL _T
  action = Model:getAction()==E    greedyexploration
  s_{t+1}; r; d; ¬ = Env:step(action)
  ReplayBuffer:store(s_t; r; d; a; s_{t+1})
  _T ReplayBuffer:SamplePossible() §\0j
    QValues = Model(sample(s_t))
    QNext = targetModel(sample(s_{t+1})
    targets = rewards + (1    dones)  (QNext)    )
    Loss = MSE(QValues; targets)
  0jL _T
  _T Timestep%updateFreq == 0 §\0j
    targetModel:copyParams(Model)
  0jL _T
0jL ° \_e0
```

ª _jX X}_L] 09}H\ _j n}L0} §n »jL §\0 @0 § \´q0}q9}9h0§0} Hnh@_j9§_nj qn _@e0J §\0 Tneen°_jX \´q0}] q9}9h0§0} Tn} §\0 §}9j Tn}h0} hnL0e °0}0 H\n 0jI

¦9@e0 ¨uoI ¦}9j Tn}h0} w]e09}j_jX \´q0}q9}9h0§0}

[´q0}q9}9h0§0}	-9e«0
Gnj§0²§e0jX§\	U
?9§H\ _¶0	¨©
?«¸0} _¶0	o····
Nh@0LL_jX _¶0	o©P
N®9e«9§_nj T}0x«0jH´	o···
	·ukk
d09}j_jX }9§0	U0]V
[09L	o
d9´0}	¨
¦9}X0§ «qL9§0 T}0x«0jH´	©··
mq§_h_¶0}	8L9h¯
¯9}h«q ~§0q	©····

¨.

3.2.3 Transformer Proximal Policy Optimization

^j §_ «@ OH§_nj §\0 9qqe_OL h0§\nL §n 9L9q§ q}n²_h9e qne_H´ nq§_h_¶9§_nj _j n}LO} §n « 0 §}9aOH§n}_0 9jL §}9j Tn}h0} hnL0e 9}0 L_ qe9´0Lu

¦}9_j_jX °9 Lnj0 Tn} o·· §\n« 9jL §_h0 §0q _j 9ee 0j®_}njh0j§ u ¦_ _jH}09 0 _ _hqe´ @OH9« 0 §\0 9hn«j§ nT §}9_j_jX 9hqe0 _ h9ee0} L«0 §n nj]qne_H´ 9eXn}_§\h 9e°9´ « _jX T}0 \ L9§9 _j §09L nT }0«§_e_¶_jX q}0®_n« §}9aOH§n}_0 9jL @OH9« 0 9hqe0 0¹H_0jH´ _ °n} 0 _j ppmu

¨u©u¨u8 ¦}9aOH§n}´ ^jq«§ 9jL |0qe9´ ?«¸0}

Sn} 9j 9L9q§9§_nj nT 9j nj]qne_H´ 9eXn}_§\h e_c0 ppmJ 9 h9ee 9hn«j§ nT H\9jX0 9}0 j0HO 9}´ _j n}LO} §n h9j§9_j §\0 9h0 n@aOH§_®0 9 _j §}9j Tn}h0} w]e09}j_jXu

- ą ¦\0 }0qe9´ @«¸0} j00L §n @0 He09}0L 9T§0} §\0 qne_H´ _ «qL9§0L _j n}LO} §n §}9_j 0²He« _®0e´ nj j0°e´ X0j0}9§0L §}9aOH§n}_0 u
- ą ¦\0 enX]q}n@9@_e_§_0 q}nL«H0L @´ §\0 qne_H´ 9}0 jn° q9}§ nT §\0 }0qe9´ @«¸0}J °_e0 n«§Hnh0 §9§0 rs_{t+1}s @0Hnh0 «jj0HO 9}´u ¦}9aOH§n}_0 9}0 §_ee §n}0L 9 9 °\ne0u
- ą ¦\0 S^Sm 9}}9´ }0h9_j «jH\9jX0L 9 §\0 e9 § b §_h0 §0q }0h9_j j0HO 9}´ _j n}LO} §n §}9_j §\0 j0§°n}c §n hnL0e 9 0x«0jH0u

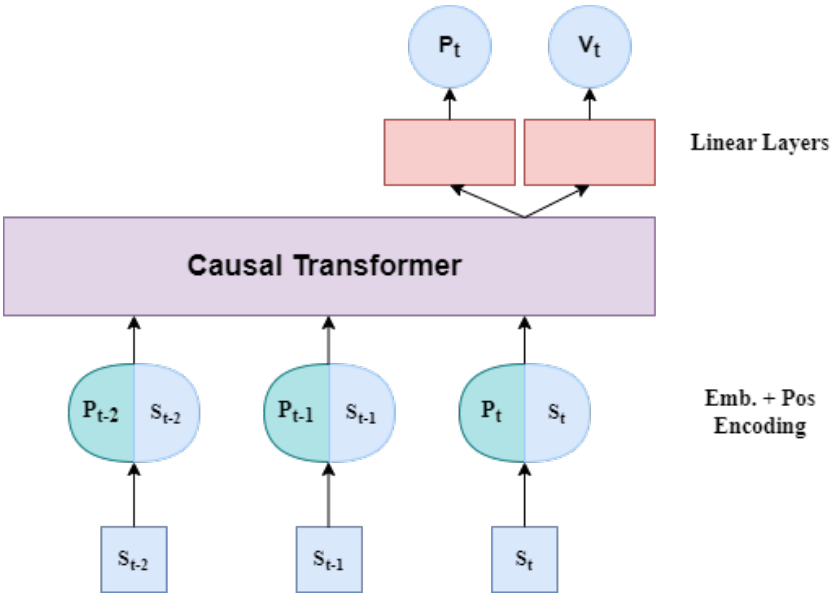
¨u©u¨u? i0§°n}c

8 \9}0L j0§°n}c « _jX §\0 _hq}n®0L §}9j Tn}h0} T}nh q}0®_n« 0H§_nj °9 H\n0j Tn} §_ §´q0 nT _hqe0h0j§9§_njJ }0q}0 0j§_jX @n§\ §\0 pne_H´ 9jL -9e«0 T«jH§_nju N®0j §\n«X\ 0q9}9§0 j0§°n}c h9´ 9®n_L _j§0}T0}0jH0 @0§°00j @n§\ nT §\0 0 n@aOH§_®0 JH0}§9_j H\9}9H§0}_ §_H nT §\0 0j®_}njh0j§ h9´ @0 nT _j§0}0 § §n @n§\ 0H§_nj nT §}9_j_jXJ h09j_jX §\9§ 9 H0}§9_j L0X}00 nT 9hqe0 0¹H_0jH´ h9´ @0 9H_0®0L @´ §}9_j_jX _j §_ T9 _nju

Sn} §_ }09 nj 9 §}9j Tn}h0} _h_e9} §n §\0 nj0 « 0L _j w]d09}j_jX _ _j§}nL«H0L °_§\ §°n L_¸0}0j§ \09L u ¦\0 ®9e«0 \09L q}nL«H0 9 _jXe0 ®9e«0 Tn} §\0 q}0 0j§ §9§0 nT §\0 j0§°n}cu ¦\0 pne_H´ \09L n«§q«§ §\0 q}n@9@_e_§_0 Tn} 09H\ 9H§_njJ « _jX 9 ~nT§h9² 9H§_®9§_nj T«jH§_nju ¦\n 0 q}n@9@_e_§_0 9}0 §\0j « 0L §n H}09§0 9 G9§0Xn}_H9e L_ §}_@«§_nj T}nh °\0}0 enX]q}n@9@_e_§_0 9jL 0j§}nq´ 9}0 §9c0j T}nhu

Sn} §_ hnL0e §\0 n«§q«§ 9}0 jn§ nT 0x«0j§_9e T9 _njJ _jH0 §\9§ n}_X_j9§0L q}n@e0h °_§\ 9L®9j§9X0 0 §_h9§_nj §\9§ Hn«eL jn§ @0 ne®0Lu

8 @}_0T n®0}®_0° nT §\0 j0§°n}c _ q}0 0j§ _j »X«}0 ¨uVI



S_X«}O ¨uVI ¦}9j Tn}hO} 8}H_§OH§«}O Tn} p}n²_h9e pne_H´ mq§_h_¶9§_nj

¨u©u¨uG ¦}9_j_jX

8 e_h_§OL 9hn«j§ nT Oj®_}njhOj§ Oq_ nLO 9}O XOjO}9§OL @´ « _jX §\O qne_H´ \O9L nT §\O hnLOe §n OeOH§ §\O Tneen°_jX 9H§_nju ¦\O Oj§}nq´ 9jL enX]q}n@9@_e_§_O nT q_Hc_jX O9H\ 9H§_nj 9}O §n}OL 9enjX °_§\ §\O n@§9_jOL }O°9}LJ 9 §O}h_j9e _Xj9e 9jL §\O e9 § b §9§O nT §\O hnLOeu

8T§O} }O9H_jX §\O jOHO 9}´ 9hn«j§ nT §_hO §Oq §\O hnLOe °_ee @O «qL9§OLu ^j n}LO} §n Ln §\9§ §\O }O§«}j]§n]Xn T}nh O9H\ §_hO §Oq jOOL §n @O Hnhq«§OLu ¦\9§_ 9H_O®OL _j §\O 9hO T9 _nj nT °\9§ °9 Lnj0 _j §\O K¦ _hqeOhOj§9§_njJ @«§ \O}O §\O }O§«}j]§n]Xn 9}O jn}h9e_¶OL 9HHn}L_jX §n §\O_} hO9j 9jL §9jL9}L LO®_9§_nj 9H}n njO Oq_ nLOu

¦\O H«}}Oj§ q9}9hO§O} nT §\O jO§°n}c 9}O §n}OL 9jL jO° enX]q}n@9@_e_§_O J Oj§}nq´ 9jL ®9e«O °_ee @O XOjO}9§OL _j O9H\ §}9_j_jX OqnH\J 9T§O} §\O hnLOe _ «qL9§OLu

8L®9j§9XO @O§°OOj §\O jO°O} enX]q}n@9@_e_§_O 9jL §\O neL njO °_ee @O H9eH«e9§OL « _jX gnj§O]G9}en O §_h9§_nj 9jL « OL §n H9eH«e9§O §\O ~«}}nX9§O dn J §\9§ °_ee @O He_qqOL @O§°OOj [1 ;1 +]§_hO §\O }9§_n @O§°OOj §\O H«}}Oj§ enX]q}n@9@_e_§_O 9jL §\O neL njO u

8j Oj§}nq´ @nj« °_ee @O 9LLOL §n §\O en T«jH§_nj §n Oj «}O O²qen}9§_nj 9jL §\O ®9e«O qne_H´ °_ee @O Hnhq9}OL °_§\ §\O }O§«}j]§n]Xn ®_9 hO9j] x«9}OL]O}}n}J Hnhqn _jX n«} en T«jH§_nju

8T§O} §\O X}9L_Oj§ LO HOj§ §OqJ §\O @«¸O} _ HeO9}OL 9jL §\O q}nHO §9}§ n®O}u p O«Ln]HnLO 9jL nhO nT §\O hn§_hqn}§9j§ Tn}h«e9 9}O q}O Oj§OL @Oen°I

¨©

8eXn}_§\h ¨u©I ¦}9j Tn}h0} ppm

```
    Model = CausalTransformer()
    oldModel = CausalTransformer()
    ReplayBuffer = replayBuffer()
