Analysis of SARSA on Windy Gridworld

Sushant Shambharkar - 153050081

November 2015

1   Introduction

This report involve analysis of variants of SARSA applied on Windy Girdworld.

Following figure shows a standard gridworld, with start and goal states, but with one difference: there is a crosswind upward through the middle of the grid. The actions are the standard four–up, down, right, and left–but in the middle region the resultant next states are shifted upward by a "wind," the strength of which varies from column to column. The strength of the wind is given below each column, in number of cells shifted upward. For example, if you are one cell to the right of the goal, then the action left takes you to the cell just above the goal. Let us treat this as an undiscounted episodic task, with constant rewards of until the goal state is reached.[1]
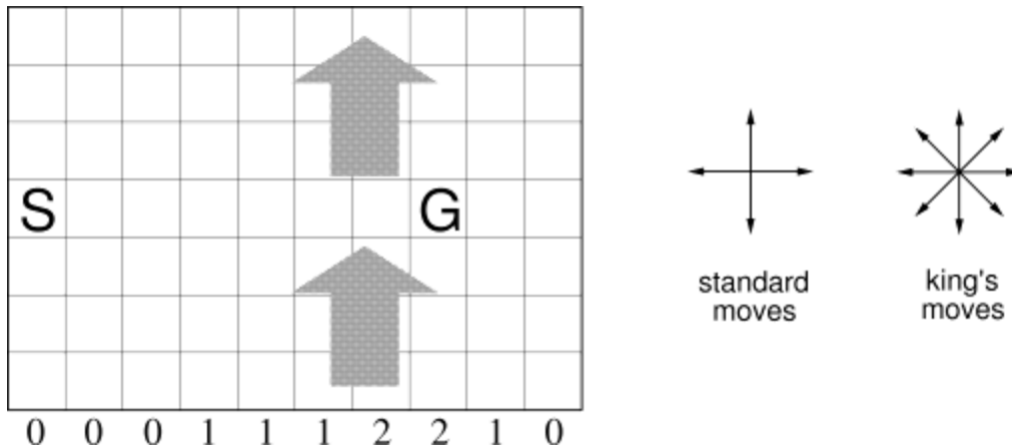

Figure 1: Gridworld in which movement is altered by a location-dependent, upward "wind."[1]

2   Observations

UCB and ϵ-soft performed equally where major efficiency found in SARSA(λ) with $\lambda = 0.1$. Combination of SARSA(λ) and UCB performed well in "Stochastic Wind with Kings Moves" almost coverged at 500th Episode.
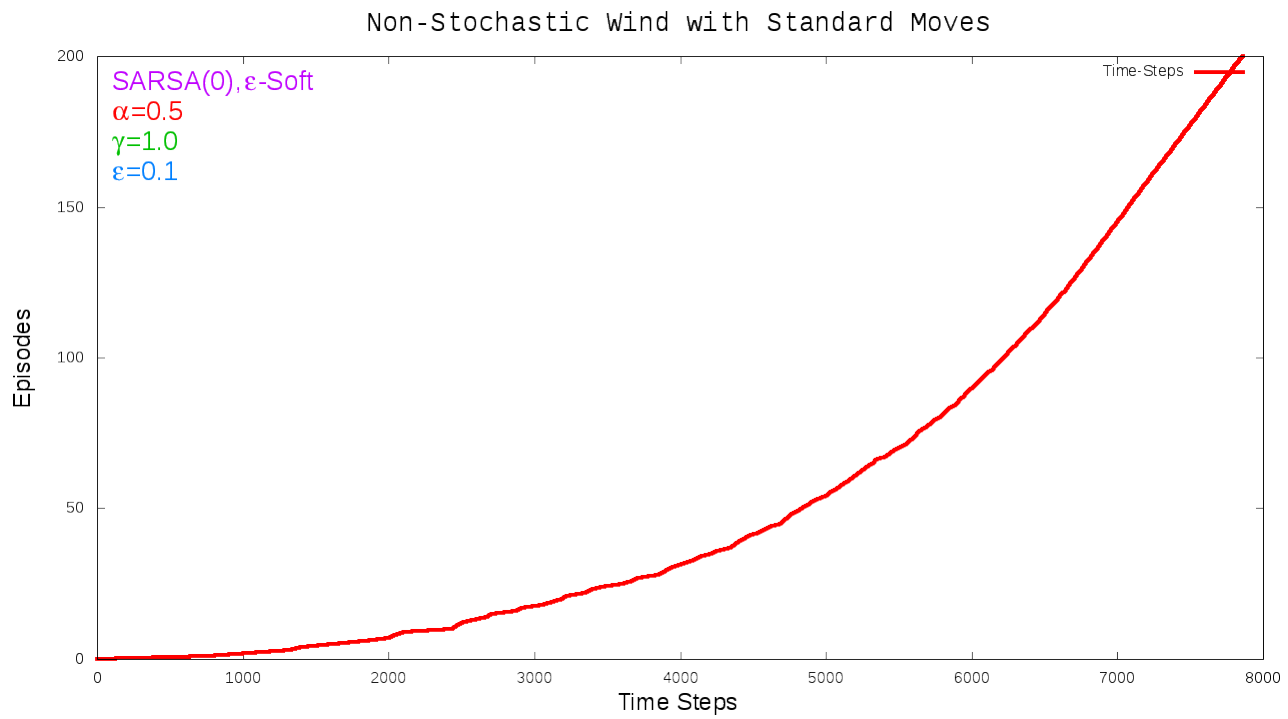
2.1 Data "D1.txt"

