Experiment 1:

Parameters of paper scaled down by 100

Minibatch size: 32

replay memory size: 10000

agent history length: 1

target network update frequency: 100

discount factor: 0.99

action repeat: 1

update frequency: 1

learning rate: 0.00025

gradient momentum: 0.95

squared gradient momentum: 0.95

min squared gradient: 0.01 (constant added to opt)

initial epislon: 1

final epsilon: 0.05 (scaled down exponentially)

epsilon decrease frame: 10000

replay start size: replay memory size(10000)

no-op max: 0

reward clipping: not used

training frames: 500000 (original 50000000)

(divide: train\_epi: 500, epilen: 1000

evaluation:

eval\_epi: 30

eval\_epilen: 500000 frames (original 50000000)

Added:

Num\_pretrainepi: 1 --- as normally state values should start from 0, this makes sure it holds.

(result: not good, maybe because each training frame too few)

Experiment 2:

Parameters of paper scaled down by 1000

Minibatch size: 32

replay memory size: 10000

agent history length: 1

target network update frequency: 100

discount factor: 0.99

action repeat: 1

update frequency: 1

learning rate: 0.00025

gradient momentum: 0.95

squared gradient momentum: 0.95

min squared gradient: 0.01 (constant added to opt)

initial epislon: 1

final epsilon: 0.05 (scaled down exponentially)

epsilon decrease frame: 20000

replay start size: replay memory size(20000)

no-op max: 0

reward clipping: not used

training frames: 50000 (original 50000000)

(divide: train\_epi: 1, epilen: 50000

evaluation:

eval\_epi: 30

eval\_epilen: 50000 frames (original 50000000)

Added:

Num\_pretrainepi: 1 --- as normally state values should start from 0, this makes sure it holds.

Experiment 3: (ORIGINAL)

Parameters of paper scaled down by 1000

Minibatch size: 32

replay memory size: 1000000

agent history length: 1

target network update frequency: 10000

discount factor: 0.99

action repeat: 1

update frequency: 1

learning rate: 0.00025

gradient momentum: 0.95

squared gradient momentum: 0.95

min squared gradient: 0.01 (constant added to opt)

initial epislon: 1

final epsilon: 0.05 (scaled down exponentially)

epsilon decrease frame: 1000000

replay start size: replay memory size(50000)

no-op max: 0

reward clipping: not used

training frames: 50000000 (original 50000000)

evaluation:

eval\_epi: 30

eval\_epilen: 50000 frames (original 50000000)

Added:

Num\_pretrainepi: 1 --- as normally state values should start from 0, this makes sure it holds.

Experiment 3:

Parameters of paper scaled down by 1000

Minibatch size: 32

replay memory size: 10000

agent history length: 1

target network update frequency: 10000

discount factor: 0.99

action repeat: 1

update frequency: 1

learning rate: 0.00025

gradient momentum: 0.95

squared gradient momentum: 0.95

min squared gradient: 0.01 (constant added to opt)

initial epislon: 1

final epsilon: 0.05 (scaled down exponentially)

epsilon decrease frame: 1000000

replay start size: replay memory size(50000)

no-op max: 0

reward clipping: not used

training frames: 50000000 (original 50000000)

evaluation:

eval\_epi: 30

eval\_epilen: 50000 frames (original 50000000)

Added:

Num\_pretrainepi: 1 --- as normally state values should start from 0, this makes sure it holds.