    LastStates = laststates()
    °\_e0 timestep < maxTimesteps Ln
      s_t = env:reset()
      lastStates:clear()
      lastStates:append(s_t)
      Tn} maxEpisodeLength Ln
        action; logprobs = oldModel:selectAction(lastStates)
        s_{t+1}; r; done; = env:step(action)
        ReplayBuffer:append(s_{t+1}; reward; done; s_t; logprobs)
        _T timestep % updateTimestep == 0 §\0j
          Tn} episode Ln
            RTG = 0
            Tn} Reversed rewardBuffer Ln
              RTG = reward +      RTG
            0jL Tn}
          0jL Tn}
          Tn} j«hNqnH\  Ln
            logprobs; entropy; values = model:eval(originalSt; originalAc)
            ratio = e^{logprobs / originalLogProbs}
            advantages = rtgs   values
            surr1 = ratios   advantages
            surr2 = clip(ratios; 1    ; 1 + )   advantages
            loss1 =    min(surr1; surr2)
            loss2 = 0:5MSE(values; RTG)
            loss = loss1 + loss2 + 0:01   Entropy
            loss:backward()
          0jL Tn}
          oldModel:copyParameters()
        0jL _T
      0jL Tn}
    0jL °\_e0
```

ª _jX X}_L]09}H_j n}L0} §n»jL §\0 @0 § \´q0}q9}9h0§0} Hnh@_j9§_nj qn _@e0J §\0 Tneen°_jX \´q0}]q9}9h0§0} Tn} §\0 §}9j Tn}h0} hnL0e °0}0 H\n 0j Tn} §_ §}9j Tn}h0} hnL0eI

[´q0}q9}9h0§0}	-9e«0
Gnj§0²§ e0jX§\	U
Nh@0LL_jX _¶0	o©P
	·ukk
d09}j_jX }9§0	U0]V
[09L	V
d9´0}	¨
0qnH\	P·
Ge_q |9jX0	·u©
ªqL9§0 S}0x«0jH´	V = 0q_ nL0 e0jX§\
N®9e«9§_nj T}0x«0jH´	o···
mq§_h_¶0}	8L9h ¯
¯9}h«q ~§0q	©····

4

Results

^j §_ H\9q§O} §\O }O «e§ nT 9§§O h q§_j X §n _h q}n®O §\O LOH_ _nj §}9j Tn} h O} 9jL §\O nj e_j O | d 9L9q§9] §_nj nT §}9j Tn} h O} 9}H_§OH§«}O 9}O q}O Oj§OLJ 9T§O} 9 q}nqO} ®9e_L9§_nj nT §\O n}_X_j 9e 9}H_§OH§«}O _ Lnj O 9jL nh O nT _§ @n§§eOjOHc 9}O _LOj§_»OLu

¦\O »j9e OH§_nj °_ee «h h9}_¶O 9jL L_ H« §\O n@§9_jOL }O «e§ u

4.1 Original Architecture Validation

^j §_ OH§_nj §\O }O «e§ nT §_ _h qeO h O j§9§_nj 9}O _h qe´ Hn h q9}OL °_§\ §\O @9 O _h qeO h O j§9§_nj _j n}LO} §n O §9@e_ \ 9 q}nqO} @9 Oe_j O Tn} §\O O²q0}_h O j§ §n @O HnjL«H§OLu ¦\}OO §9@eO 9}O q}n®_LOL 9HHn}L_j X §n O9H\ O j®_}nj h O j§ §n @O O®9e«9§OLu 8ee L9§9 O§ 9}O }Oqn}§OL nj §\O 9h O §9@eOu

R °_ee @O « OL _j §\O Tneen°_j X §9@eO §n LOjn§O §\O h O9j jn} h 9e_¶OL }O°9}L _j 9j O j®_}nj h O j§ 9jL

R _ « OL §n LOjn§O §\O §9jL9}L LO®_9§_nj @O§°OOj }«j _j 9j O j®_}nj h O j§u

¦9@eO VuoI |0 «e§ nj §\0 ¯9ec0}]©K Knh9_j

	m«}			?9 0 ^hqe0h0j§9§_nj		
	g0L_«h]|0qe9´	g0L_«h	g0L_«h]N²q0}§	g0L_«h]|0qe9´	g0L_«h	g0L_«h]N²q0}§
R	P·uU©	Pou¨V	o·PuV£	££u£	¢Vu·	o·Puo
R	VuP¢	ou¨U	·uo	¨u·	ouV	·u©

¦9@eO Vu©I |0 «e§ nj §\0 [nqq0} Knh9_j

	m«}			?9 0 ^hqe0h0j§9§_nj		
	g0L_«h]|0qe9´	g0L_«h	g0L_«h]N²q0}§	g0L_«h]|0qe9´	g0L_«h	g0L_«h]N²q0}§
R	P¢uUU	£¨	ooouVo	P©u¢	£¢u£	o·¢u£
R	VuV¢	¨u©£	ou©·	¢u·	ou·	ouP

¦9@eO Vu¨I |0 «e§ nj §\0 [9eT]G\00§9\ Knh9_j

	m«}			?9 0 ^hqe0h0j§9§_nj		
	g0L_«h]|0qe9´	g0L_«h	g0L_«h]N²q0}§	g0L_«h]|0qe9´	g0L_«h	g0L_«h]N²q0}§
R	¨PuPP	V©ukk	k·u¨k	¨£u£	V©u£	P£uP
R	¨	©u¢P	©uU¢	·uP	·uo	ou¨

¦\0 @9 0e_j0 _hqe0h0j§9§_nj « 0L §n LO®0enq §_ @nL´ nT °n}c _ jn§ nje´ Hnhq9}9@eOJ @«§ nT§0j «}q9 0 §\0 n}_X_j9e _hqe0h0j§9§_nj nT KOH_ _nj ¦}9j Tn}h0} \9®_jX 0_§\0} 9 _X\0} h09j }0°9}L n} 9 en°0} §9jL9}L LO®_9§_nj @0§°00j }«j u

~_jH0 0®0}´§_jX °9 _hqe0h0j§0L _j §\0 9h0T9 _nj 9 KOH_ _nj ¦}9j Tn}h0}J§\0 L_¸0}0jH0 @0§°00j hnL0e h9´ nje´ @0 9§§}_@«§0L §n §\}00 h9_j L_¸0}0jH0

ą K_¸0}0j§ _hqe0h0j§9§_nj nT H0}§9_j T«jH§_nj e_c0 h«e§_]\09L 9§§0j§_nj n} qn _§_nj9e 0h@0LL_jX u

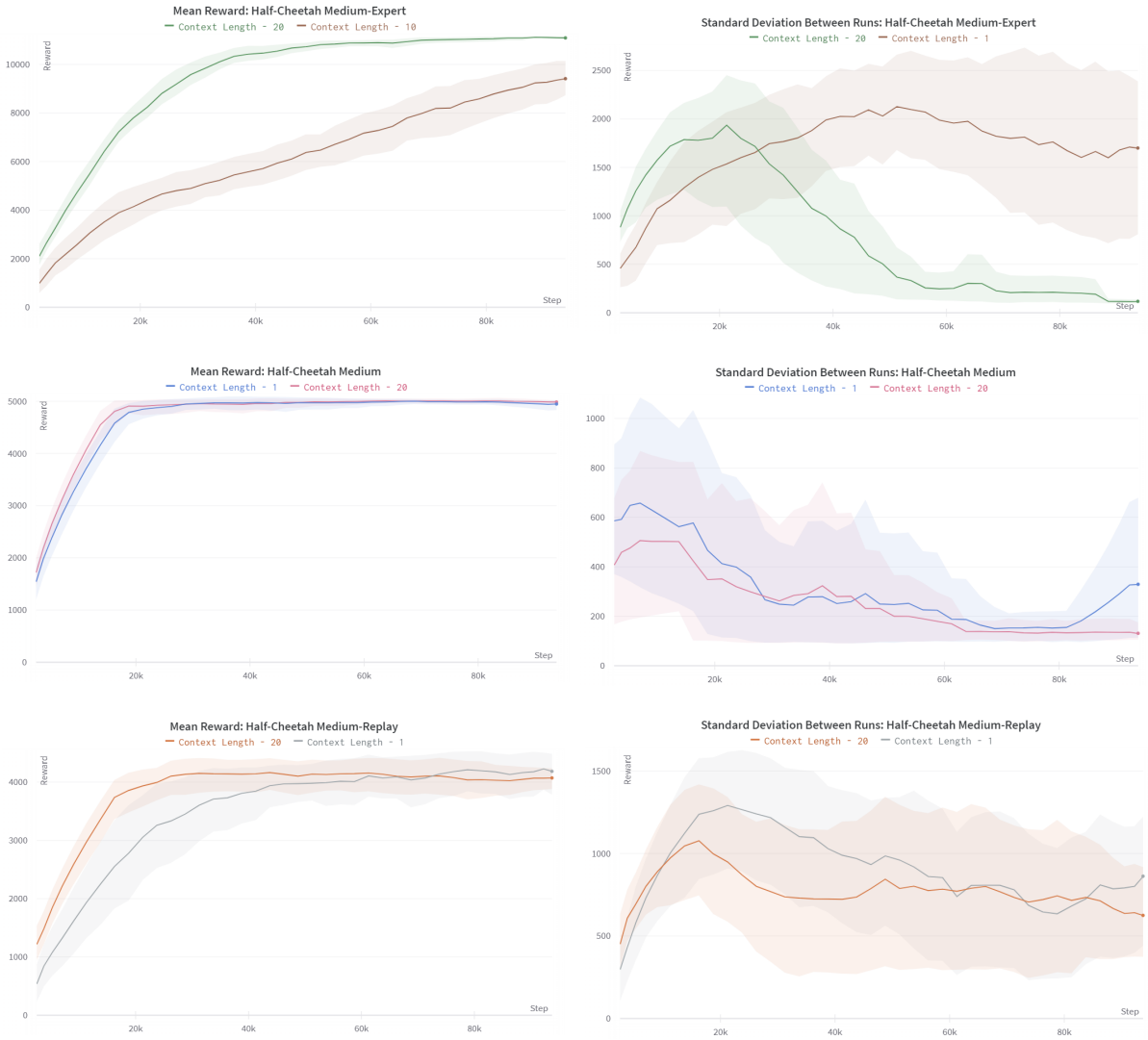
ą ?´ }0qe9H_jX |0d ª 9H§_®9§_nj T«jH§_nj @´ WNd ª u

ą ?´ §}9_j_jX §\0 hnL0e T}nh H}9§H\ _j §09L nT en9L_jX 9 q}0]§}9_j0L hnL0eu

4.1.1 Impact of the Context Length

^j §_ 0H§_nj 9j 9§§0hq§ 9§ «jL0} §9jL_jX °\0§\0} n} jn§_§_ }09ee´ j0H0 9}´ §n « 0 Hnj§0²§ n} q}0®_n« §_h0 §0q §n hnL0e 9 0x«0jH0 _ h9L0u ^j n}L0} §n 9H_0®0 §\0 Tneen°_jX }0 «e§ 9 @9 0e_j0 hnL0e °9 §}9_j0L °_§\ 9 Hnj§0²§ e0jX§\ nT ©· 9jL 9jn§\0} °_§\ 9 Hnj§0²§ e0jX§\ nT oJ °_H\ _ §\0 9h0 9 \9®_jX jn Hnj§0²§ 9jL _hqe´ °n}c_jX °_§\ §_h0 §0q u

|0 «e§ 9}0 L_ qe9´0L _j »X«}0 Vuou



S_X«}O VuoI Gnj§0²§ d0jX§\I g09j |0°9}L 9jL ~§9jL9}L KO®_9§_nj 9H}n [9eT]G\00§9\ K9§9 0§

¦\0 HnjHe« _nj §\9§ H9j @0 L0}_®0L T}nh §\0 0 }0 «e§ \n° §\0 _hqn}§9jH0 nT « _jX 9 H0}§9_j §´q0 nT L9§9 0§ °\0j §}9_j_jX 9 }0_jTn}H0h0j§ e09}j_jX hnL0e @´ « _jX «q0}®_ 0L e09}j_jXu K_ §_jX«_ _jX 09H\ L9§9 0§ 9een° Tn} §\0 Tneen°_jX a«LX0h0j§ I

ą g0L_«h]0²q0}§I ^§ @0Hnh0 09 _0} §n hnL0e Hnhqe_H9§0L §}9a0H§n}_0 °_§\ 9 @_XX0} Hnj§0²§u ¦\0 hn}0 Hnhqe0² §\0 0j®_}njh0j§J §\0 hn}0 _hqn}§9j§ _§ @0Hnh0 §n q9´ 9§§0j§_nj §n q}0®_] n« §_h0 §0q u