Non-Stochastic Wind with Standard Moves

SARSA(0),ε-Soft
α=0.5
γ=1.0
ε=0.1

Time-Steps

Figure 2: Episodes Vs Time-Step : Standard move

Non-Stochastic Wind with Standard Moves

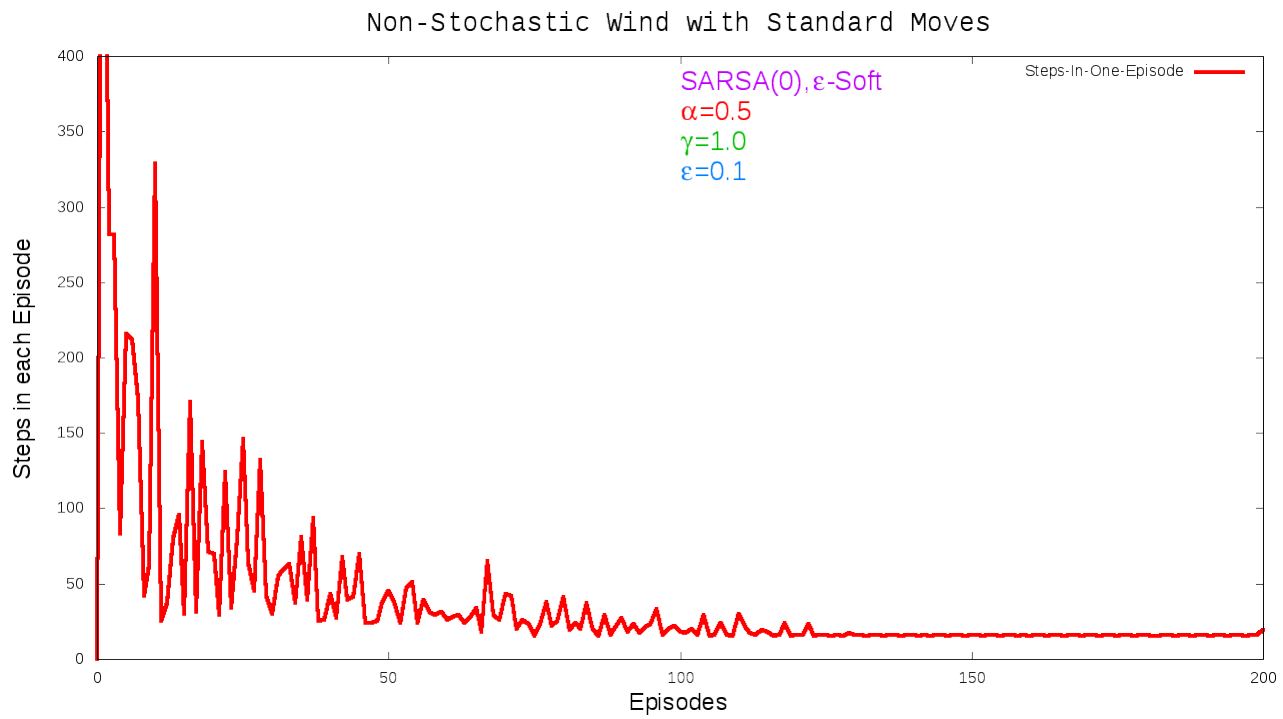SARSA(0),ε-Soft
α=0.5
γ=1.0
ε=0.1

Steps-In-One-Episode

Figure 3: Episodes Vs Steps-in-each-episode : Standard move

State Values
    -13.91 -13.45 -12.84 -12.0 -11.0 -10.0 -09.0 -08.0 -07.0 -06.0
    -13.99 -13.51 -12.87 -11.99 -11.0 -10.0 -09.0 -08.0 -07.0 -05.0
    -14.0 -14.0 -13.0 -12.0 -11.0 -9.99 -8.99 -3.62 -04.8 -04.0
    -15.0 -14.0 -13.0 -12.0 -10.95 -9.94 -08.9 00000 -3.13 -3.04
    -14.2 -13.49 -12.62 -11.87 -10.78 -9.73 00000 -00.5 -01.0 -02.0
    -13.46 -12.96 -12.25 -11.56 -10.37 00000 00000 00000 -0.94 -01.5
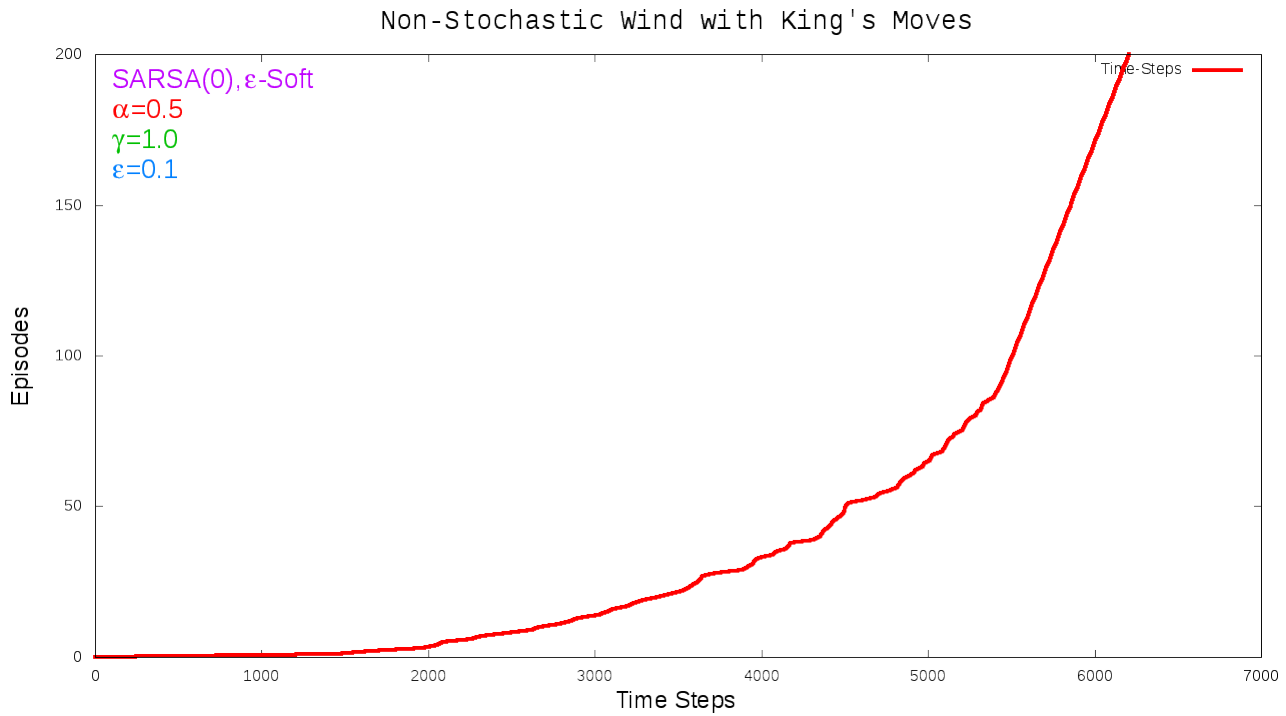   -12.95 -12.43 -11.7 -10.96 00000 00000 00000 00000 00000 -0.88

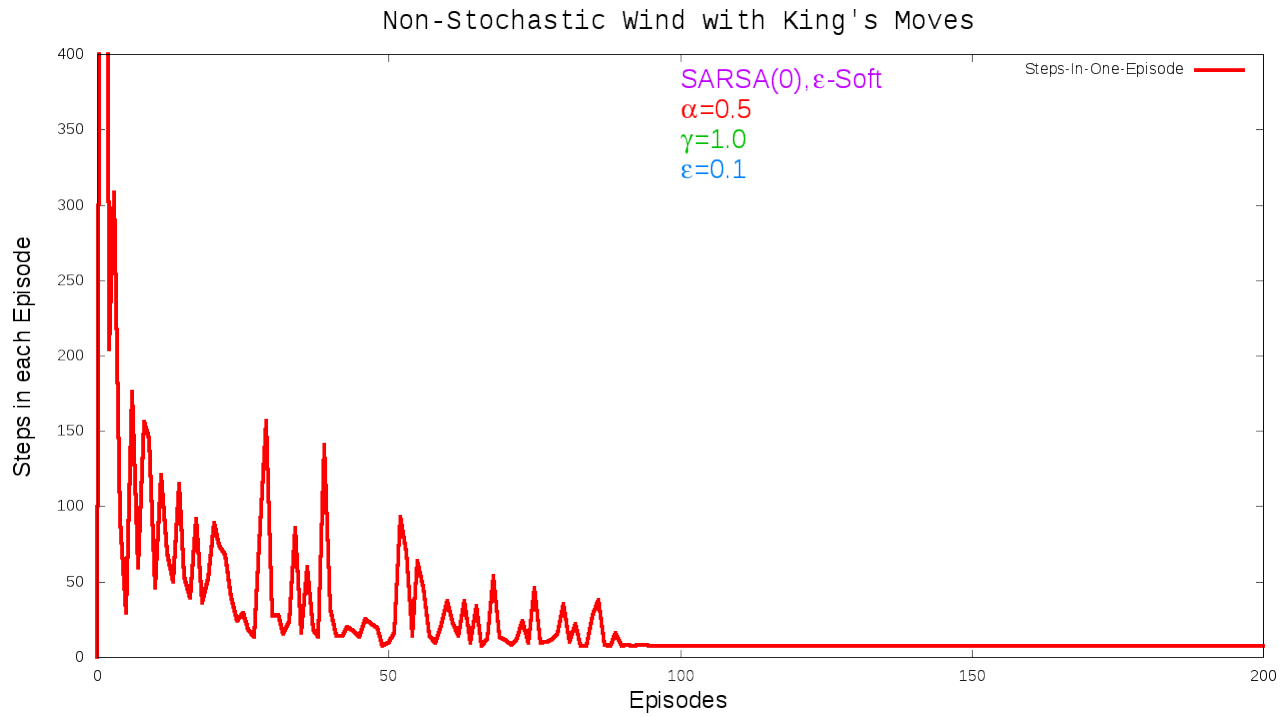## 2.2   Data "D2.txt"

Figure 4: Episodes Vs Time-Step : King's move

Figure 5: Episodes Vs Steps-in-each-episode : King's move

State Values
-5.69 -5.76 -6.34 -6.97 -7.36 -6.91 -5.99 -05.0 -4.72 -3.97
-5.66 -5.79 -5.78 -6.05 -06.0 -6.17 -5.65 -05.0 -04.0 -03.0
-5.71 -06.0 -5.46 -5.46 -5.34 -4.85 -4.92 -3.49 -2.51 -02.0
-06.0 -5.36 -05.0 -5.11 -4.52 -3.85 -3.63 00000 -01.0 -1.22
-5.34 -5.24 -4.52 -04.0 -03.0 -02.0 -01.0 00000 -00.5 -00.5
-4.97 -4.81 -4.38 -3.78 -2.85 -1.91 -0.94 00000 00000 00000
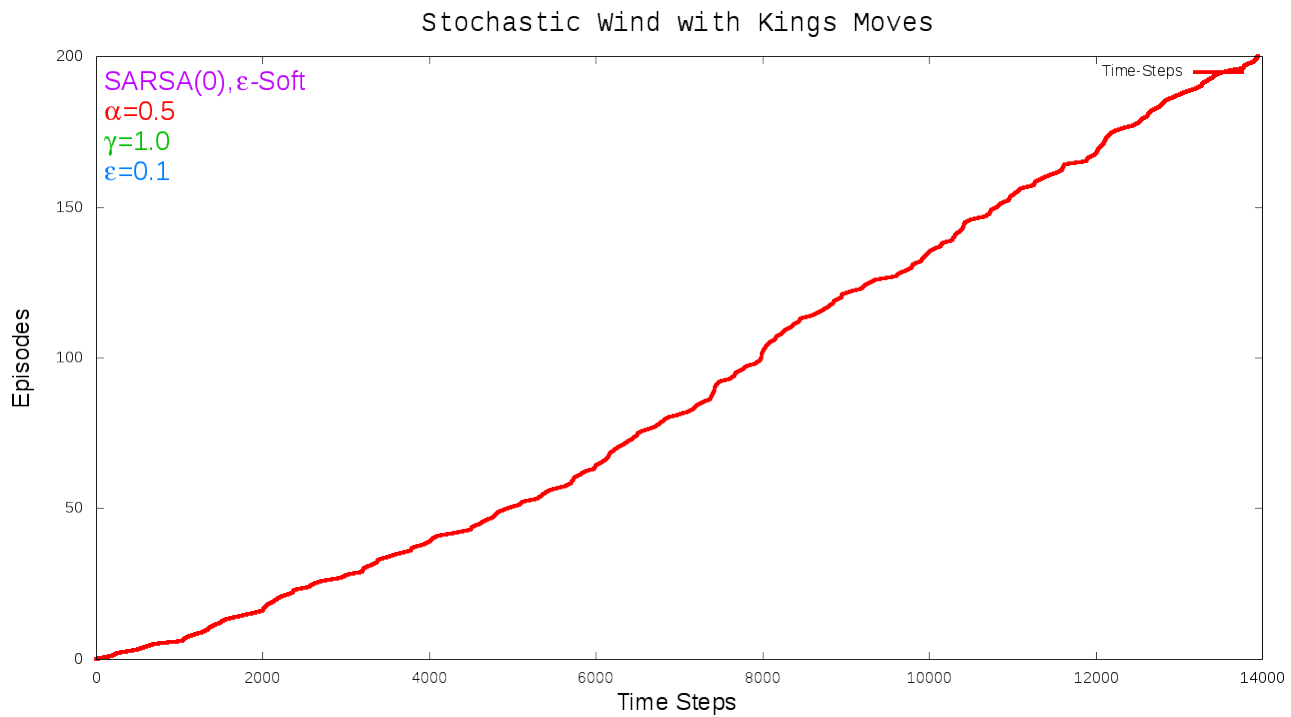-4.88 -4.67 -4.24 -3.48 00000 00000 00000 00000 00000 00000