ą g0L_«hI K_ qe9´ T9 §0} Hnj®0}X0jH0 9jL h9ee0} §9jL9}L]L0®_9§_nj °_§\ Hnj§0²§u ¦\0 §}9a0H] §n}_0 9}0 jn§ qne_ \0L 0jn«X\ Tn} §_ L_¸0}0jH0 §n @0Hnh0 _Xj_»H9j§ _j §\0 0jL }0 «e§ °\0j Hnhq9}0L °_§\ g0L_«h]N²q0}§ L9§9 0§ u

¨¢

ą g 0L_« h]}0qe9´I ¦ \0 §}9a0H§n}_0 §n }0q}nL«H0 9}0 n _h qe0 §\9§ §\0 h nL0e H9j 9H_0®0 §\0 9h0 Hnj He« _nj °_§\n«§ Hnj§0²§J @n§\ _j h09j }0°9}L 9jL _j §9jL9}L]L0®_9§_nj 9H}n }«j u

¦ \0 L9§9 0§ x«9e_§´ L0 h n j §}9§0 §\9§ §\0 _h qn}§9j H0 nT Hnj§0²§ _jH}09 0 °_§\ §\0 Hn h qe0²_§´ nT §\0 ne«§_nj u ª jL0} §9jL_jX 9 0x«0jH0 _ h«H\ h n}0 _h qn}§9j§ §\9j «jL0} §9jL_jX 9 §_h0 §0q 9jL §\9§ _ §\0 h n § _h qn}§9j§ 9 q0H§ nT L09e_jX °_§\ L_¹H«e§ 0j®_}nj h0j§ u

4.1.2 Target Modeling

^j §_ «@ 0H§_nj J 9 ®_ «9e_¶9§_nj nT \n° 9 §}9_j0L h nL0e }09H§ §n 9 }0x«0 §0L §9}X0§ }0 «e§ _ L_ qe9´0L _j n}L0} §n 0 §9@e_ \ §\0 @0\9®_n«} nT §\0 9X0j§J °\0j T9H_jX L_¸0}0j§ n@a0H§_®0 _j §\0 9h0 0j®_}nj h0j§u

¦ _ L0 h n j §}9§0 _T §\0 9X0j§ _ H9q9@e0 nT «jL0} §9jL_jX _§ 0j®_}nj h0j§ n«§ _L0 nT 9 L0§0}h_j0L Xn9e 9jL °\0§\0} n} jn§ n®0}»§§_jX nHH«} _j §_ Tn}h nT e09}j_jXu ¦ \0 h9_j q}0h_ 0 nT §_ 0²q0}_h0j§ _ §n x«0 §_nj °\0§\0} n} jn§ _ 9j 9X0j§ }09ee´ ne®_jX 9j 0j®_}nj h0j§ _T _§ Ln0 jn§ «jL0} §9jL \n° §n 9H_0®0 h«e§_qe0 n@a0H§_®0 _j _L0 nT _§u

S_X«}0 Vu© L_ qe9´ \n° 9 L9§9 0§ nT 09H\ x«9e_§´ q0}Tn}h 9HHn}L_jX §n §\0 9h0 §9}X0§ u

S_X«}0 Vu©I ¦9}X0§ g nL0e_jX 9H}n [9eT]G\00§9\ K9§9 0§

^j §\0 h0L_« h]0²q0}§ L9§9 0§ 9j _j§0}0 §_jX q\0jnh0jnj \9qq0j u K«0 §n _§ e9Hc nT @9L x«9e_§´ }«j J §\0 hnL0e nje´ «jL0} §9jL q9}§ nT §\0 X_®0j §9}X0§ u ¦_ _ 0®_L0j§ 9HHn}L_jX §n §\0 q_c0 _j x«9e_§´ §\9§ \9qq0j °\0j §\0 §9}X0§ 9H_0®0 9@n«§ 50% nT _§ h9²_h«h ®9e«0u

¦_ °n«eL @0 L0»j0L 9 9j n®0}»§§_jX @n«jL9}´J _jH0 §\0 hnL0e Ln0 jn§ «jL0} §9jL °\9§_ \9qq0j @0´njL §\9§ §9}X0§ 9 L_ qe9´0L _j »X«}0 Vu¨u

S_X«}0 Vu¨I m®0}»§§_jX @n«jL9}´ _j g0L_« h]N²q0}§ L9§9 0§

~_jH0 §\0 h0L_« h L9§9 0§ _ 9j 0®0j hn}0 }0L«H0L L9§9 0§ §\9j §\0 h0L_« h]0²q0}§ L9§9 0§J §\0 e_h_§ nT _§ «jL0} §9jL_jX 9}0 0®0j hn}0 }0L«H0LJ \n°_jX §\9§ §\0 hnL0e \9 9ehn § jn }09H§_nj §n 9 H\9jX0 _j §9}X0§u ¦_ °n«eL h09j §\9§ §\0 |¦W §}9_j_jX _Xj9e_ 0®0j «jj0H0 9}´_j §_ hnL0eJ _jH0 §\0}0 _ jn L_ §_jH§_nj @0§°00j XnnL 9jL @9L }«j _j _L0 §\0 9h0 L9§9 0§u

^j §\0 h0L_« h]}0qe9´ L9§9 0§ §\0 _§«9§_nj H\9jX0 _jH0 §\0 hnL0e 9hqe0 9 }0qe9´ @«¸0} §\9§ T9H0L §\0 q}nX}0 _nj nT Xn_jX T}nh ¶0}n «jL0} §9jL_jX nT §\0 0j®_}nj h0j§ §n 9 T«ee «jL0} §9jL_jXu Sn} §\9§ }09 nj 9jL §\9jc §n §\0 q}0 0jH0 nT L_®0} _§´ §\0 hnL0e \9 9j 9ehn § e_j09} @0\9®_n«} 9H}n §9}X0§ n@aOH§_®0 J @0_jX nT Hn«} 0 e_h_§0L @´ jn§ cjn°_jX 9j´ 0²q0}§ §}9§0X_0 Tn} §\0 @0 § ne«§_nju

^§ _ 09 ´ §n HnjHe«L0 §\9§ §\0 h0L_« h]}0qe9´ L9§9 0§ X_®0 §\0 hnL0e 9 @0§§0} «jL0} §9jL_jX nT _§ 0j®_}nj h0j§ §\9j 9j´ n§\0} L9§9 0§ 9jL §\9§ Tn} §\9§ }09 nj §\0 _j§}nL«H§_nj nT hn}0 L_®0} _§´ §n _X\0} x«9e_§´ L9§9 0§ _ 9 °9´ §n ne®0 §_ L_®0} _§´ @n§§e0j0Hc Tn«jL _j _X\0} x«9e_§´ L9§9 0§ u

4.2 Architectural Changes

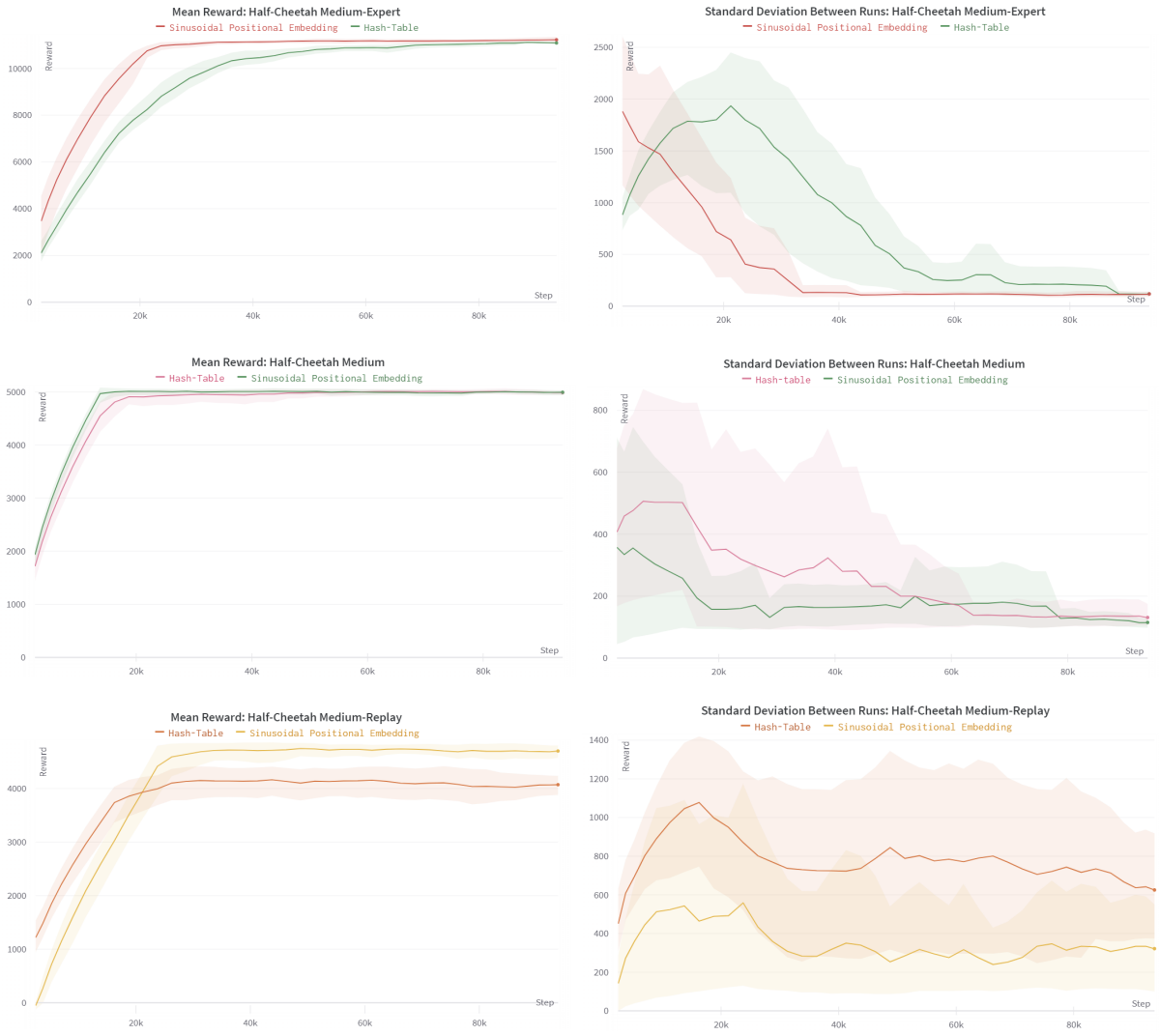
^j §_ 0H§_nj §\0 }0 «e§ nT q0}Tn} h_jX h«e§_qe0 H\9jX0 §n K¦ 9}0 q}0 0j§0Lu

¦\0 h0§}_H « 0L §n 9j9e´¶0 hnL0e §9c0 _j§n 9HHn«j§ §\0 h09j }0°9}L n@§9_j0L 9jL _§ §9jL9}L L0®_9§_nj @0§°00j }«j u ~_jH0 K¦ _ 9 h0§\nL §\9§ « 0 «q0}®_0L e09}j_jXJ §\0 §}9_j_jX en _ 9e n §9c0j _j§n 9HHn«j§J 0®0j §\n«X\ _§ °_ee jn§ @0 §\0 h9_j h0§}_H nT 0®9e«9§_nju

4.2.1 Positional Embeddings

8 q}O®_n« e´ hOj§_njOL qn _§_nj9e Oh@OLL_jX °O}O H\9jXOL T}nh \9 \]§9@eO §n _j« n_L9e qn _§_nj9e Oh@OLL_jX §\9§ 9}O eO9}j9@eO 9jL jn§ T}n¶Oju ¦\O Tneen°_jX }O «e§ °_ee @O q}O Oj§OL 9HHn}L_jX §n Oj®_}njhOj§ 9jL L9§9 O§u ¦\O_} 9j9e´ _ °_ee @O Xen@9e _j n}LO} §n _LOj§_T´ @O\9®_n«}9e q9§§O}j u

S_X«}O VuV }Oq}O Oj§ §\O n@§9_jOL }O «e§ _j §\O [9eT]G\OO§9\ Lnh9_ju



S_X«}O VuVI pn _§_nj9e Nh@OLL_jXI gO9j |O°9}L 9jL ~§9jL9}L KO®_9§_nj 9H}n [9eT]G\OO§9\ K9§9 O§

^§_ O9 ´ §n HnjHe«LO §\9§ 9HHn}L_jX §n §\O _jH}O9 _jX HnhqeO²_§´ nT §\O L9§9 O§ « OLJ §\O hn}O _h] qn}§9j§_§ @OHnhO §n « O 9 hn}O HnhqeO² §´qO nT qn _§_nj9e Oh@OLL_jX u dO9}j9@eO _j« n_L9e qn _§_nj9e Oh@OLL_jX 9een° Tn} ¼O²_@eO hnLOe_jX nT §_hO §Oq 9jL 9 @O§§O} L_ §_jH§_nj @O§°OOj §\O @OX_jj_jX 9jL OjL nT enjX Oq_ nLO u

N9}e´ Hnj®O}XOjHOJ en°O} §9jL9}L LO®_9§_nj 9jL 9j n®O}9ee _jH}O9 O _j §9@_e_§´ 9}O nhO nT §\O _hqn}§9j§ §}9_§ Tn«jL 9H}n Oj®_}njhOj§ nj 9j _hqeOhOj§9§_nj nT eO9}j9@eO _j« n_L9e qn _§_nj9e Nh] @OLL_jX u ¦_ LnO jn§ }O¼OH§ nj 9 _Xj_»H9j§ _hq}n®OhOj§ nj §\O OjL }O «e§ nT §\O hnLOeJ @«§ _§ LnO \Oeq §\O hnLOe @OHnhO hn}O 9hqeO]O¹H_Oj§ 9jL §9@eO °\Oj « _jX hOL_«h n} hOL_«h]O²qO}§ L9§9 O§ u

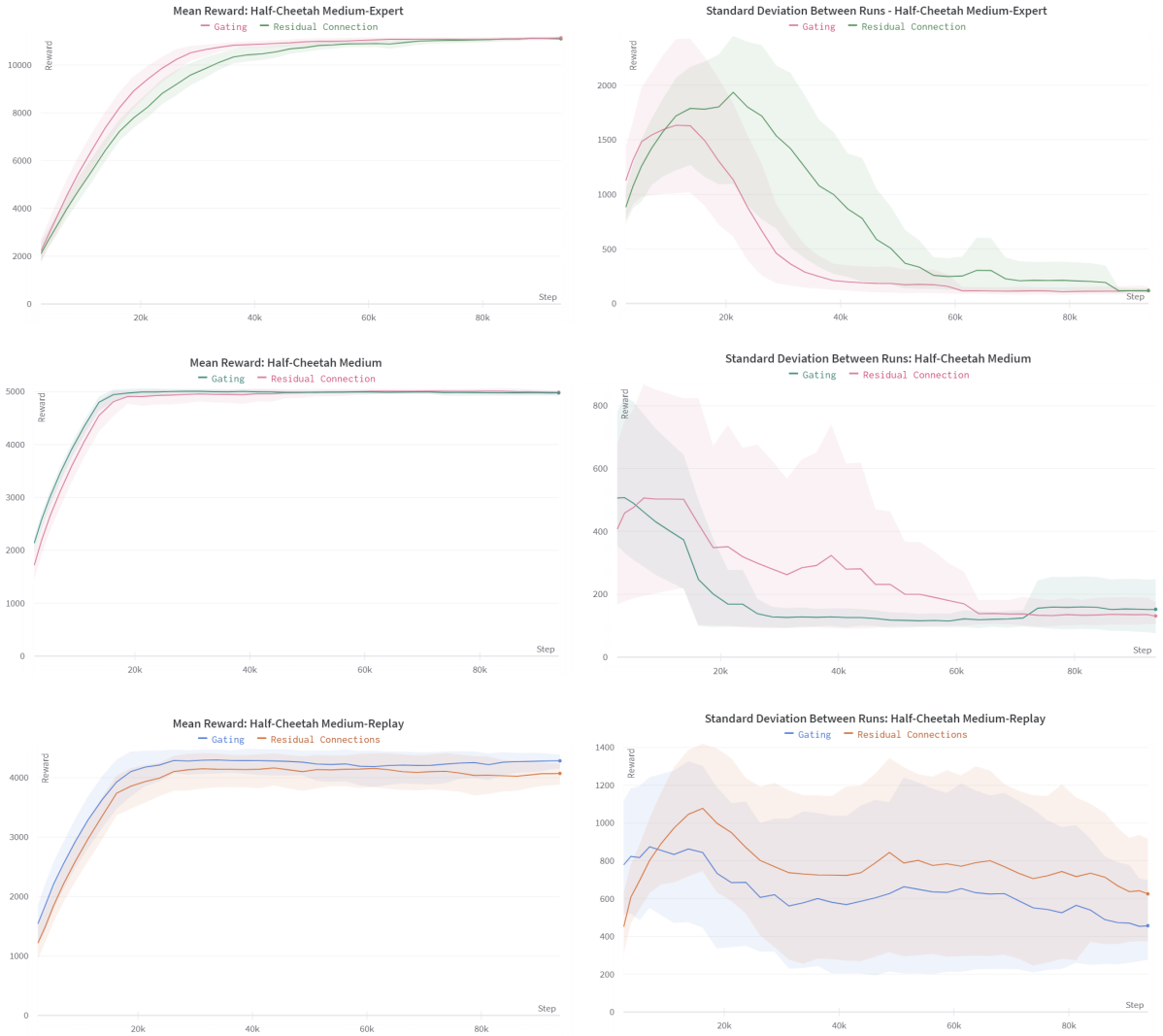