2.3 Data "D3.txt"

Stochastic Wind with Kings Moves

SARSA(0),ε-Soft
α=0.5
γ=1.0
ε=0.1

Time-Steps

Episodes

Time Steps

Figure 6: Episodes Vs Time-Step : Stochastic King's move

Stochastic Wind with Kings Moves

SARSA(0),ε-Soft
α=0.5
γ=1.0
ε=0.1

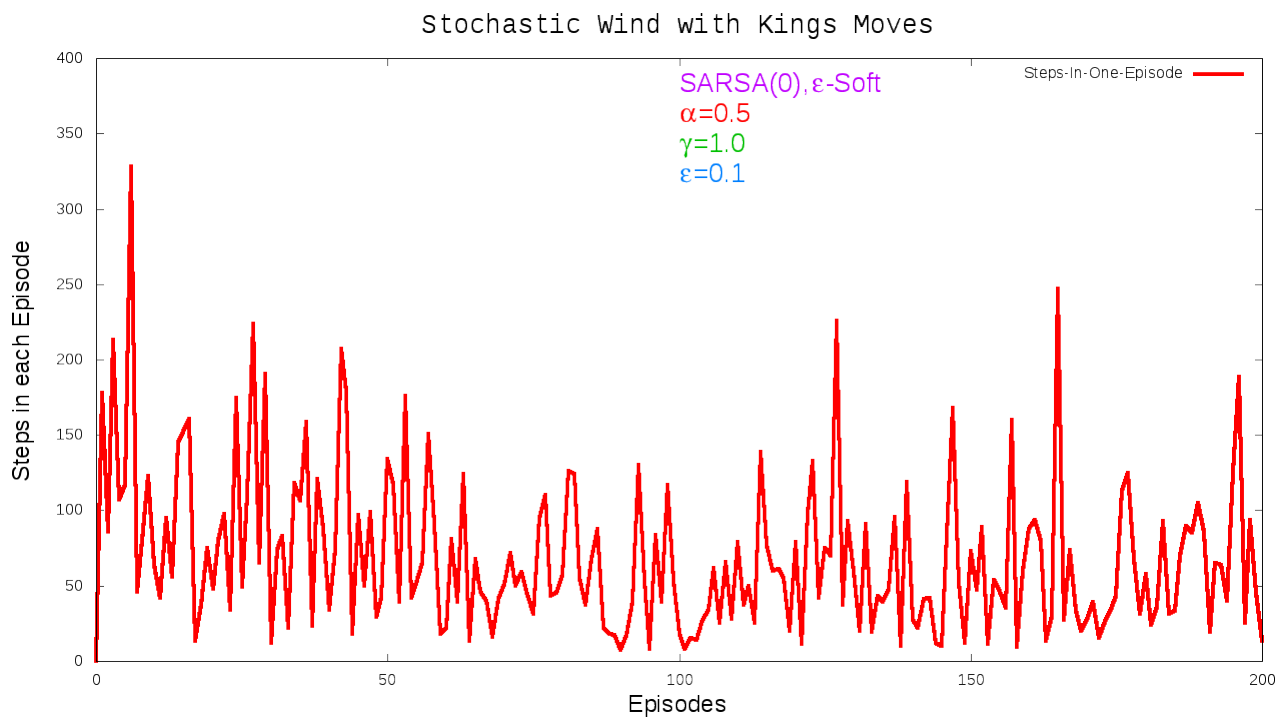Steps-In-One-Episode

Steps in each Episode

Episodes

Figure 7: Episodes Vs Steps-in-each-episode : Stochastic King's move

State Values

    -11.53 -11.74 -11.87 -11.93 -12.46 -11.37 -11.82 -10.44 -11.45 -10.23
    -11.6 -11.58 -11.57 -11.91 -11.81 -12.17 -6.49 -11.17 -10.65 -9.18
    -11.44 -11.9 -11.3 -11.5 -11.12 -11.36 -11.12 -11.25 -9.48 -8.33
    -11.64 -11.22 -11.14 -11.15 -10.69 -10.4 -10.05 00000 -9.14 -8.38
    -11.01 -10.9 -10.54 -9.78 -9.67 -7.03 -7.27 -9.84 -2.82 -7.88
    -10.77 -10.49 -10.25 -9.59 -6.36 -6.57 -06.4 -6.43 -7.53 -7.15
    -10.6 -10.42 -9.97 -8.86 -6.79 -5.07 -3.28 -3.77 -6.85 -6.81
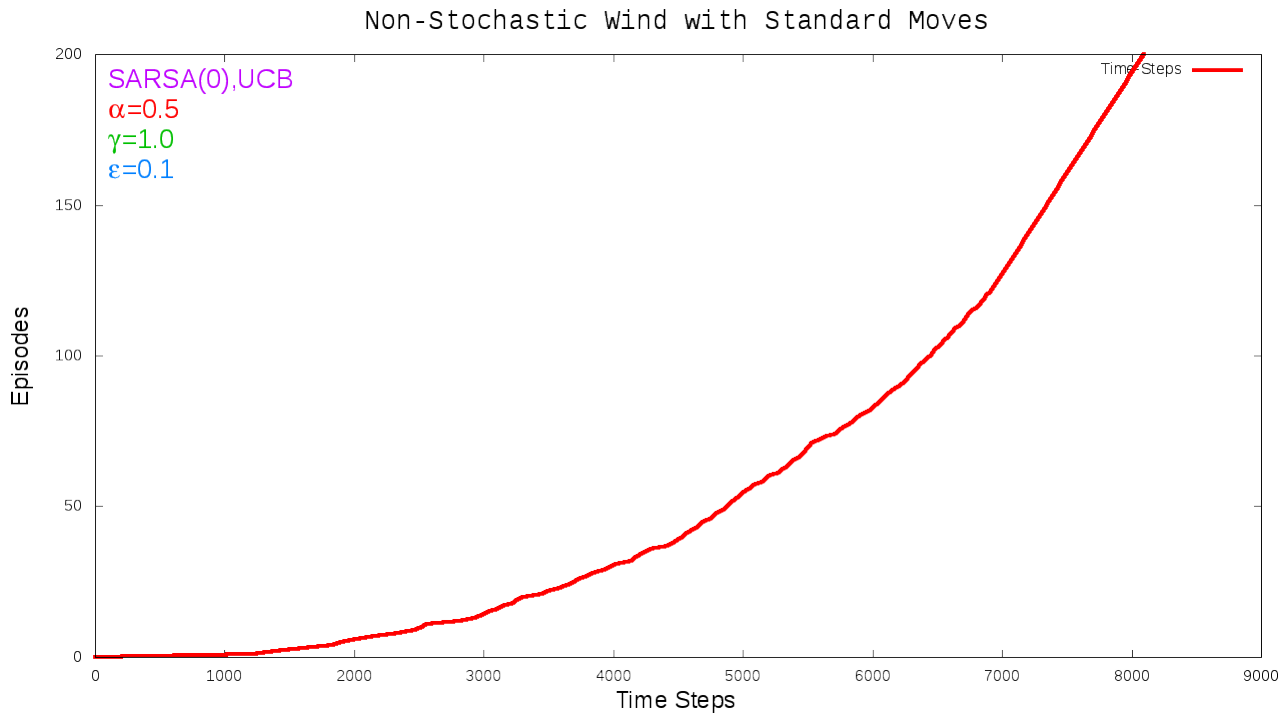
2.4   Data "D4.txt"

Figure 8: Episodes Vs Time-Step : Standard move

Non-Stochastic Wind with Standard Moves
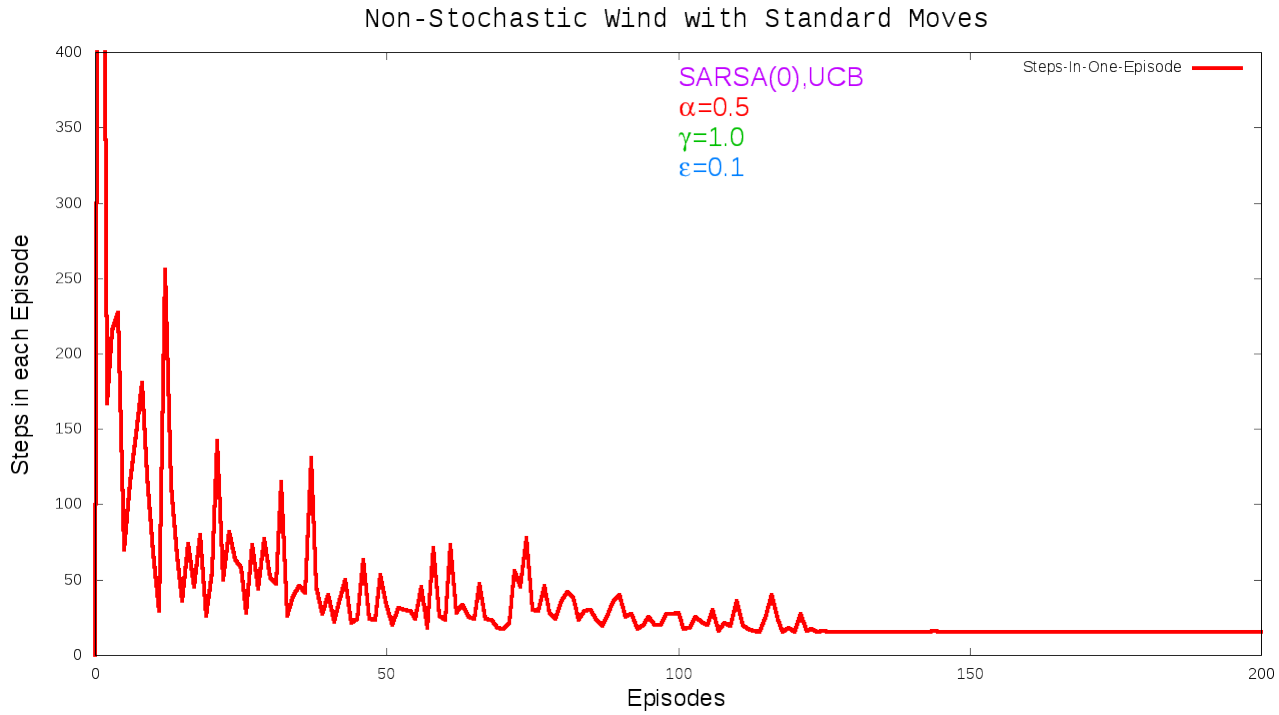
SARSA(0),UCB
α=0.5
γ=1.0
ε=0.1

Steps-In-One-Episode

Figure 9: Episodes Vs Steps-in-each-episode : Standard move

State Values
-13.9 -13.85 -13.24 -12.52 -11.39 -10.15 -09.0 -08.0 -07.0 -06.0
-14.1 -13.69 -13.3 -12.31 -11.06 -10.0 -9.35 -8.24 -7.22 -05.0
-14.34 -14.3 -13.18 -12.07 -11.0 -10.53 -9.26 -6.37 -5.75 -04.0
-15.0 -14.0 -13.0 -12.0 -11.6 -9.98 -9.05 00000 -4.92 -03.0
-14.39 -13.67 -12.78 -12.24 -10.88 -9.56 00000 -0.97 -01.0 -02.0
-13.83 -13.2 -12.25 -11.78 -10.37 00000 00000 00000 -1.81 -2.34
-13.32 -12.57 -11.72 -10.93 00000 00000 00000 00000 -0.75 -1.71
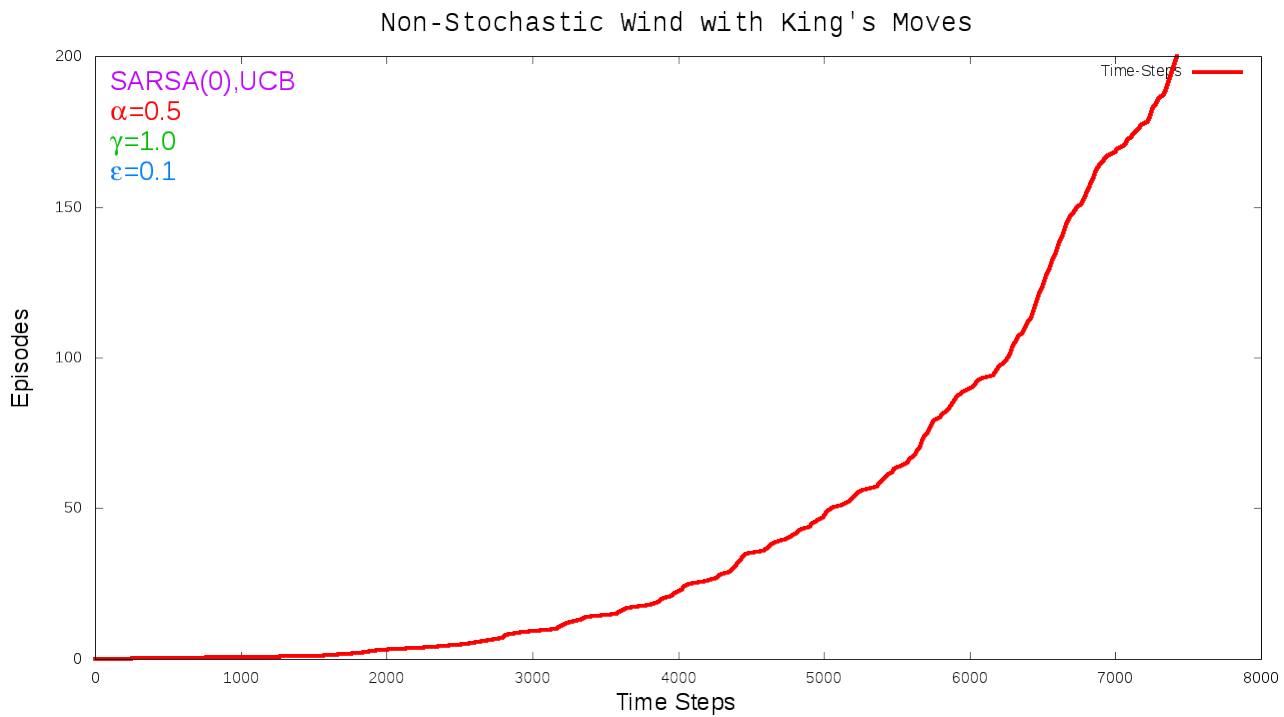
2.5 Data "D5.txt"

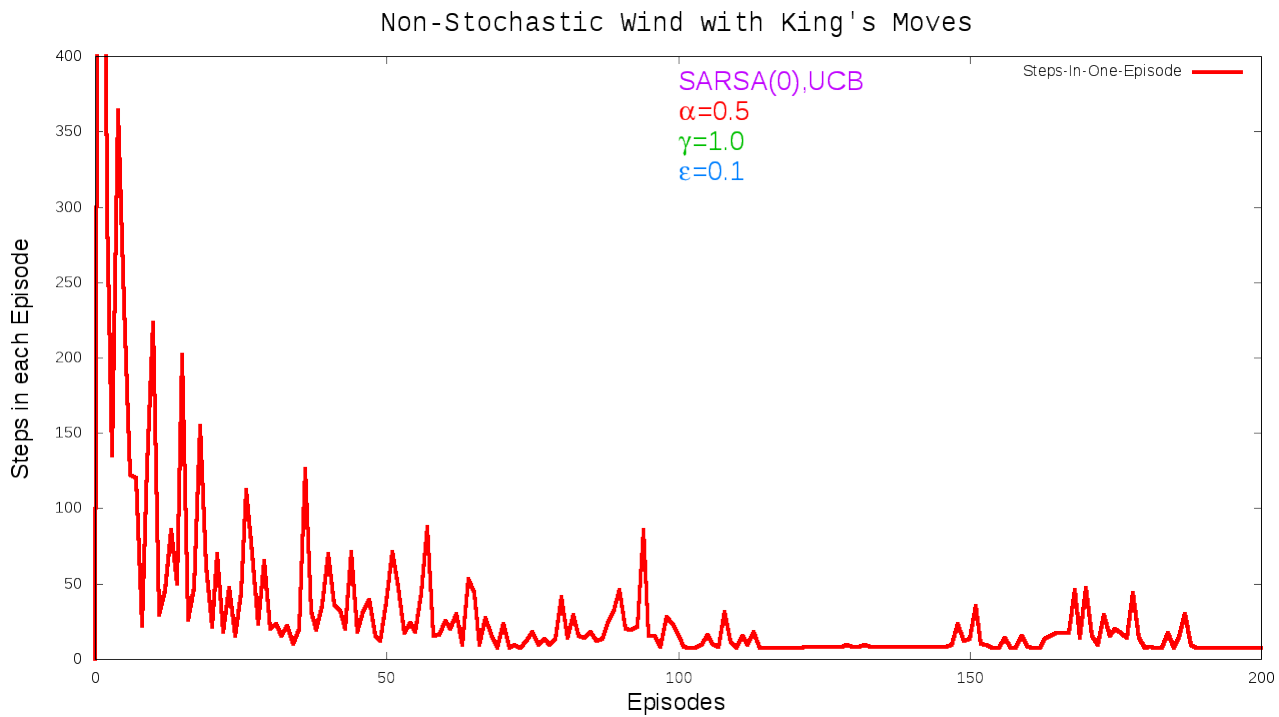Figure 10: Episodes Vs Time-Step : King's move

Figure 11: Episodes Vs Steps-in-each-episode : King's move

State Values

```
-7.54 -7.58 -7.73 -8.18 -07.9 -07.0 -06.0 -05.0 -5.25 -04.1
-7.31 -7.23 -07.5 -7.72 -7.28 -6.93 -06.0 -5.08 -04.0 -03.0
-7.43 -7.69 -7.03 -07.3 -6.85 -6.33 -5.83 -04.6 -3.35 -02.0
-7.07 -6.93 -6.52 -6.66 -06.0 -5.23 -5.01 00000 -01.0 -1.65
-06.7 -6.03 -5.19 -5.64 -4.63 -02.0 -01.0 -00.5 -0.75 -0.88
-6.36 -6.11 -5.01 -04.0 -03.0 -1.91 -0.97 00000 00000 00000
-6.38 -5.91 -5.27 -4.64 00000 00000 00000 00000 00000 00000
```
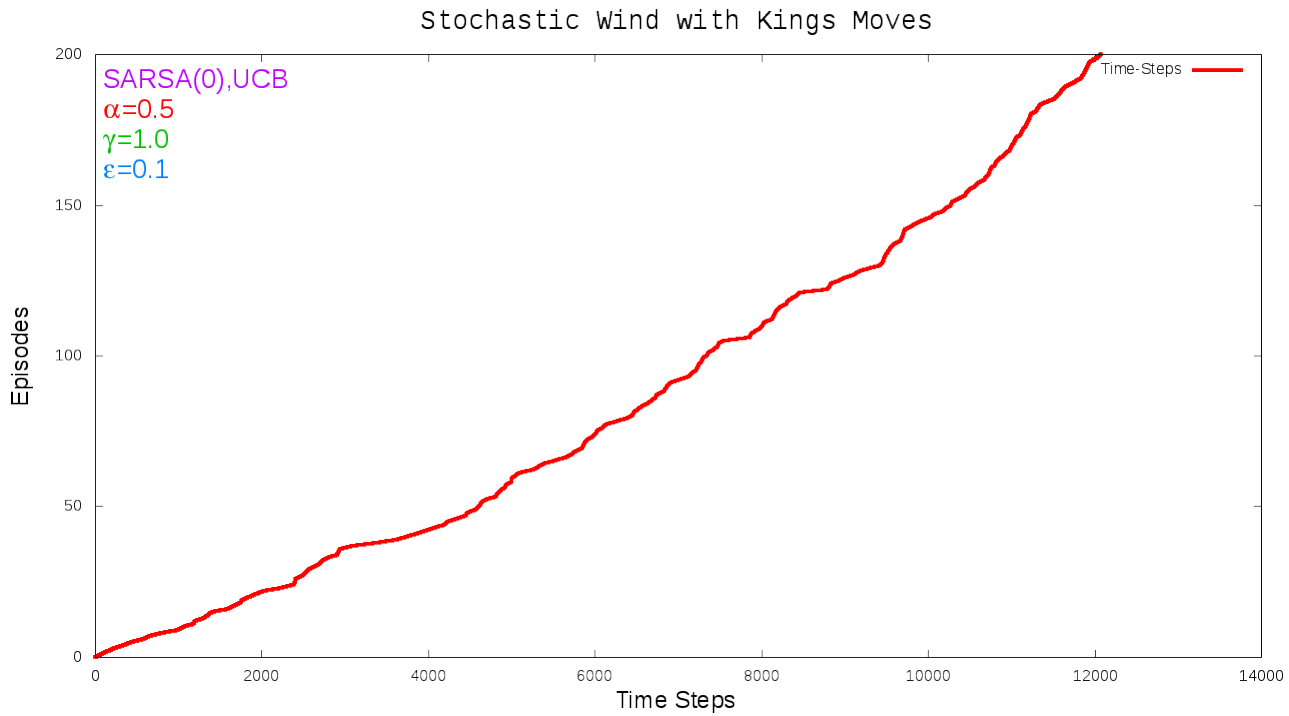
2.6 Data "D6.txt"