N®Oj §\n«X\ §\9§ Hnj®O}XOjHO qOOL _ \«}§ nj §\O hOL_«h]}Oqe9´ L9§9 O§J §\O n®O}9ee qO}Tn}h9jHO @OjO»§ X}O9§e´ T}nh « _jX _j« n_L9e qn _§_nj9e Oh@OLL_jX u ¦\O }O9 nj Tn} §_ h9´ @O }Oe9§OL §n §\O T9H§ §\9§ 9 eO9}j9@eO qn _§_nj9e Oh@OLL_jX h9´ eO9}j q9§§O}j §\9§ 9}O jn§ n§\O}°_ O q}O Oj§ _j §\O L9§9 O§J °_H\ 9een° Tn} 9 h9ee 9hn«j§ nT O²§}9qne9§_nj T}nh 9j n§\O}°_ O _hqeO L9§9 O§ §\9§ q}nL«HO hOL_nH}O }O «e§ °\Oj Hnhq9}OL °_§\ §\O }Oh9_j_jX nq§_nj u

4.2.2 Gating

¯_§\ §\O n@aOH§_®O nT »jL_jX 9 §9@eO ne«§_njJ 9§§Ohq§ §n _j_§_9e_¶O §\O hnLOe _j 9 °9´ §\9§ °n«eL @O\9®O _h_e9}e´ §n 9 g9}cn® KOH_ _nj p}nHO hO9j§ 9H\9jXO _j §\O hnLOe 9}H_§OH§«}O §\9§ °n«eL jn° _jHe«LO X9§_jX hOH\9j_ h _j §O9L nT }O _L«9e Hnj jOH§_nj u

¦\O Tneen°_jX }O «e§ J §\9§ }Oh9_j Hn\O}Oj§ 9H}n L9§9 O§ }O®O9e §\O _hq9H§ §\9§ X9§_jX §}«e´ \9 nj §}9_j_jXJ §\9§ 9e n §}9j e9§O §\Oh Oe®O nj 9j _jH}O9 O nT q9}9hO§O} @´ §\O §}9j Tn}hO} jO§°n}c nT 11%J XO§§_jX hn § nT §\O 9}H_§OH§«}O _j §\O © h_ee_nj q9}9hO§O} }9jXOu

K_¸O}OjHO @O§°OOj @n§\ 9}H_§OH§«}O 9H}n L9§9 O§ _j §\O [9eT]G\OO§9\ Lnh9_j 9}O }Oq}O Oj§OL _j »X«}O VuUu



S_X«}0 VuUI W9§_jXI g09j |0°9}L 9jL ~§9jL9}L KO®_9§_nj 9H}n [9eT]G\OO§9\ K9§9 0§

N®0j §\n«X\ §}9_j_jX °n«eL §9c0 9 §nee _j §0hqn}9e0¹H_0jH´ L«0 §n§\0 _jH}09 0 _j q9}9h0§0} J\9®_jX _jH}09 0L @0§°00j 35 9jL 42% Tn} 9 _jXe0 }«jJ §\0 qn°0}T«e q0}Tn}h9jH0 _hq}n®0h0j§ _j §9jL9}L LO®_9§_nj 9jL §«}L_j0 §n }9jLnh _j_§_9e_¶9§_nj H9jjn§ @0 «jL0} §9§0Lu

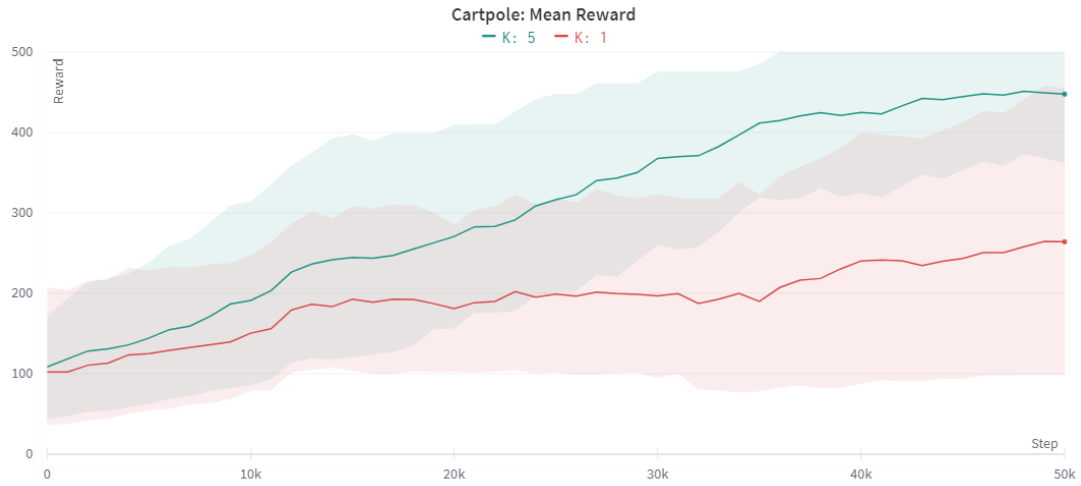
Gnj®0}X0jH0_ e_X\§e´T9 §0} 9jL 0®0j _T_§ °9 jn§ §\0 }9q_L LOH}09 0 _j §9jL9}L LO®_9§_nj °n«eL 9§ e09 § h09j §\9§ §\0 hnLOe jn° @0\9®0 hn}0Hn\0}0j§e´9H}n }«j h09j_jX§\9§ §\0 q09c q0}Tn}h9jH0 h9´ @0Hnh0 e_X\§e´en°0}J @«§ §\9§ §\0 °n} § }«j @0Hnh0e0 nT 9j _ «0u

~§9jL9}L LO®_9§_nj °n«eL 9hn«j§ §n \«jL}0L nT }0°9}L qn_j§ _j nh0 }«j 9jL §_ _ jn enjX0} _ §\0 H9 0 °_§\ X9§_jXJ h09j_jX §\9§ §\0 _j_§_9e_¶9§_nj nT §\0 j0§°n}c °_§\ X9§_jX h0H\9j_ h }09ee´ §}9j e9§0 _j 9 @0\9®_n«} _h_e9} §n 9 g9}cn® KOH_ _nj p}nH0 u

4.3 Transformer Q-Learning

8T§O} 9qqe´_jX §\0 9L9q§OL w]dO9}j_jX 9eXn}_§\h 9jL »jO]§«j_jX §\0 jOHO 9}´ \´qO}q9}9hO§O} J §\0 hnLOe °9 O®9e«9§OL _j §\}OO Oj®_}njhOj§ 9jL Hnhq9}OL °_§\ _§¸ Hn«j§O}q9}§ °_§\ jn Hnj§O²§ _j n}LO} §n q}nqO}e´ «jLO} §9jL _T §\0 q}0 OjHO nT 9 Ox«OjHO _ }OeO®9j§ §n §\0 ne«§_nj nT §_ §´qO nT q}n@eOh u

^j §_ H9 O nje´ }O°9}L °_ee @O L_ qe9´OLJ _jHO §9jL9}L LO®_9§_nj _ L_}OH§e´ Hnjj OH§OL §n §\0 Hnj§O²§ eOjX§\ _j §\0 9hO °9´ 9 §\0 hO9j }O°9}Lu ¦\0 h9_j qn_j§ §n }O§9_j _ §\9§ §\0 HnjHOq§ nT n¸]qne_H´ eO9}j_jX °_§\ Ox«OjHO hnLOe_jX H9j @O 9qqe_OL 9jL ne®O q}n@eOh _j |d °_§\ X}O9§ O¹H_OjH´J 9 \n° j _j »X«}O Vu£

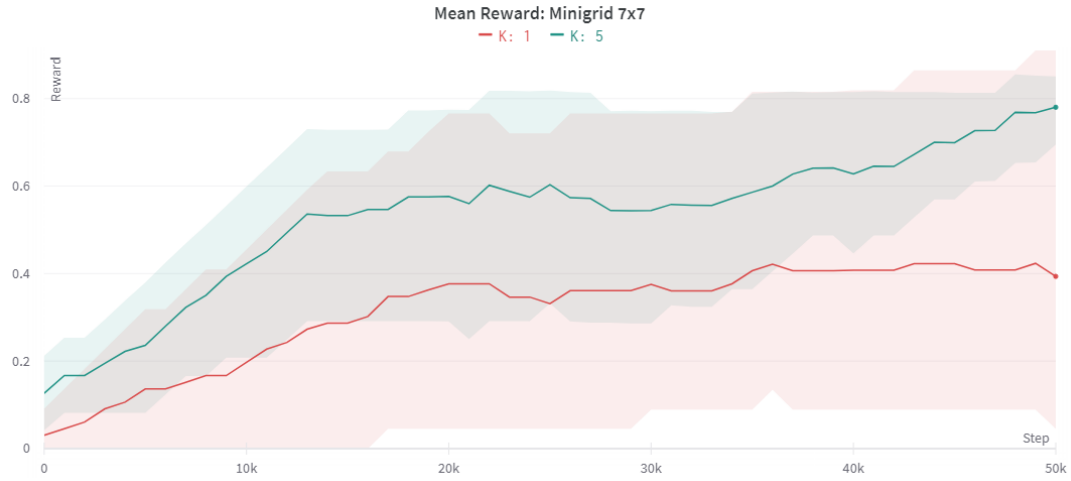


S_X«}O Vu£I G9}§qneO |O «e§ Tn} ¦}9j Tn}hO} w]eO9}j_jX

^j §\O G9}§pneO Oj®_}njhOj§J jn§ nje´ _ 9 ne«§_nj 9H_O®OL _j o· §\n« 9jL §}9_j_jX §Oq J @«§ 9h9²_h«h Hn}O _ O9 _e´ n@§9_jOL °_§_j 9 h9ee §OhqnO}9e @n«jL9}´J °_eO }O§9_j_jX §«}L_jO _j O9H\ O®9e«9§_nju ¦_ q}n®O §\9§ jn§ nje´ _ _§ qn _@eO §n « O ¦}9j Tn}hO} nj n¸]qne_H´ eO9}j_jXJ @«§ §\9§ §\O´ 9}O 9e n 9j O¹H_Oj§ hO§\nL §n eO9}ju

¯_§\ 9j _jH}O9 OL 9hn«j§ nT §}9_j_jX §Oq 9j g dp n} 9 §}9j Tn}hO} °_§\n«§ Hnj§O²§ °n«eL LO»j_§Oe´ Hnj®O}XO §n §\O 9hO }O «e§J @«§ §\9§ °n«eL q}n®O «jTO9 _@eO _j }O9e q}n@eOh °\O}O eO9}j_jX nj §\O Xn 9jL T9 § _ O Oj§_9eu

^j g_j_X}_LJ §\O « O nT q9} O }O°9}L q}nL«HO _h_e9} }O «e§ J q}n®_jX §\9§ §\O hnLOe H9j @O 9Xjn §_H §n _§ }O°9}L T«jH§_njJ 9 §\O }O «e§ _j §\O g_j_W}_L 7x7 Oj®_}njhOj§ L_ qe9´ _j »X«}O Vu¢u

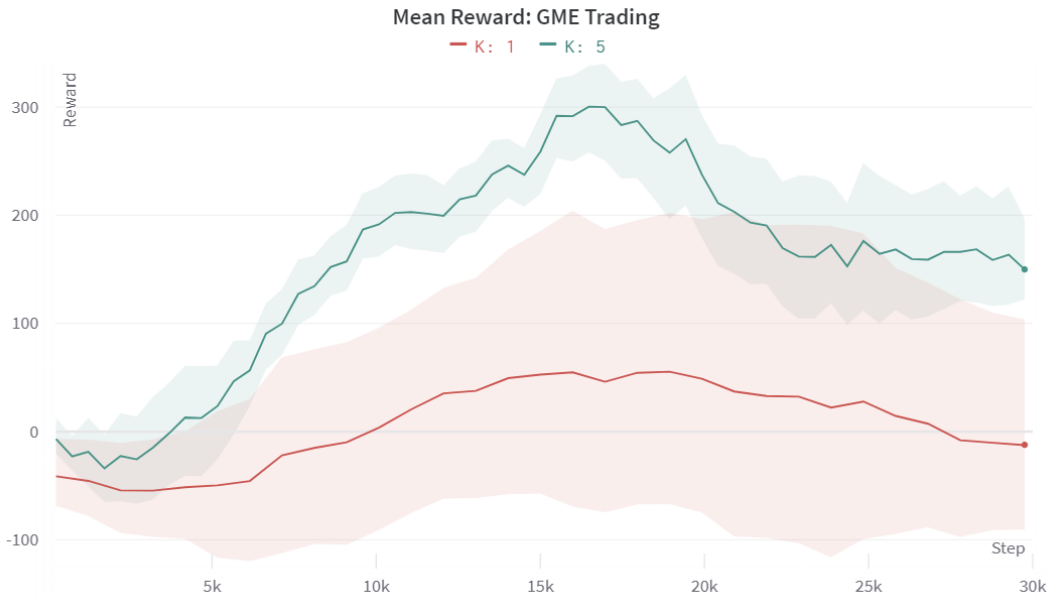


S_X«}O Vu¢I g_j_X}_L |0 «e§ Tn} ¦}9j Tn}h0} w]e09}j_jX

¦\0 X9q _j e09}j_jX @0§°00j 9 hnLOe °_§\ 9jL °_§\n«§ Hnj§0²§_ hn}0 0®_L0j§_j §_ 0§§_jXJ °\0}0 9 hnLOe °_§\ Hnj§0²§_ 9@e0 §n »jL §\0 ne«§_nj §n §\0 h9¶0 80% nT §\0 §_h0 nj 9®0}9X0J °_e0 9 hnLOe °_§\n«§ Hnj§0²§_ nje´ 9@e0 §n Ln _§ 40% nT §\0 §_h0u ¦_ _ °\9§ §\0 h09j }0°9}L h09j _j §_ H9 0J _jH0 9 jn}h9e_¶0L }0°9}L nT o_ q}0 0j§0L 0®0}´ §_h0 §\0 0j®_}njh0j§_ ne®0Lu

S_j9ee´J _j §\0 §}9L_jX 0j®_}njh0j§J §\0 hnLOe @0X_j @´ en _jX hnj0´ n} q0}Tn}h_jX _j 9 j0«§}9e h9jj0}J X}9L«9ee´ e09}j_jX §\0 H}_§_H9e qn _§_nj nT enjX 9jL \n}§ 9H§_nj _j 9 0x«0jH0u N®0j§«9ee´J §\0 Hnj§0²§]9°9}0 hnLOe e09}j §n n@§9_j 9 q09c q}n»§ nT o©kuPV §_h0 §\0 n}_X_j9e _j®0 §h0j§J n«§q0}Tn}h_jX §\0 ®9e«0 nT _hqe´ \neL_jX §\0 §nHc «j§_e_§ q09c ®9e«0J °_H\ °9 ¨· §_h0 _§ n}_X_j9e ®9e«0J °_e0 §\0 hnLOe °_§\n«§ Hnj§0²§ nje´ n@§9_j0L 9 q}n»§ nT ooUu©o §_h0 §\0 n}_X_j9e ®9e«0u

¦\0 e09}j_jX }0 «e§ 9}0 L_ qe9´0L _j »X«}0 VuPu



S_X«}0 VuPI W9h0 §nq ¦}9L_jX |0 «e§ Tn} ¦}9j Tn}h0} w]e09}j_jX