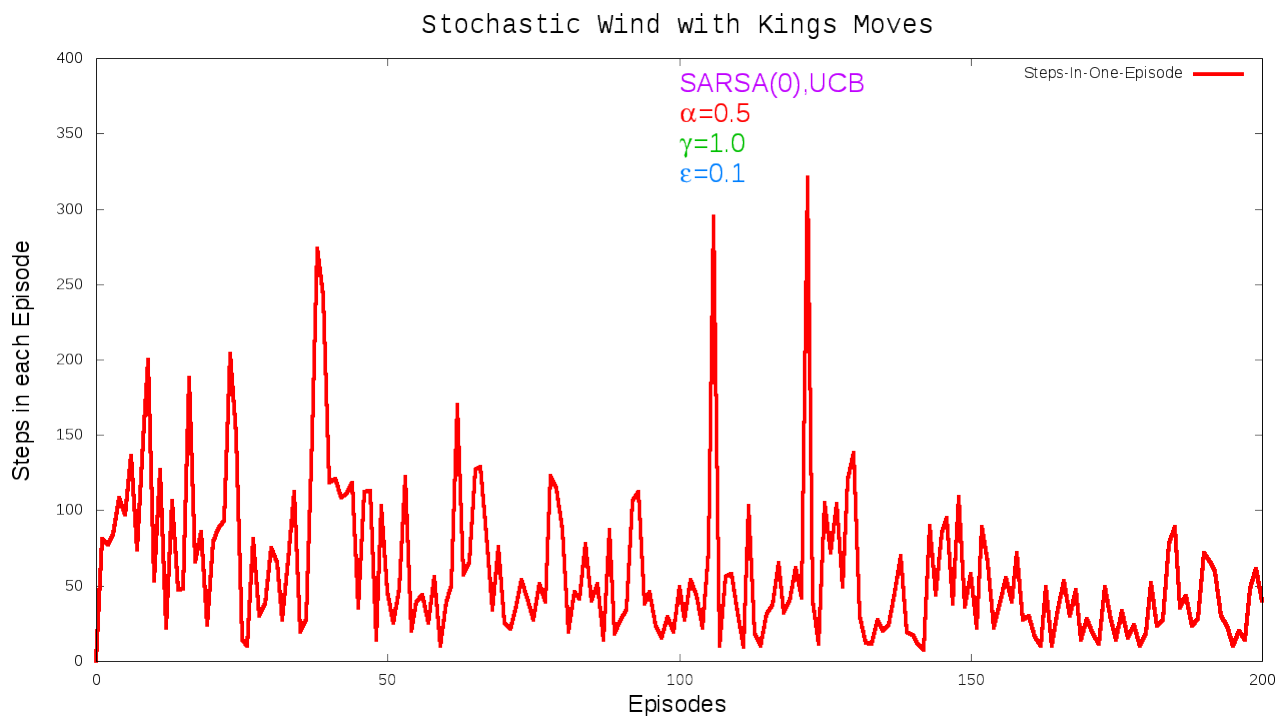Figure 12: Episodes Vs Time-Step : Stochastic King's move)

Figure 13: Episodes Vs Steps-in-each-episode : Stochastic King's move)

State Values

-11.05 -11.13 -11.28 -11.35 -11.24 -9.29 -9.32 -9.19 -9.01 -7.35
-10.84 -10.73 -10.74 -10.84 -10.34 -10.7 -8.81 -8.86 -8.44 -6.84
-10.81 -11.29 -10.47 -10.57 -9.83 -9.68 -8.53 -6.77 -3.89 -07.5
-11.3 -9.19 -10.41 -10.13 -09.3 -09.0 -8.65 00000 -7.01 -06.5
-10.33 -10.16 -8.71 -9.36 -9.28 -7.05 -5.91 -7.22 -6.27 -5.42
-9.98 -9.92 -9.27 -8.07 -6.72 -3.59 -3.56 -01.9 -03.8 -05.0
-9.82 -09.8 -9.35 -8.15 -6.25 -1.55 00000 -2.11 -4.44 -4.46
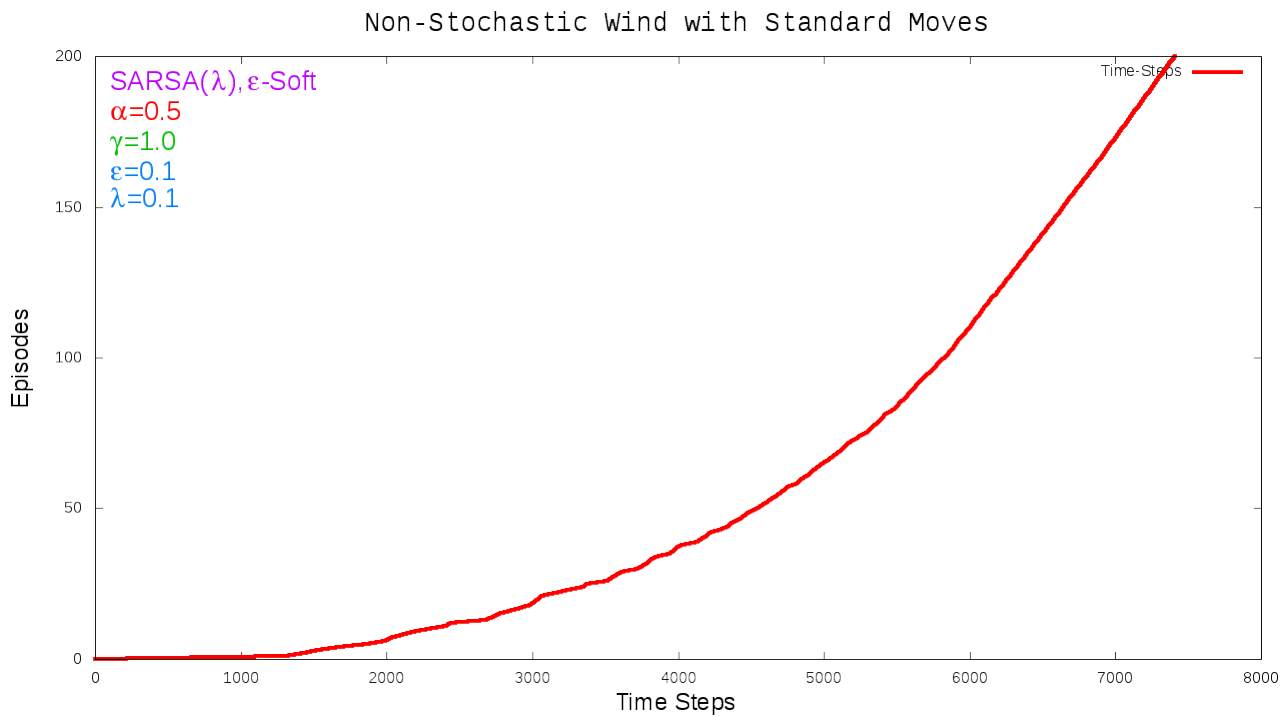
2.7 Data "D7.txt"
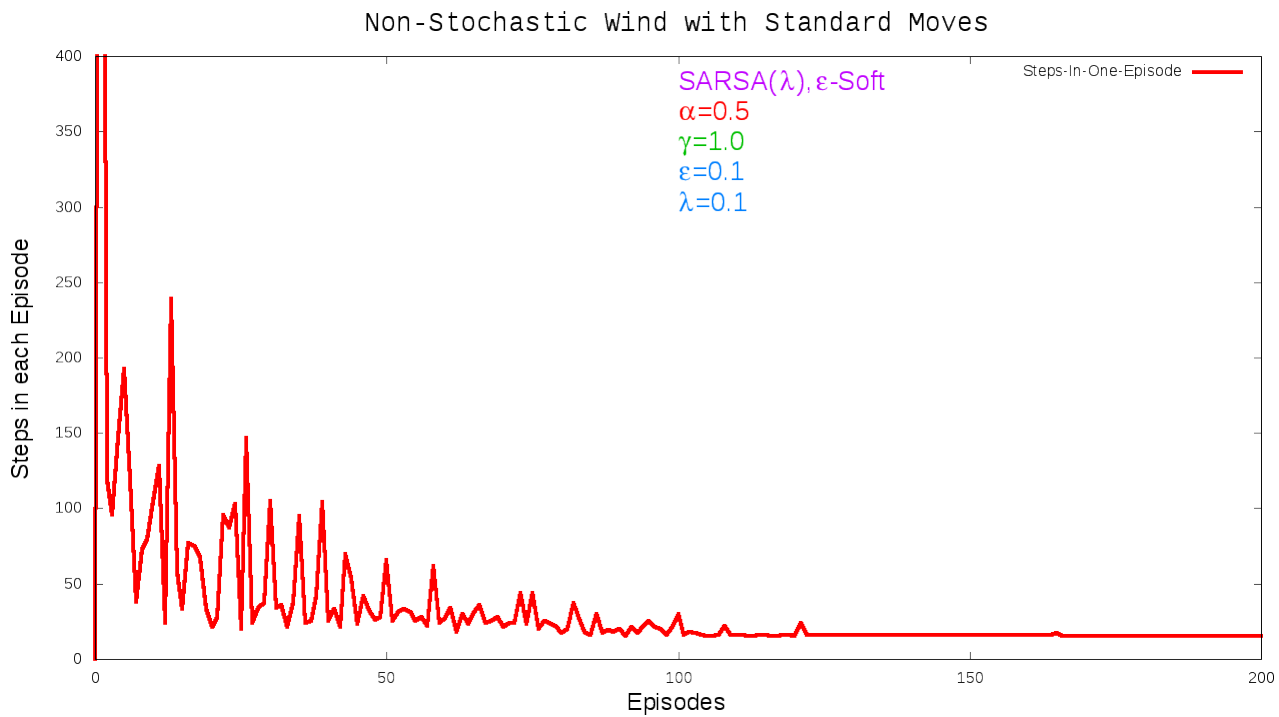

Figure 14: Episodes Vs Time-Step : Standard move

Figure 15: Episodes Vs Steps-in-each-episode : Standard move

State Values
    -13.81 -13.68 -12.96 -12.05 -11.02 -10.0 -09.0 -08.0 -07.0 -06.0
    -13.84 -13.59 -13.0 -12.03 -11.0 -10.0 -9.04 -8.03 -7.03 -05.0
    -14.03 -14.0 -13.0 -12.0 -11.0 -10.05 -9.02 -2.72 -3.94 -04.0
    -15.0 -14.0 -13.0 -12.0 -11.05 -9.94 -8.76 000.0 -3.35 -03.0
    -14.19 -13.51 -12.74 -12.02 -10.76 -9.41 000.0 000.0 -01.0 -02.0
    -13.57 -12.93 -12.15 -11.47 -10.18 000.0 000.0 000.0 -0.87 -1.41
    -13.17 -12.48 -11.83 -11.04 000.0 000.0 000.0 000.0 000.0 -1.06
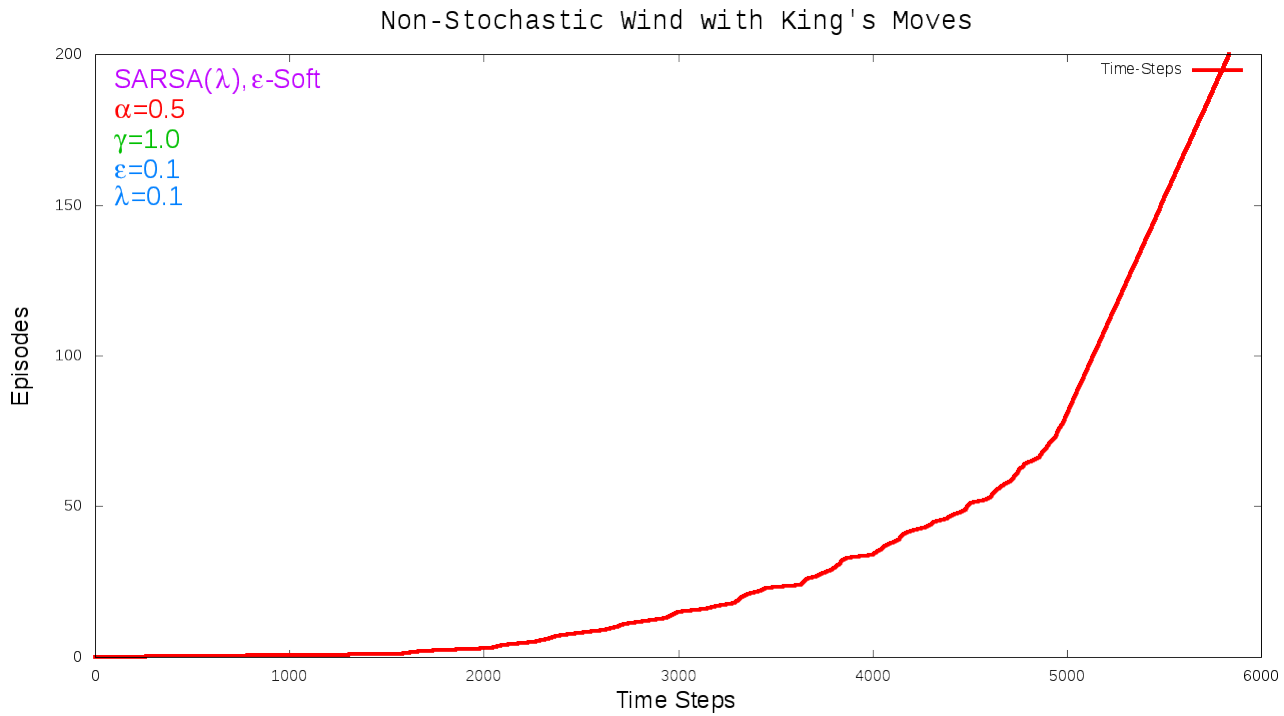
2.8   Data "D8.txt"

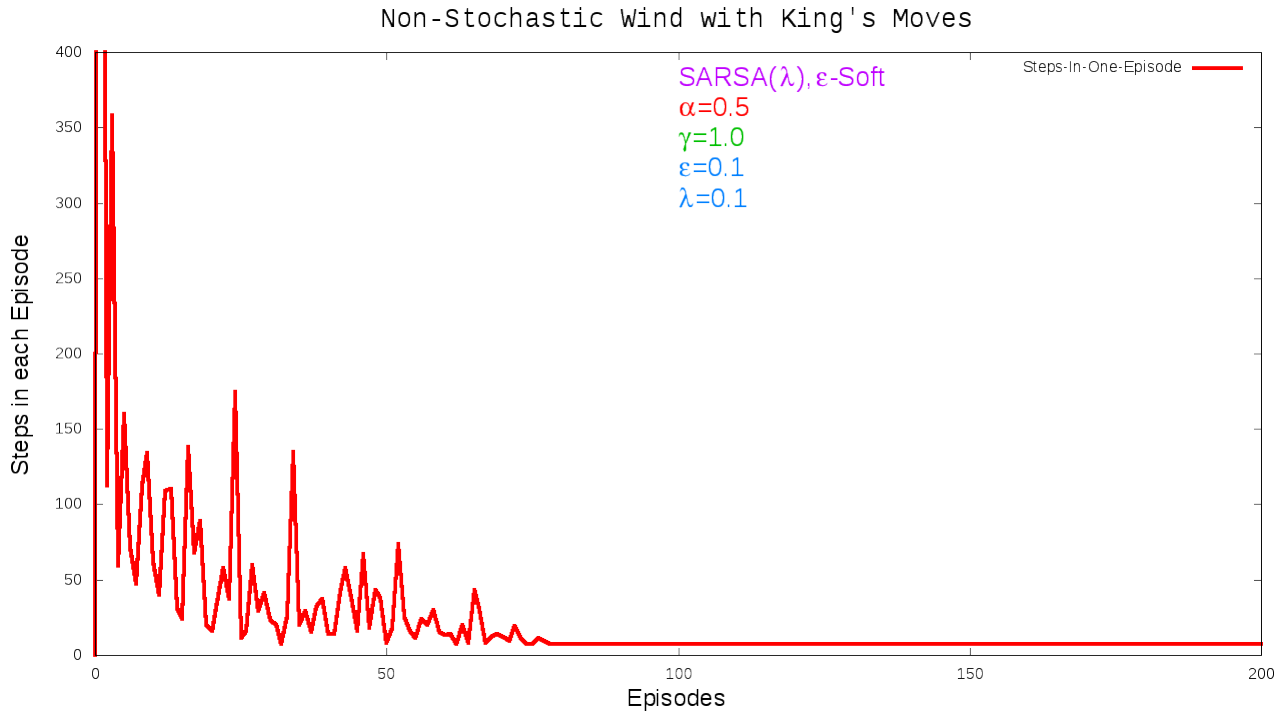Figure 16: Episodes Vs Time-Step : King's move