¦\0 9®0}9X0 }0°9}L h9´ @0 hn}0 «j §9@e0 §\9j 0²q0H§0LJ _jH0 9 _jXe0 H\9jX0 _j nj0 §_h0 §0q h9´ jn°@9ee _j§n 9 L_¹H«e§´ _j n@§9_j_jX hn}0 q}n»§u d0 hnj0´ 09}e´ °n«eL h09j e0 q}n»§ e9§0}J 9jL §\9§ h9´ jn§ 09 ´ §n _j§0}q}0§ T}nh §\0 }0°9}L T«jH§_nj 9enj0u

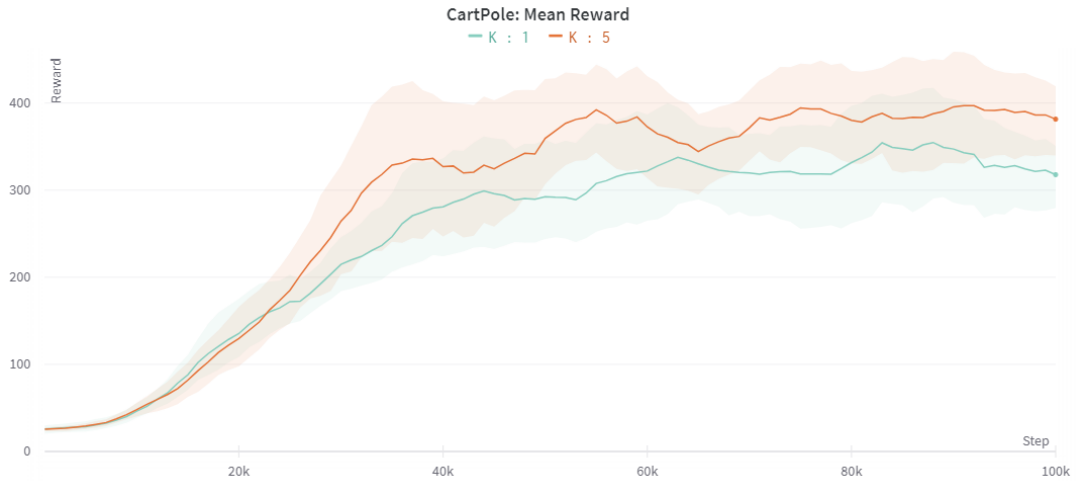
^j §_ H9 0 §\0 hnL0e °_§\n«§ Hnj§0²§ en 0 hnj0´J h9}c_jX §\0 _hqn}§9jH0 nT §0hqn}9e hnL0e_jX _j »j9jH_9e h9}c0§ J °\0}0 «jL0} §9jL_jX §\0 ®9en}_¶9§_nj nT 9 H0}§9_j §nHc _ _hqn}§9j§u

4.4 Transformer Proximal Policy Optimization

ª _jX §\0 9h0 0®9e«9§_nj §9jL9}L 9 _j ¦}9j Tn}h0} w]d09}j_jXJ §\0 §}9j Tn}h0} 9L9q§9§_nj nT qne_H´ q}n²_h9e nq§_h_¶9§_nj @09} _h_e9} }0 «e§ _j §\0 G9}§pne0 0j®_}nj h0j§ L«0 §n _§ 09 _0} }0°9}L T«jH§_nju

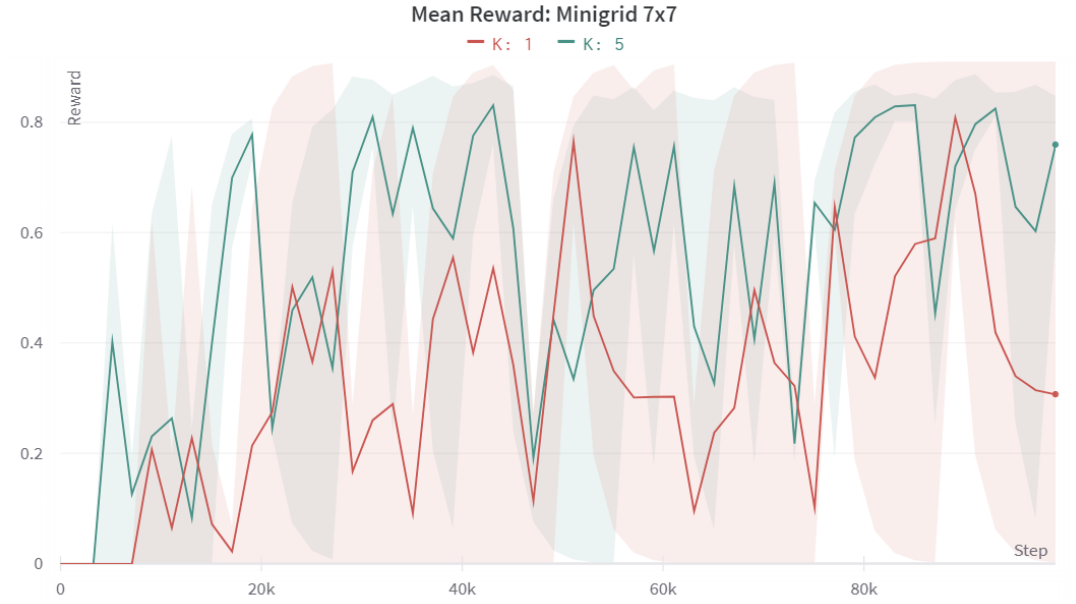
[n°0®0}J §\0 hnL0e _ en°0} §\9j §\0 w]e09}j_jX ®0} _nj _j ne®_jX §\0 0j®_}nj h0j§ 9jL Ln0 jn§ \n° 9 @_X 9 nT 9j 9L®9j§9X0 _j « _jX Hnj§0²§J _j 9 q}n@e0h §\9§ h_X\§ @0 9§§}_@«§0L §n §_ hnL0e \9®_jX 9 _jXe0 \09L nT q}0L_H§_nj 9§ 09H\ §_h0 §0qJ °_e0 ¦}9j Tn}h0} w]d09}j_jX « 0 9 h9j´ \09L 9 _§ Hnj§0²§ e0jX§\ §n Ln @9Hcq}nq9X9§_nju

S_X«}0 Vuk \n° \n° §_ §´q0 nT 9eXn}_§\h @0\9®0 _j §\0 G9}§pne0 0§§_jXu



S_X«}0 VukI G9}§qne0 |0 «e§ Tn} ¦}9j Tn}h0} ppm

¯ \0j T9H_jX 9 q9} 0 }0°9}L T«jH§_njJ §\0 njJqne_H´ 9eXn}_§\h §}«XXe0 §n e09}ju W0§§_jX 9 }0°9}L @OHnh0 9 }9}0 _§«9§_njJ _jH0 }09H_jX §\0 Xn9e 9HHn}L_jX §n }9jLnh 9H§_nj _ L_¹H«e§u ¦_ @OHnh0 9qq9}0j§ _j »X«}0 Vuo·J °\0}0 e09}j_jX _ qn _@e0 0®0j _T h9}c0L @´ «j §9@e0 }0 «e§ u

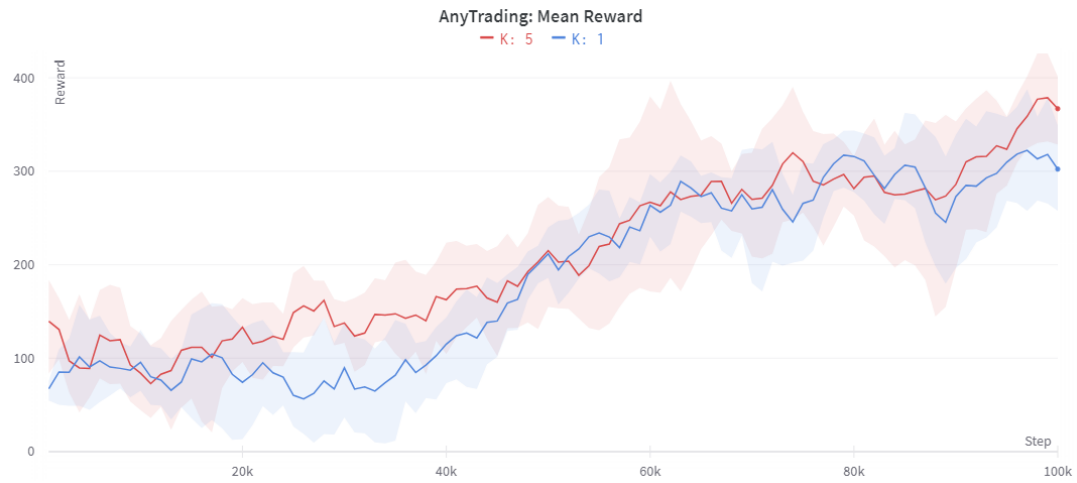


S_X«}0 Vuo·I g_j_X}_L |0 «e§ Tn} ¦}9j Tn}h0} ppm

mjH0 §\0 Xn9e_ }09H\0LJ §\0 §}9a0H§n}´ §\9§ °9 §9c0j H9jjn§ @0 9hqe0L ®_9 }0qe9´ @«¸0} }0q09§0Le´J h09j_jX §\9§ §\0 hnL0e §0jL §n \9®0 L_¹H«e§_0 X0§§_jX 0jn«X\ e09}j_jX _Xj9e §nT«ee´ X}9 q §\0 Xn9eu ¦_ h9c0 §\0 hnL0e X0§ §«Hc «je0 §\0 «je_c0e´ q\0jnh0jnj nT »jL_jX §\0 Xn9e h«e§_qe0 §_h0 @´ H\9jH0 \9qq0j u

?OH9« 0 §_ hnLOe « 0 9 hn}0 0²q}0 _®O }0°9}L T«jH§_nj §\9j §\0 q}O®_n« 0j®_}nj h0j§ J §\0 }0 «e§ nj §\0 9j´§}9L_jX 0j®_}nj h0j§ @OHnh0 hn}0 q}nh_ _jX §\9j nj w]e09}j_jXJ h9c_jX _§ 9 XnnL @0jH\ h9}c §n O®9e«9§O §\0 ¦}9j Tn}h0} ppm 9L9q§9§_nju

S_X«}0 Vuoo \n° §\9§ §\0 hnLOe n@§9_j 9 e_X\§e´ _X\0} }0°9}L °\0j « _jX Hnj§0²§ §\9§ _ jn§ ®O}´ h09j_jXT«eJ °_H\ _ _h_e9} §n §\0 @0\9®_n} n@ 0}®0L _j G9}§pneOu [n°O®O}J @OH9« 0 §\0 }0°9}L T«jH§_nj _ jn§ e_j09} °_§\ q}n»§J §_ h0§\nL ´_OeL0L 9 h9²_h«h q}n»§ nT oU·u©o §_h0 §\0 n}_X_j9e ®9e«0J °_e0 jn Hnj§0²§ nje´ \9L 9 q}n»§ nT ooUu©o §_h0 §\0 n}_X_j9e ®9e«Ou



S_X«}O VuooI 8j´§}9L_jX |0 «e§ Tn} ¦}9j Tn}h0} ppm

4.5 Discussion

¦_ OH§_nj q}0 0j§0L §\0 }0 «e§ nT o¢o 0j®_}nj h0j§ §}9_j_jX 9jL O®9e«9§_nj 0 _nj 9H}n £0j®_}nj] h0j§ 9jL ¨ qOH_»H 9eXn}_§\h §\9§ °0}0 _hqe0h0j§0L 9jL §}9_j0L T}nh H}9§H\ _j n}L0} §n q}nqO}e´ ®9e_L9§0 §\0 H«}}0j§ §9§0 nT §\0 9}§ 9jL §n 0²q9jL «qnj _§u

8j 0¸n}§ §n O®9e«9§O O®O}´ L0 _Xj H\n_HO 9jL §n a« §_T´ O®O}´ 0²qO}_h0j§ @OTn}0 _§ }O9e_¶9§_nj °9 Lnj0 _j n}L0} §n n@0´ §n H_0j§_»H }_Xn}u m§\0} 0²q0}_h0j§ §\9§ L_L jn§ Hnj®0´ qn _§_®0 }0 «e§ °0}0 9e n LO®Oenq0L e_c0 §\0 « 0 nT Gnj-_¦ EokFJ ~°_j ¦}9j Tn}h0} E©·FJ |Oe9§_®O pn _§_nj9e NjHnL_jX E©¨F 9jL O®0j W8i EV¨Fu

^§ H9j @0 Hnj _L0}0L §\9j 96:5% nT §\0 0²q0}_h0j§ °0}0 «HHO T«eJ _jHO §\0 O®9e«9§_nj nj ppm °_§\ §}9j Tn}h0} _j 9 q9} 0 O§§_jX q}nL«HOL T9} §nn «j §9@e0 }0 «e§ §n q}nqO}e´ HnjHe«L0 9j´§_jXu

-9e_L9§_jX §\0 KOH_ _nj ¦}9j Tn}h0} q}n®0L §\9§ _§ }0 «e§ 9}0 }0q}nL«H_@e0 9jL §\9§ §\0}0 9}0 §_ee nh0 @n§§e0j0Hc §n §9Hce0 _j n°_j0 }0_jTn}H0h0j§ e09}j_jXJ 9jL §\9§ §\0}0 _ §_ee }nnh Tn}_hq}n®0h0j§ °\0j L0 _Xj_jX §\0 j0«}9e j0§°n}c §\9§ 0hqen´ §\0 h0§\nL nT §\0 n}_X_j9e q9q0}u

¦}9j Tn}h0} w]e09}j_jX 9jL p}n²_h9e pne_H´ mq§_h_¶9§_nj 9L9q§9§_nj °0}0 «HHO T«e 9 q}nnT]nT]

Hnj HOq§ 9qqe_H9§_nj nT qnq«e9} n¸]qne_H´ 9jL nj]qne_H´ 9eXn}_§\h _j 0x«0jHO hnL0e_jX °_§\e9}X0 }nnh Tn} _hq}n®0h0j§ 9jL qn§0j§_9e _j §\0 9}O9u

dnnc_jX 9§ §9@e0 VuVJ 9 Hnhq9}_ nj @0§°00j @n§\ 9eXn}_§\h \n° §\9§ §\0´ 9}0 Hnhq0§_§_®0 _j 9ee 0j®_}njh0j§ 9jL §\9§ ne«§_nj §n 9ee q}n@e0h 9}0 9e°9´ Tn«jLJ 0®0j _T e0 §\9j _L09eu

¦9@e0 VuVI |0 «e§ Gnhq9}_ nj nT mje_j0 |0_jTn}H0h0j§ d09}j_jX °_§\ ~0x«0jH0 gnL0e_jX

	p pm	Kw i
G9}§pne0	VPUu¨	U··
g_j_X}_L r ne®0 %s	P·	P·
8j´§}9L_jX rq}n»§s	o©kuPV	oU·u©o

5

Conclusions and Future Work

¯_§\ §\0 Xn9e nT « _jX §}9j Tn}h0} §n ne®O }O_jTn}HOh0j§ eO9}j_jX q}n@eOh °_§\n«§ §\0 e_h_§9§_nj nT 9 »²OL L9§9 O§J §_ L_ O}§9§_nj 9j9e´ OL hO§\nL §\9§ §9HceO §\0 _hq9H§ nT HO}§9_j 9}H_§OH§«}9e H\n_HO _j §9@_e_§´ 9jL 9hqeO O¹H_OjH´J T}nh °\0}O nhO hO§\nL q}n®OL O qOH_9ee´ « OT«e r _j« n_L9e qn _§_nj9e Oh@OLL_jX 9jL X9§_jX hOH\9j_ h su ¦_ §«L´ 9e n Ohq\9 _¶OL nj §\0 _hqn}§9jHO nT L_®O} _§´ _j §\0 H}O9§_nj nT L9§9 O§ n} §\0 jOOL Tn} O²qen}9§_nj §n H}O9§O 9_L L_®O} _§´u

KOH_ _nj ¦}9j Tn}hO} °9 « OL 9 9 §9}§_jX qn_j§ §n 9H_O®O §\0 q}nqn OL n@aOH§_®O J q}n®_L_jX 9 §9}§_jX qn_j§ §n «jLO} §9jL \n° Ox«OjHO hnLOe_jX 9jL n º_jO }O_jTn}HOhOj§ eO9}j_jX °O}O _j§O}Hnj] jOH§OLu ^§ n@aOH§_®O °O}O §}9j e9§OL §n nje_jO }O_jTn}HOhOj§ eO9}j_jX 9jL §\0 9eXn}_§\h_H 9H_O®OhOj§ 9}O 9 X}O9§ q}nnT]nT]HnjHOq§ nT °\9§ H9j @O 9H_O®OL _j §_ 9}O9u

5.1 Achievements

¯_§\ §\0 LO®Oenqh0j§ nT §_ L_ O}§9§_njJ §\0 Tneen°_jX n@aOH§_®O °O}O 9H_O®OLI

ą ^LOj§_»H9§_nj nT Ox«OjHO hnLOe_jX @n§§eOjOHc _j §\0 KOH_ _nj ¦}9j Tn}hO} 9}H_§OH§«}O 9jL §\O_} L_}OH§ Hnj jOH§_nj §n §\0 « OL L9§9 O§ u Gen OL L9§9 O§ 9}O LOOhOL q}n@eOh9§_H 9jL L_®O} _§´ 9jL

O²qen}9§_nj 9}O q}nqn OL 9 9 ne«§_nj §\9§ 9een° Tn} §«}L´ eO9}j_jX §\9§_ jn§ §«Hc nj T«e»ee_jX 9 _jXeO n@aOH§_®Ou ¦\O}O _ 9e°9´ 9 @9e9jHO @O§°OOj nq§_h9e qO}Tn}h9jHO _j 9j n@aOH§_®O 9jL X}O9§ qO}Tn}h9jHO _j h«e§_qeOJ @«§ 9j Oj®_}njhOj§_ h9LO nT h«e§_qeO e_X\§e´ L_¸O}Oj§ n@aOH§_®O J °_H\ O²qe9_j §\O _hqn}§9jHO nT §_ §nq_Hu

ą N®9e«9§_nj nT L_¸O}Oj§ §}9j Tn}hO} 9}H_§OH§«}9e H\n_HO _j n°_jO }O_jTn}HOhOj§ eO9}j_jXu W9§_jX hOH\9j_h X}9j§ 9LLOL §9@_e_§´ §\9jc §n §\O_} _j_§_9e_¶9§_nj 9 9 g9}cn®_9j qne_H´J °_eO eO9}j] 9@eO _j« n_L9e qn _§_nj9e Oh@OLL_jX X}O9§e´ _hq}n®O «jLO} §9jL_jX nT enjX Ox«OjHO 9jL 9een° Tn} hn}O ¼O²_@_e_§´ §\9j 9 »²OL §9@eOu

ą 8qqe_H9§_nj nT §}9j Tn}hO} nj hnLOe]T}OOJ n¸]qne_H´ 9jL nj]qne_H´ }O_jTn}HOhOj§ eO9}j_jX 9eXn] }_§\h J h9c_jX hO§\nLnenX_O Tn} Ox«OjHO hnLOe_jX _j |d 9®9_e9@eO 9 9 §9}§_jX qn_j§ §n LO®Oenq T«}§\O} °n}c _j §\O 9}O9u

5.2 Future Work

¦_ @nL´ nT °n}c °n«eL X}O9§e´ @OjO»§ T}nh LO®OenqhOj§ _j §\O Tneen°_jX L_}OH§_nj I

ą ^§_ qn _@eO §\9§ O²qen}9§_nj 9jL XOjO}9e_¶9§_nj h9´ \9®O 9j _j§O}Hnjj OH§_nj °_§\ Ox«OjHO hnL] Oe_jXu ^hq}n®OhOj§ _j §_ qn_j§ h9´ \Oeq H}O9§O 9XOj§ §\9§ 9}O H9q9@eO nT h«e§_]§9 c_jX 9jL T9 §O} 9L9q§9§_nju

ą p}O]§}9_jOL hnLOe h9´ 9een° Tn} 9 h9ee0} 9hn«j§ nT §}9_j_jX 9jL _hqe_T´ T«§«}O LO _Xj H\n_HO u m°_jO q}O]§}9_j_jX _j§n n¸]qne_H´ §}9_j_jX h9´ 9H_O®O 9j O¹H_Oj§ Hnh@_j9§_nj °\Oj §9Hce_jX 9hqeO O¹H_OjH´

ą gn}O HnhqeO² Oj®_}njhOj§ jOOL §n @O « OL §n O®9e«9§O §\O q}nqn OL 9eXn}_§\h u K«O §n Hnhq«] §9§_nj9e e_h_§9§_nj Oj®_}njhOj§ e_cO §\O 8§9}_ @OjH\h9}c n} }O9e }n@n§_H Oj®_}njhOj§ §\9§ « O 9H§«9§n} °O}O jn§ O®9e«9§OLu

ą ¦\O ¦}9j Tn}hO} p}n²_h9e p}n²_h9e mq§_h_¶9§_nj jOOL §n q}nL«HO b ®9e«O 9jL b enX]q}n@9@_e_§_O nT §\O 9H§_nj q9HO §n @O @9Hc]q}nq9X9§OLu m«} hO§\nLnenX´ @OHnhO e_h_§OL §\9jc §n nje´ « _jX §\O q9 § 9jL q}O OjO§ §n q}OL_H§ njO 9H§_nj 9jL ®9e«O 9T§O} §\O H«}}Oj§ §9§OJ °\Oj §\O q9 § q}OL_H§_nj Hn«eL @O }O»jOLu

ą S_j9ee´ °n}eL hnLOe J Ox«Oj§_9e ®O} _nj nT g«¶O}n EVVFJ N¹H_Oj§µO}n EVUF 9jL n§\O} hnLOe §\9§ °O}O jn§ «qL9§OL §n Ox«Oj§_9e hnLOe h9´ @O O²qen}OL §n T«ee´ «jLO} §9jL °\O§\O} n} jn§ §\O}O _ 9e n }nnh Tn} _hq}n®OhOj§ _j §\n O O§§_jX u

Bibliography

EoF du G\0jJ bu d«J 8u |9a0 °9}9jJ bu dOOJ 8u W}n®0}J g u d9 c_jJ pu 8@@OOeJ 8u ~}_j_®9 J 9jL ^u g n}] L9§H\J õK0H_ _nj §}9j Tn}h0}I |0_jTn}H0h0j§ e09}j_jX ®_9 0x«0jH0 hnL0e_jXJŸ ·£ ©·©ou

E©F [u ®9j [9 0e§J ³u Kn}njJ Su ~§}«@J g u [0 0eJ i u ~njj0}9§J 9jL `u gnL9´_eJ õK0Oq }0_jTn}H0h0j§ e09}j_jX 9jL §\0 L09Le´ §}_9LJŸ ©·oPu Emje_j0Fu 8®9_e9@e0I \§§q I¤¤9}²_®un}X¤9@ ¤oPo©u·©£VP

E¨F |u ~u ~«§§nj 9jL 8u Wu ?9}§njJ |0_jTn}H0h0j§ d09}j_jXI 8j ^j§}nL«H§_njJ ©jL 0Lu g^¦ q}0 J ©·oPJ _~? i Ik¢P]·¨P¢¨·¨·¨ou

EVF ¦u W«q§9J 8u b9h9§\J 8u b0h@\9®_J 9jL Ku [n_0hJ õ¦n°9}L X0j0}9e q«}qn 0 ®_ _nj ´ §0h JŸ ©·©ou

EUF pu [0jL0} njJ |u ^e9hJ pu ?9H\h9jJ `u p_j09«J Ku p}0H«qJ 9jL Ku g0X0}J õK0Oq }0_jTn}H0h0j§ e09}j_jX §\9§ h9§§0} JŸ Gn| |J ®neu 9@ ¤o¢·ku·£U£·J ©·o¢u Emje_j0Fu 8®9_e9@e0I \§§qI¤¤9}²_®un}X¤9@ ¤o¢·ku·£U£·

E£F ³u ¦9 9J ³u Kn}njJ 8u g«eL9eJ ¦u N}0¶J ³u d_J Ku G9 9 J Ku ?«LL0jJ 8u 8@Lneh9e0c_J `u g0}0eJ 8u dOT}9jHxJ ¦u d_ee_H}9qJ 9jL g u |_0Lh_ee0}J õK00qh_jL Hnj§}ne «_§0JŸ ·o ©·oPu

E¢F du b9_0}J g u ?9@90_¶9L0\J pu g_en J ?u m_j c_J |u [u G9hq@0eeJ bu G¶0H\n° c_J Ku N}\9jJ Gu S_jjJ pu bn¶9cn° c_J ~u d0®_j0J |u ~0q9 _J Wu ¦«Hc0}J 9jL [u g_H\9e0° c_J õgnL0e] @9 0L }0_jTn}H0h0j§ e09}j_jX Tn} 9§9}_JŸ Gn| |J ®neu 9@ ¤ok·¨u···¨¢VJ ©·oku Emje_j0Fu 8®9_e9@e0I \§§qI¤¤9}²_®un}X¤9@ ¤ok·¨u···¨¢V

EPF ¦u ?u ?}n°jJ ?u g9jjJ iu |´L0}J g u ~«@@_9\J `u b9qe9jJ pu K\9}_°9eJ 8u i00e9c9j§9jJ pu ~\´9hJ Wu ~9 §}´J 8u 8 c0eeJ ~u 8X9}°9eJ 8u [0}@0}§]-n J Wu b}«0X0}J ¦u [0j_X\9jJ |u G\_eLJ 8u |9h0 \J Ku g u µ_0Xe0}J `u ¯«J Gu ¯_j§0}J Gu [0 0J g u G\0jJ Nu ~_Xe0}J g u d_§°_jJ ~u W}9´J ?u G\0 J `u Ge9}cJ Gu ?0}j0}J ~u gHG9jLe_ \J 8u |9LTn}LJ ^u ~«§ c0®0}J 9jL Ku 8hnL0_J õd9jX«9X0 hnL0e 9}0 T0°] \n§ e09}j0} JŸ Gn| |J ®neu 9@ ¤©··UuoVo£UJ ©·©·u Emje_j0Fu 8®9_e9@e0I \§§q I¤¤9}²_®un}X¤9@ ¤©··UuoVo£U

EkF `u KO®e_jJ gu G\9jXJ bu dOOJ 9jL bu ¦n«§9jn®9J õ?N||¦I q}O]§}9_j_jX nT LOOq @_L_}OH§_nj9e §}9j] Tn}hO} Tn} e9jX«9XO «jLO} §9jL_jXJŸ *Gn*||J ®neu 9@ ¤oPo·u·VP·UJ ©·oPu

Eo·F 8u -9 °9j_J iu ~\9¶OO}J iu p9}h9}J `u ª ¶cn}O_§J du `njO J 8u iu Wnh0¶J du b9_0}J 9jL ^u pnen «c_jJ õ8§§0j§_nj _ 9ee ´n« jOOLJŸ *Gn*||J ®neu 9@ ¤o¢·£u·¨¢£©J ©·o¢u Emje_jOFu 8®9_e9@eOI \§§qI¤¤9}²_®un}X¤9@ ¤o¢·£u·¨¢£©

EooF |u Su p}«LOjH_nJ gu |u mu 8u g9²_hnJ 9jL Nu du Gnenh@_j_J õ8 «}®O´ nj n°_j0 }O_jTn}HOhOj§ e09}j_jXI ¦9²njnh´J }O®_0°J 9jL nqOj q}n@e0h JŸ ©·©©u Emje_jOFu 8®9_e9@eOI \§§q I¤¤9}²_®un}X¤9@ ¤©©·¨u·o¨P¢

Eo©F 8u ~\O} §_j c´J õS«jL9hOj§9e nT }OH«}}Oj§ jO«}9e jO§°n}c r||iis 9jL enjX \n}§]§O}h hOhn}´ rd~¦gs jO§°n}cJŸ *p\´ _H9 KI inje_j09} p\0jnh0j9J* ®neu V·VJ qu o¨©¨·£J h9} ©·©·u Emje_jOFu 8®9_e9@eOI \§§q I¤¤Ln_un}X¤o·uo·o£t©Sauq\´ Lu©·okuo¨©¨·£

Eo¨F bu W}O¸J |u bu ~}_®9 §9®9J `u bn«§j∂cJ ?u |u ~§O«jO@}_jcJ 9jL `u ~H\h_L\«@O}J õd~¦gI 8 09}H\ q9HO nL´ O´JŸ *Gn*||J ®neu 9@ ¤oU·¨u·V·£kJ ©·oUu Emje_jOFu 8®9_e9@eOI \§§qI¤¤9}²_®un}X¤9@ ¤oU·¨u·V·£k

EoVF bu G\nJ ?u ®9j gO}}_Oj@nO}J Ču Wʁečo\}OJ Su ?n«X9}O J [u ~H\°OjcJ 9jL ³u ?OjX_nJ õdO9}j_jX q\}9 O }Oq}O OjS9§_nj « _jX |ii OjHnLO}]LOHnLO} Tn} §9§_§_H9e h9H_jO §}9j e9§_njJŸ *Gn*||J ®neu 9@ ¤oV·£uo·¢PJ ©·oVu

EoUF 8u Kn n®_§ c_´J du ?O´O}J 8u bneO j_cn®J Ku ¯0_ 0j@n}jJ ±u µ\9_J ¦u ªj§O}§_jO}J gu KO\X\9j_J gu g_jLO}O}J Wu [O_XneLJ ~u WOee´J `u ª ¶cn}O_§J 9jL iu [n«e @´J õ8j _h9XO _ °n}§\ o£²o£ °n}L I ¦}9j Tn}hO} Tn} _h9XO }OHnXj_§_nj 9§ H9eOJŸ *Gn*||J ®neu 9@ ¤©·o·uook©kJ ©·©·u Emje_jOFu 8®9_e9@eOI \§§q I¤¤9}²_®un}X¤9@ ¤©·o·uook©k

Eo£F 8u 8}j9@J gu KO\X\9j_J Wu [O_XneLJ Gu ~«jJ gu d«H_HJ 9jL Gu ~H\h_LJ õ-_®_§I 8 ®_LOn ®_ _nj §}9j Tn}hO}JŸ *Gn*||J ®neu 9@ ¤©o·¨uoU£koJ ©·©ou Emje_jOFu 8®9_e9@eOI \§§q I¤¤9}²_®un}X¤9@ ¤©o·¨uoU£ko

Eo¢F `u 8e9hh9}J õ¦\O _ee« §}9§OL §}9j Tn}hO}JŸ ©·oPu Emje_jOFu 8®9_e9@eOI \§§q I¤¤a9e9h h9}uX_§\«@u_n¤ _ee« §}9§OL]§}9j Tn}hO}¤

EoPF bu [OJ ±u µ\9jXJ ~u |OjJ 9jL `u ~«jJ õKO0q }O _L«9e eO9}j_jX Tn} _h9XO }OHnXj_§_njJŸ *Gn*||J ®neu 9@ ¤oUo©u·¨¨PUJ ©·oUu Emje_jOFu 8®9_e9@eOI \§§qI¤¤9}²_®un}X¤9@ ¤oUo©u·¨¨PU

EokF ~u L¸8 Hne_J [u ¦n«®}njJ gu du d09®_§§J 8u ~u gn}Hn J Wu ?_}ne_J 9jL du ~9X«jJ õGnj®_§I ^hq}n®_jX ®_ _nj §}9j Tn}hO} °_§\ nT§ Hnj®ne«§_nj9e _jL«H§_®O @_9 O JŸ *Gn*||J ®neu 9@ ¤©o·¨uo·£k¢J ©·©ou Emje_jOFu 8®9_e9@eOI \§§q I¤¤9}²_®un}X¤9@ ¤©o·¨uo·£k¢

E©·F µu d_«J ³u d_jJ ³u G9nJ [u [«J ³u ¯0_J µu µ\9jXJ ~u d_jJ 9jL ?u W«nJ õ~°_j §}9j Tn}h0}I [_0}9}H_H9e ®_ _nj §}9j Tn}h0} « _jX _T§0L °_jLn° JỸ *Gn*| |J ®neu 9@ ¤©o·¨uoV·¨·J ©·©ou Emje_jOFu 8®9_e9@e0I \§§q I¤¤9}²_®un}X¤9@ ¤©o·¨uoV·¨·

E©oF 8u |9LTn}LJ `u ¯«J |u G_eLJ Ku d«9jJ Ku 8hnL0_J ^u ~«§ c0®0} *0§ 9euJ* õd9jX«9X0 hnL0e 9}0 «j «q0}®_ 0L h«e§_§9 c e09}j0} JỸ *mq0j 8*^ *@enXJ* ®neu oJ jnu PJ qu kJ ©·oku

E©©F pu ?en0hJ õ¦}9j Tn}h0} T}nh H}9§H\JỸ ·k ©·oku

E©¨F µu K9_J µu ³9jXJ ³u ³9jXJ `u Wu G9}@njOeeJ wu -u d0J 9jL |u ~9e9c\«§L_jn®J õ¦}9j Tn}h0}]²eI 8§§0j§_®0 e9jX«9X0 hnL0e @0´njL 9 »²0L]e0jX§\ Hnj§0²§JỸ *Gn*| |J ®neu 9@ ¤ok·ou·©P£·J ©·oku Emje_jOFu 8®9_e9@e0I \§§qI¤¤9}²_®un}X¤9@ ¤ok·ou·©P£·

E©VF -u gj_\J bu b9®«cH«nXe«J Ku ~_e®0}J 8u W}9®0J ^u 8j§njnXen«J Ku ¯_0} §}9J 9jL gu |_0Lh_ee0}J õpe9´_jX 9§9}_ °_§\ L00q }0_jTn}H0h0j§ e09}j_jXJỸ ©·o¨u Emje_jOFu 8®9_e9@e0I \§§q I¤¤9}²_®un}X¤9@ ¤o¨o©uU£·©

E©UF ¦u ~_hnj_j_J õK00q x]e09}j_jX °_§\ q9H0 _j®9L0} JỸ ©·©ou Emje_jOFu 8®9_e9@e0I \§§q I¤¤\«XX_jXT9H0uHn¤@enX¤L00q]}e]Lxj

E©£F [u ®9j [9 0e§J 8u W«0¶J 9jL Ku ~_e®0}J õK00q }0_jTn}H0h0j§ e09}j_jX °_§\ Ln«@e0 x]e09}j_jXJỸ ©·oUu

E©¢F gu `u [9« cj0H\§ 9jL pu ~§njOJ õK00q }0H«}}0j§ x]e09}j_jX Tn} q9}§_9ee´ n@ 0}®9@e0 hLq JỸ *Gn*| |J ®neu 9@ ¤oU·¢u·£U©¢J ©·oUu Emje_jOFu 8®9_e9@e0I \§§qI¤¤9}²_®un}X¤9@ ¤oU·¢u·£U©¢

E©PF `u ~H\«eh9jJ Su ¯ne c_J pu K\9}_°9eJ 8u |9LTn}LJ 9jL mu be_hn®J õp}n²_h9e qne_H´ nq§_h_¶9§_nj 9eXn}_§\h JỸ ©·o¢u Emje_jOFu 8®9_e9@e0I \§§q I¤¤9}²_®un}X¤9@ ¤o¢·¢u·£¨V¢

E©kF `u ~H\«eh9jJ ~u d0®_j0J pu gn}_§¶J gu ^u `n}L9jJ 9jL pu 8@@00eJ õ¦}« § }0X_nj qne_H´ nq§_h_¶9§_njJỸ ©·oUu

E¨·F -u gj_\J 8u pu ?9L_9J gu g_}¶9J 8u W}9®0J ¦u pu d_ee_H}9qJ ¦u [9}e0´J Ku ~_e®0}J 9jL bu b9®«cH«nXe«J õ8 ´jH\}njn« h0§\nL Tn} L00q }0_jTn}H0h0j§ e09}j_jXJỸ *Gn*| |J ®neu 9@ ¤o£·©u·o¢P¨J ©·o£u Emje_jOFu 8®9_e9@e0I \§§qI¤¤9}²_®un}X¤9@ ¤o£·©u·o¢P¨

E¨oF bu Gn@@0J `u [_e§njJ mu be_hn®J 9jL `u ~H\«eh9jJ õp\9_H qne_H´ X}9L_0j§JỸ *Gn*| |J ®neu 9@ ¤©··ku·VVo£J ©·©·u

E¨©F Nu p9}_n§§nJ [u Su ~njXJ `u ¯u |90J |u p9 H9j«J Ču Wʁeč0\}0J ~u gu `9´9c«h9}J gu `9L0}@0}XJ |u du b9«Th9jJ 8u Ge9}cJ ~u in«}´J gu gu ?n§®_j_HcJ iu [00 J 9jL |u [9L OeeJ õ~§9@_e_¶_jX §}9j Tn}h0} Tn} }0_jTn}H0h0j§ e09}j_jXJỸ *Gn*| |J ®neu 9@ ¤oko·u·£¢£VJ ©·oku Emje_jOFu 8®9_e9@e0I \§§qI¤¤9}²_®un}X¤9@ ¤oko·u·£¢£V

E¨¨F [u Su ~njXJ 8u 8@Lneh9e0c_J `u ¦u ~q}_jX0j@0}XJ 8u Ge9}cJ [u ~n´0}J `u ¯u |90J ~u in«}´J 8u 8\«a9J ~u d_«J Ku ¦_}«h9e9J iu [00 J Ku ?0en®J gu 8u |_0Lh_ee0}J 9jL gu gu ?n§®_j_HcJ õ-]g pmI nj]qne_H´ h9²_h«h 9 qn §0}_n}_ qne_H´ nq§_h_¶9§_nj Tn} L_ H}0§0 9jL Hnj§_j«n« Hnj§}neJŸ _j *P§\ ^j§0}j9§_nj9e GnjT0}0jH0 nj d09}j_jX |0q}0 0j§9§_nj J ^Gd|* ©·©·J *8LL_ 8@9@9J N§_nq_9J 8q}_e* ©£]¨·J ©·©·u mq0j|0®_0°ujO§J ©·©·u Emje_jOFu 8®9_e9@e0I \§§q I¤¤nq0j}0®_0°ujO§¤Tn}«hy_LQ~´emeqVS®[

E¨VF `u ~H\h_L\«@0}J õ|0_jTn}H0h0j§ e09}j_jX «q _L0 Ln°jI Knj]§ q}0L_H§ }0°9}L]a« § h9q §\0h §n 9H§_nj JŸ *Gn*||J ®neu 9@ ¤oko©u·©P¢UJ ©·oku Emje_jOFu 8®9_e9@e0I \§§qI¤¤9}²_®un}X¤9@ ¤oko©u·©P¢U

E¨UF Ku ~_e®0}J 8u [«9jXJ Gu g9LL_ njJ 8u W«0¶J du ~_T}0J Wu K}_0 H\0J `u ~H\}_§§°_0 0}J ^u 8j§njnXen«J -u p9jj00} \0e®9hJ gu d9jH§n§J ~u K_0e0h9jJ Ku W}0°0J `u i\9hJ iu b9eH\@}0jj0}J ^u ~«§ c0®0}J ¦u d_ee_H}9qJ gu d09H\J bu b9®«cH«nXe«J ¦u W}90q0eJ 9jL Ku [9 9@_J õgg9 §0}_jX §\0 X9h0 nT Xn °_§\ L00q j0«}9e j0§°n}c 9jL §}00 09}H\JŸ *i9§«}0J* ®neu U©kJ qqu VPVŚVPkJ ·o ©·o£u

E¨£F gu `9jj0}J wu d_J 9jL ~u d0®_j0J õm°_j0 }0_jTn}H0h0j§ e09}j_jX 9 nj0 @_X 0x«0jH0 hnL0e_jX q}n@e0hJŸ _j *8L®9jH0 _j i0«}9e ^jTn}h9§_nj p}nH0 _jX ~´ §0h* J ©·©ou

E¨¢F gu |0_LJ ³u ³9h9L9J 9jL ~u W«J õG9j °_c_q0L_9 \0eq n°_j0 }0_jTn}H0h0j§ e09}j_jXyŸ ·o ©·©©u

E¨PF 8u ?9j_jnJ 8u pu ?9L_9J `u Gu ¯9ec0}J ¦u ~H\ne§0 J `u g_§}n®_HJ 9jL Gu ?e«jL0eeJ õGn@0}eI Gnj§}9 §_®0 ?N|¦ Tn} }0_jTn}H0h0j§ e09}j_jXJŸ *Gn*||J ®neu 9@ ¤©o·¢u·UV¨oJ ©·©ou Emje_jOFu 8®9_e9@e0I \§§q I¤¤9}²_®un}X¤9@ ¤©o·¢u·UV¨o

E¨kF wu µ\0jXJ 8u µ\9jXJ 9jL 8u W}n®0}J õmje_j0 L0H_ _nj §}9j Tn}h0}JŸ _j *p}nH00L_jX nT §\0 ¨k§\ ^j§0}j9§_nj9e GnjT0}0jH0 nj g9H_j0 d09}j_jXJ* 0}u p}nH00L_jX nT g9H_j0 d09}j_jX |0 09}H\J bu G\9«L\«}_J ~u `0X0ec9J du ~njXJ Gu ~¶0q0 ®9}_J Wu i_«J 9jL ~u ~9@9§nJ NL wJ ®neu o£©u p g d|J o¢Ś©¨ `«e ©·©©J qqu ©¢·V©Ś©¢·Uku Emje_jOFu 8®9_e9@e0I \§§q I¤¤q}nH00L_jX uhe}uq}0 ¤®o£©¤¶\0jX©©Hu\§he

EV·F Nu ¦nLn}n®J ¦u N}0¶J 9jL ³u ¦9 9J õg«anHnI 8 q\´_H 0jX_j0 Tn} hnL0e]@9 0L Hnj§}neJŸ _j ©·*o*© ^*NNN*¤|~` ^*j§0}j9§_nj9e GnjT0}0jH0 nj ^j§0ee_X0j§ |n@n§ 9jL ~´ §0h* J ©·o©J qqu U·©£ŚU·¨¨u

EVoF `u S«J 8u b«h9}J mu i9H\«hJ Wu ¦«Hc0}J 9jL ~u d0®_j0J õKV}eI K9§9 0§ Tn} L00q L9§9]L}_®0j }0_jTn}H0h0j§ e09}j_jXJŸ ©·©·u

EV©F |u 8X9}°9eJ Ku ~H\««}h9j J 9jL gu in}n«¶_J õ8j nq§_h_§_H q0} q0H§_®0 nj n°_j0 }0_jTn}H0h0j§ e09}j_jXJŸ ©·oku Emje_jOFu 8®9_e9@e0I \§§q I¤¤9}²_®un}X¤9@ ¤ok·¢u·VUV¨

EV¨F ^u `u WnnLT0een°J `u pn«X0§]8@9L_0J gu g_}¶9J ?u ±«J Ku ¯9}L0]S9}e0´J ~u m¶9_}J 8u Gn«}®_ee0J 9jL ³u ?0jX_nJ õW0j0}9§_®0 9L®0} 9}_9e j0§°n}c JŸ ©·oVu Emje_jOFu 8®9_e9@e0I \§§q I¤¤9}²_®un}X¤9@ ¤oV·£u©££o

EVVF ~H\}_§§°_0 0}J 8u `uJ [u ^uJ 9jL ¦u 0§ 9euJ õg9 §0}_jX 9§9}_J XnJ H\0 9jL \nX_ @´ qe9jj_jX °_§\ 9 e09}j0L hnL0eJŸ *i 9§«}0J* ®neu UPPJ qqu £·VŚ£·kJ o· ©·©·u

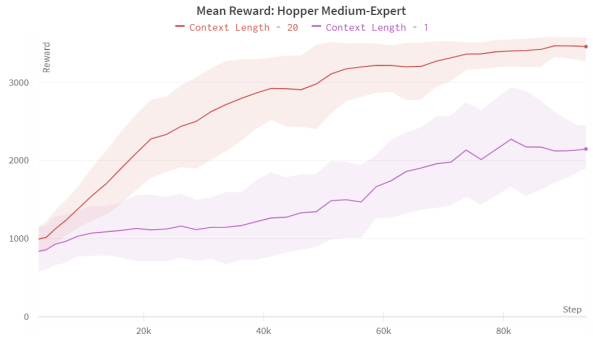
EVUF ¯u ³0J ~u d_«J ¦u b«}«§9H\J pu 8@@00eJ 9jL ³u W9nJ õg9 §0}_jX 9§9}_ X9h0 °_§\ e_h_§0L L9§9JŸ *Gn* | |J ®neu 9@ ¤©ooou··©o·J ©·©ou Emje_j0Fu 8®9_e9@e0I \§§q I¤¤9}²_®un}X¤9@ ¤©ooou··©o·

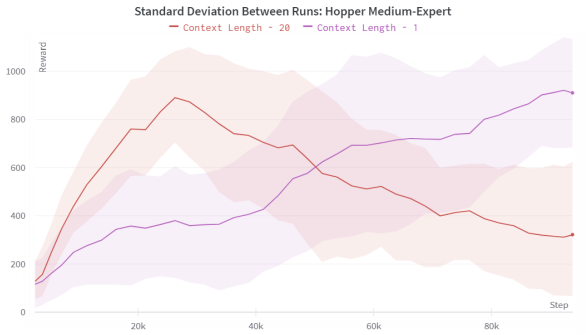
A

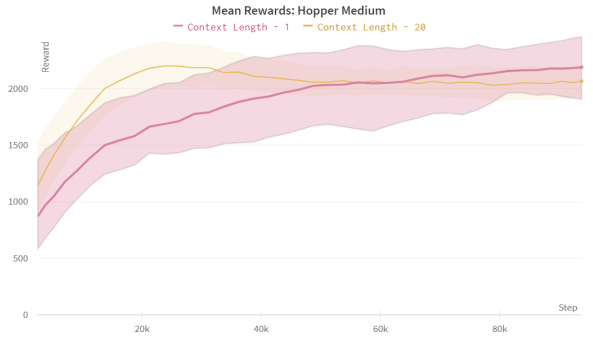
Decision Transformer validation results

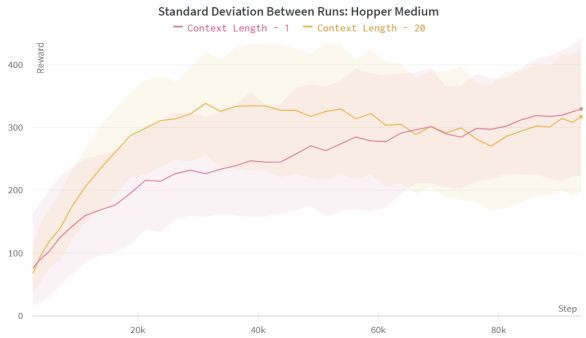
U¢

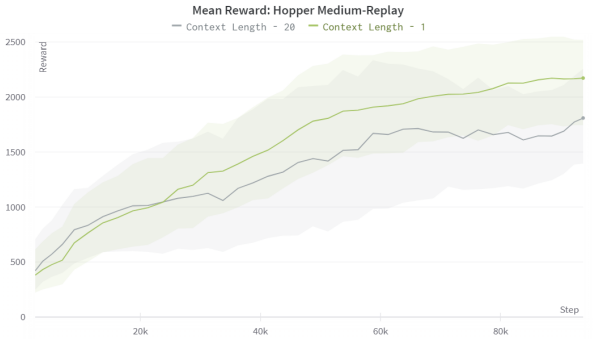
A.1 Hopper Environment Results

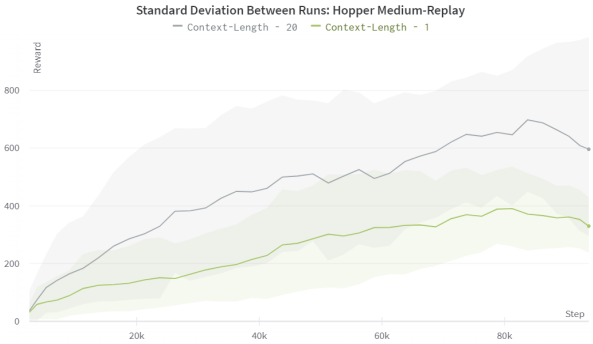






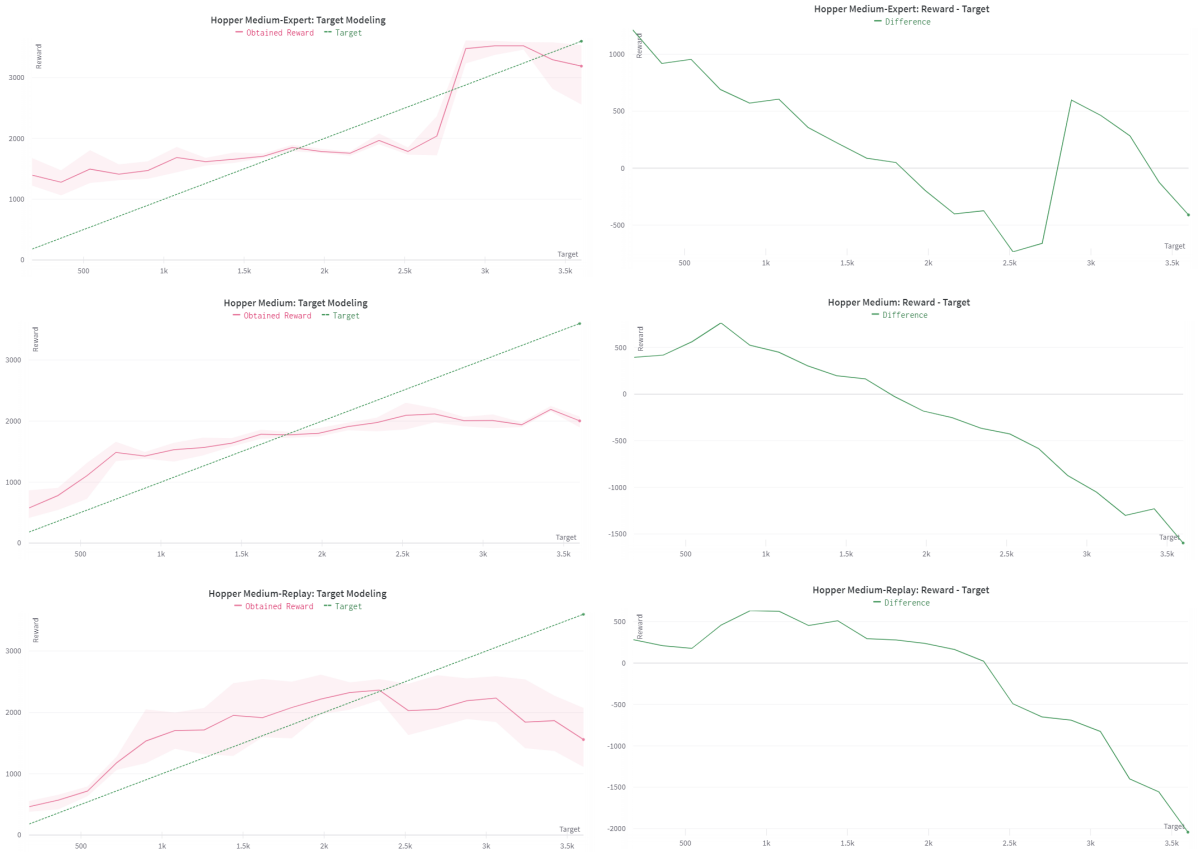






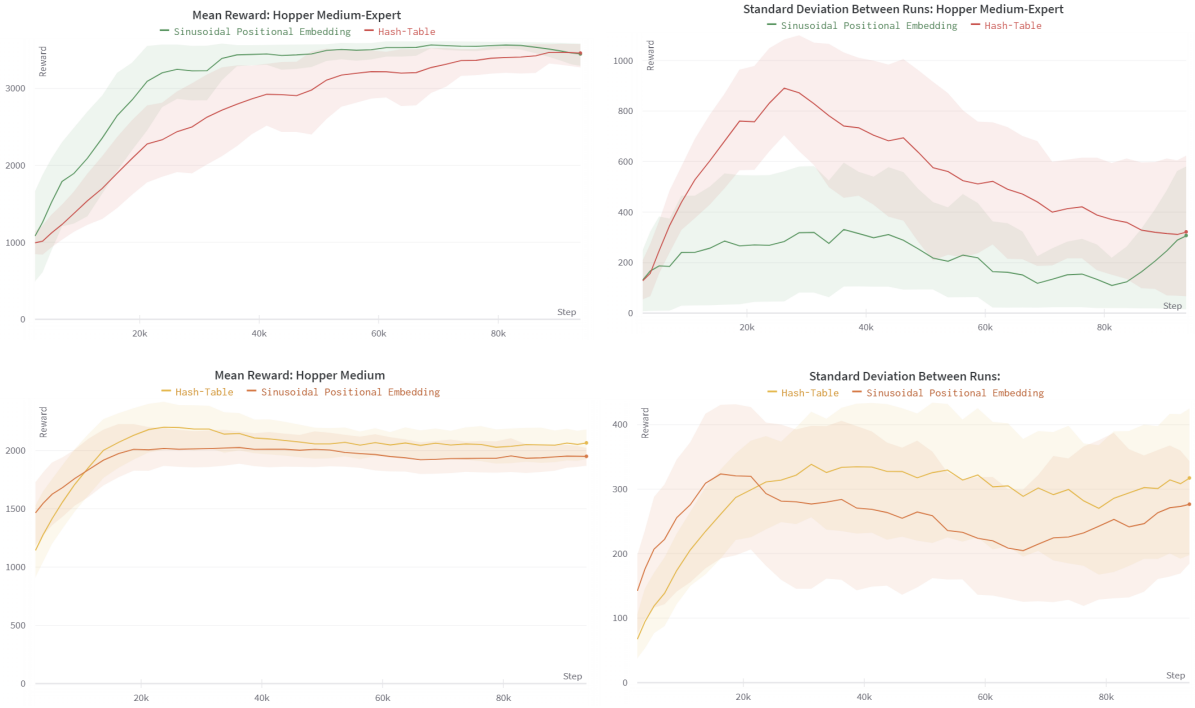
S_X«}O 8uoI Gnj§0²§ d0jX§\I g09j |0°9}L 9jL ~§9jL9}L K0®_9§_nj 9H}n [nqq0} K9§9 0§

UP



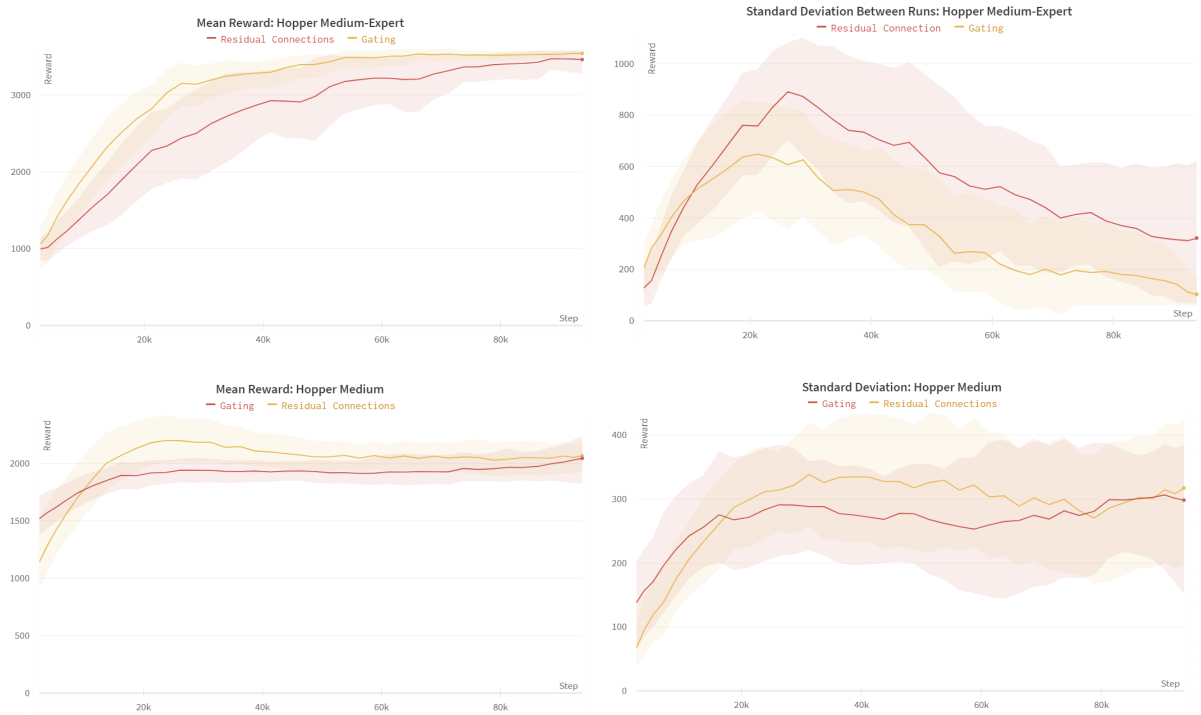
S_X«}O 8u©I ¦9}XO§ gnLOe_jX 9H}n [nqqO} K9§9 O§

Uk



S_X«}O 8ᵤ¨I pn _§_nj9ₑ Nh@0LL_jXI g09j |0°9}L 9jL ~§9jL9}L KO®_9§_nj 9ℍ}n [nqq0} K9§9 0§

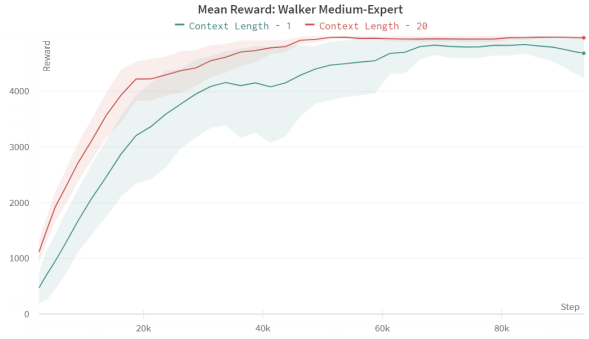
£·

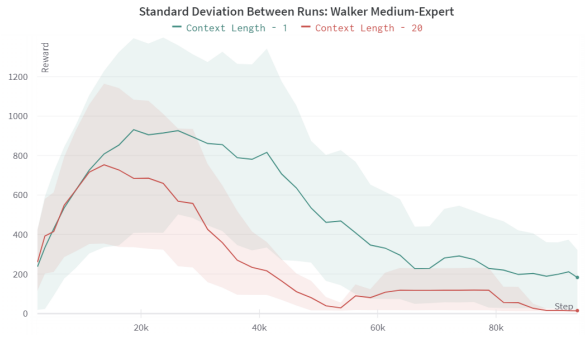


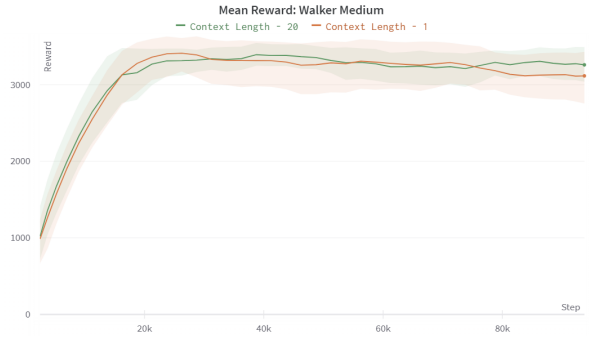
S_X«}O 8uVI W9§_jXI g09j |0°9}L 9jL ~§9jL9}L K0®_9§_nj 9H}n [nqq0} K9§9 0§

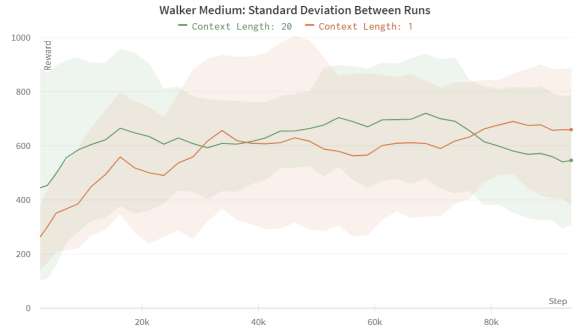
£o

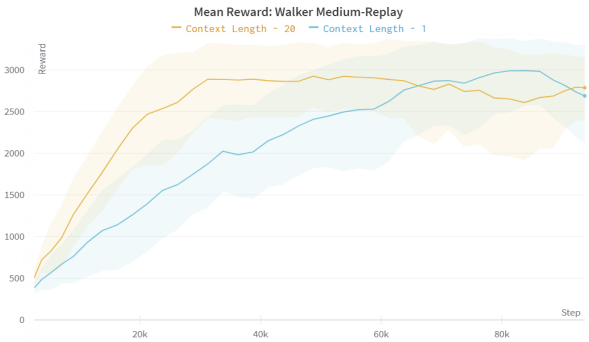
A.2 Walker2d Environment Results

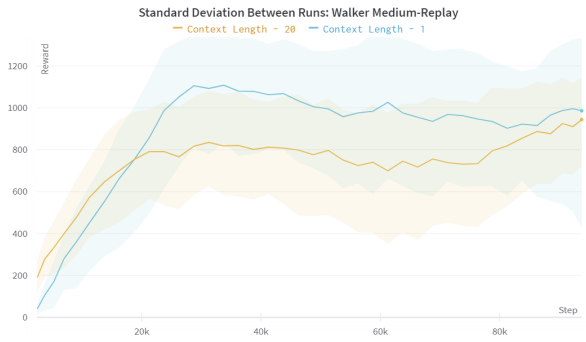








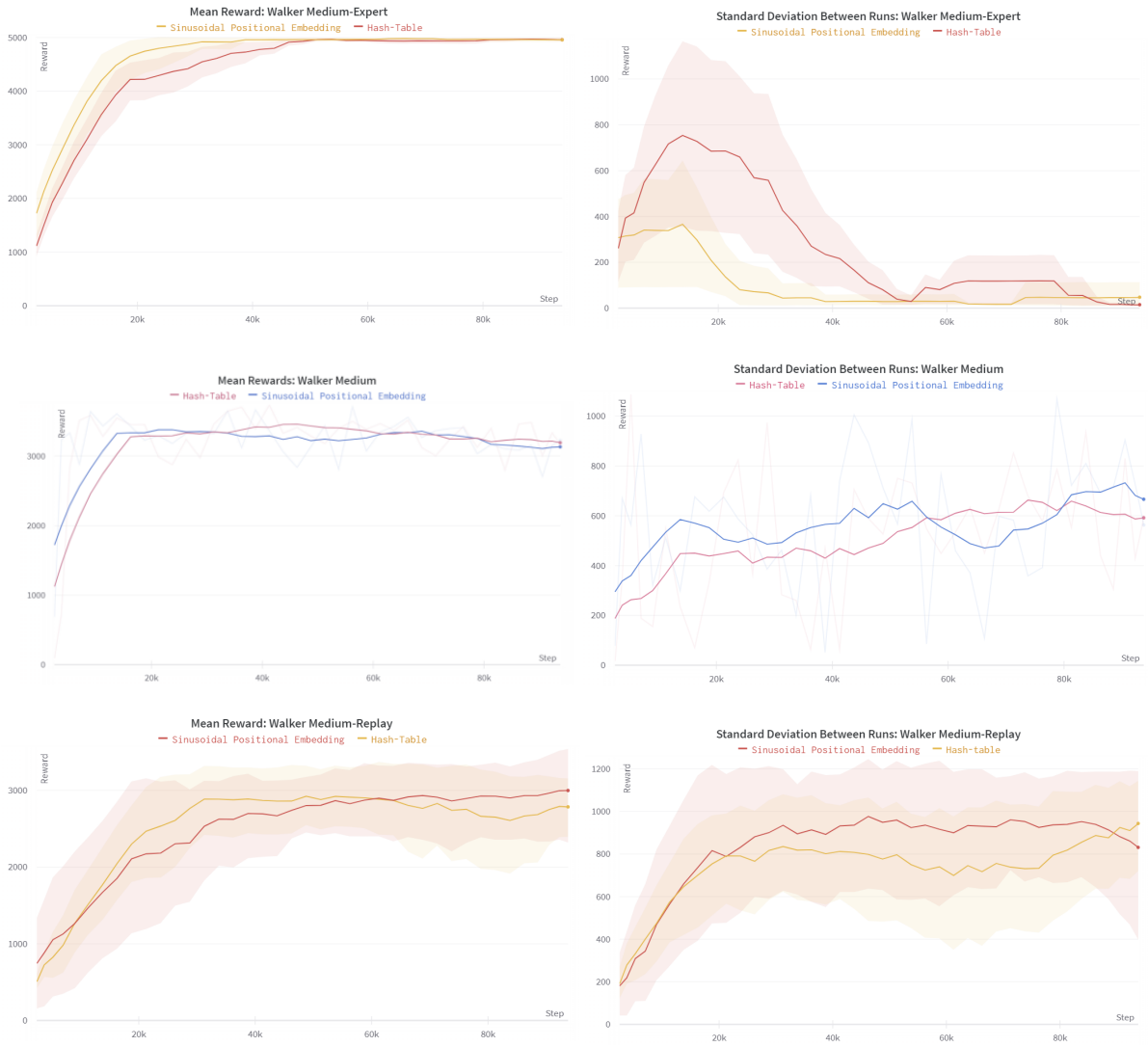




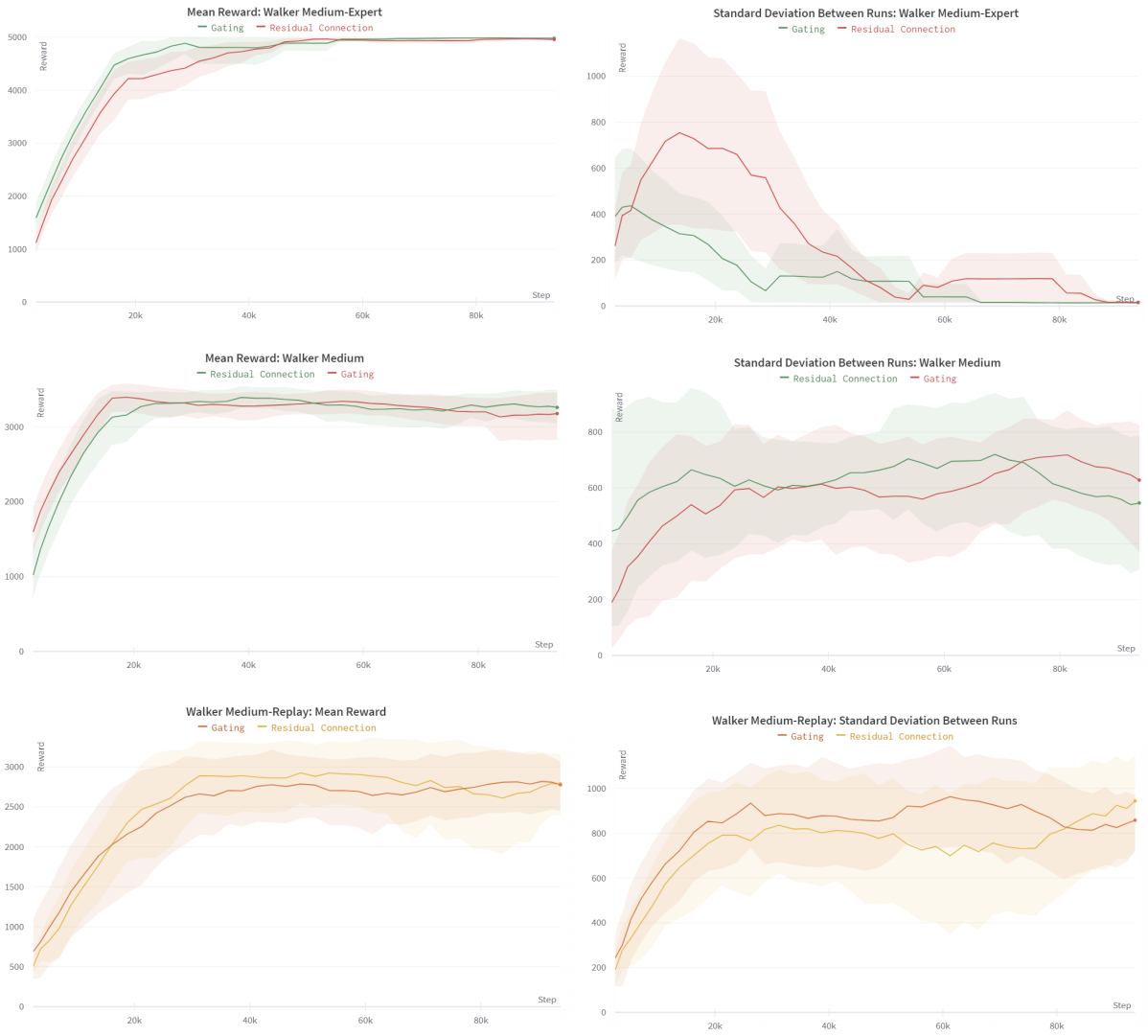
S_X«}O 8uUI Gnj§0²§ d0jX§\I g09j |0°9}L 9jL ~§9jL9}L KO®_9§_nj 9H}n ¯9ec0}©L K9§9 0§

S_X«}O 8u£I ¦9}XO§ gnLOe_jX 9H}n ¯9ecO} K9§9 0§

£¨



S_X«}O 8u¢I pn _§_nj9e Nh@0LL_jXI g09j |0°9}L 9jL ~§9jL9}L K0®_9§_nj 9H}n ¯9ec0}©L K9§9 0§



S_X«}O 8uPI W9§_jXI g09j |0°9}L 9jL ~§9jL9}L K0®_9§_nj 9H}n ¯9ec0}©L K9§9 0§

£U