Figure 17: Episodes Vs Steps-in-each-episode : King's move

State Values
        -5.74 -5.88 -6.55 -7.15 -7.53 -7.02 -6.04 -5.05 -5.06 -04.2
        -5.63 -5.76 -5.71 -6.11 -06.3 -6.33 -5.94 -5.05 -4.04 -3.04
        -05.7 -06.0 -5.47 -5.54 -5.49 -05.2 -4.41 -3.67 -02.8 -2.04
        -06.0 -5.48 -05.0 -4.98 -4.36 -3.92 -3.59 000.0 -1.03 -1.12
        -5.32 -5.13 -4.52 -04.0 -03.0 -02.0 -01.0 000.0 000.0 -0.57
        -4.88 -4.69 -4.41 -3.61 -2.67 -1.78 -00.9 000.0 000.0 000.0
        -4.67 -4.52 -04.2 -3.52 000.0 000.0 000.0 000.0 000.0 000.0
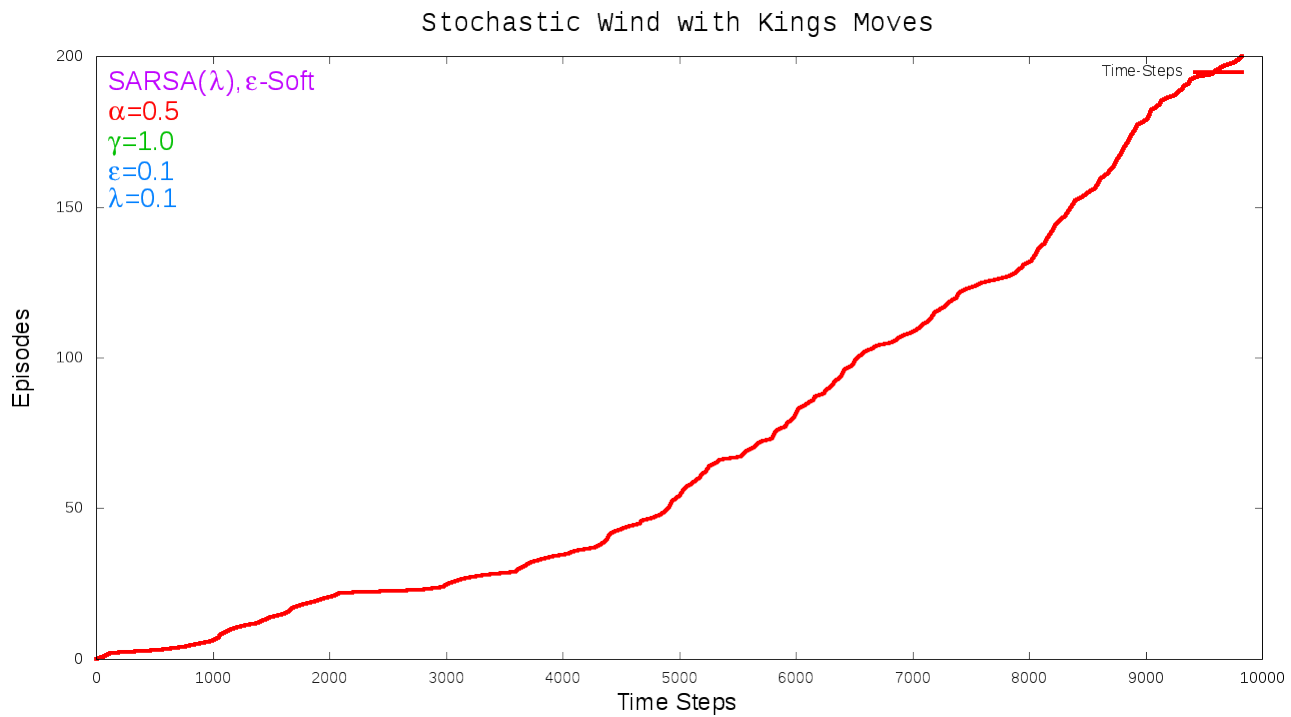
2.9   Data "D9.txt"

Stochastic Wind with Kings Moves

Figure 18: Episodes Vs Time-Step : Stochastic King's move

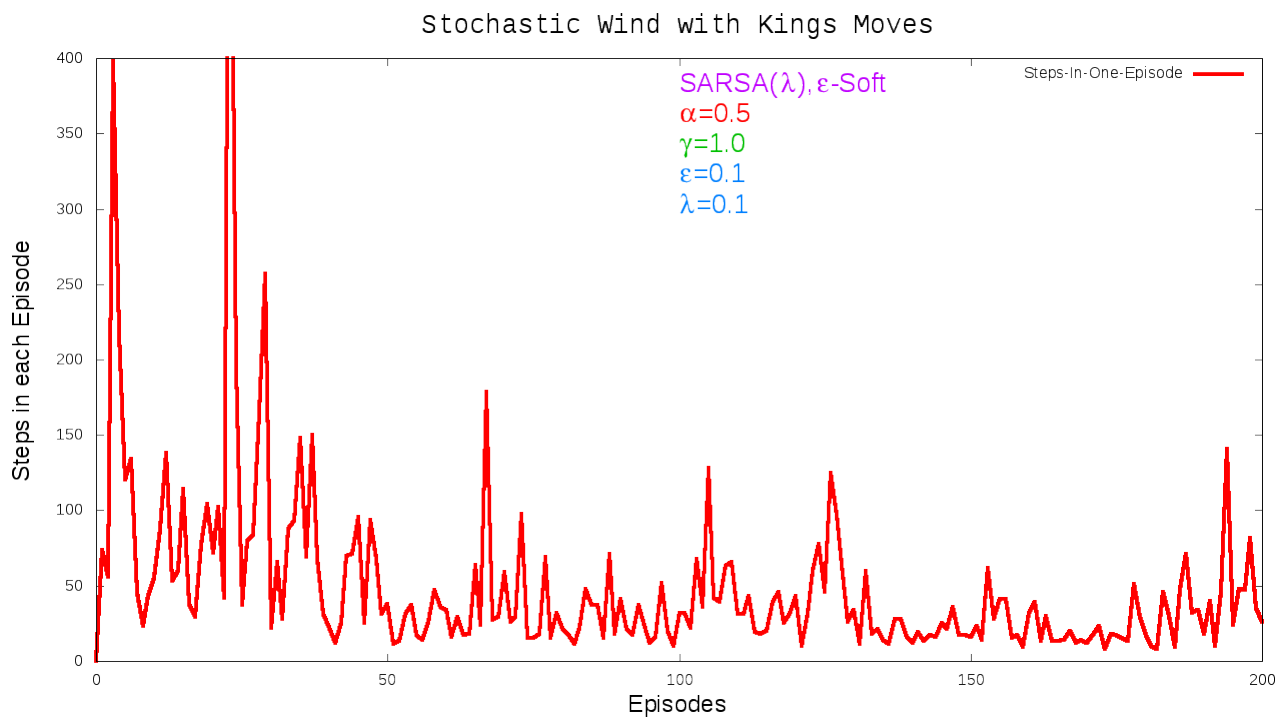Stochastic Wind with Kings Moves

Figure 19: Episodes Vs Steps-in-each-episode : Stochastic King's move

14

State Values
    -9.23 -9.16 -9.48 -9.61 -09.6 -9.03 -8.76 -7.55 -08.4 -8.24
    -9.11 -9.17 -9.01 -9.02 -9.09 -8.52 -6.57 -8.24 -8.19 -7.38
    -9.06 -9.49 -08.8 -8.49 -8.14 -8.16 -7.75 -02.5 -7.03 -6.85
    -9.47 -8.67 -8.61 -7.98 -08.0 -6.96 -7.86 000.0 -6.64 -6.18
    -8.74 -8.68 -7.96 -7.73 -7.19 -6.93 -5.53 -6.17 -3.45 -5.43
    -8.37 -8.15 -7.89 -7.83 -5.38 -5.93 -1.42 -3.91 -2.86 -4.77
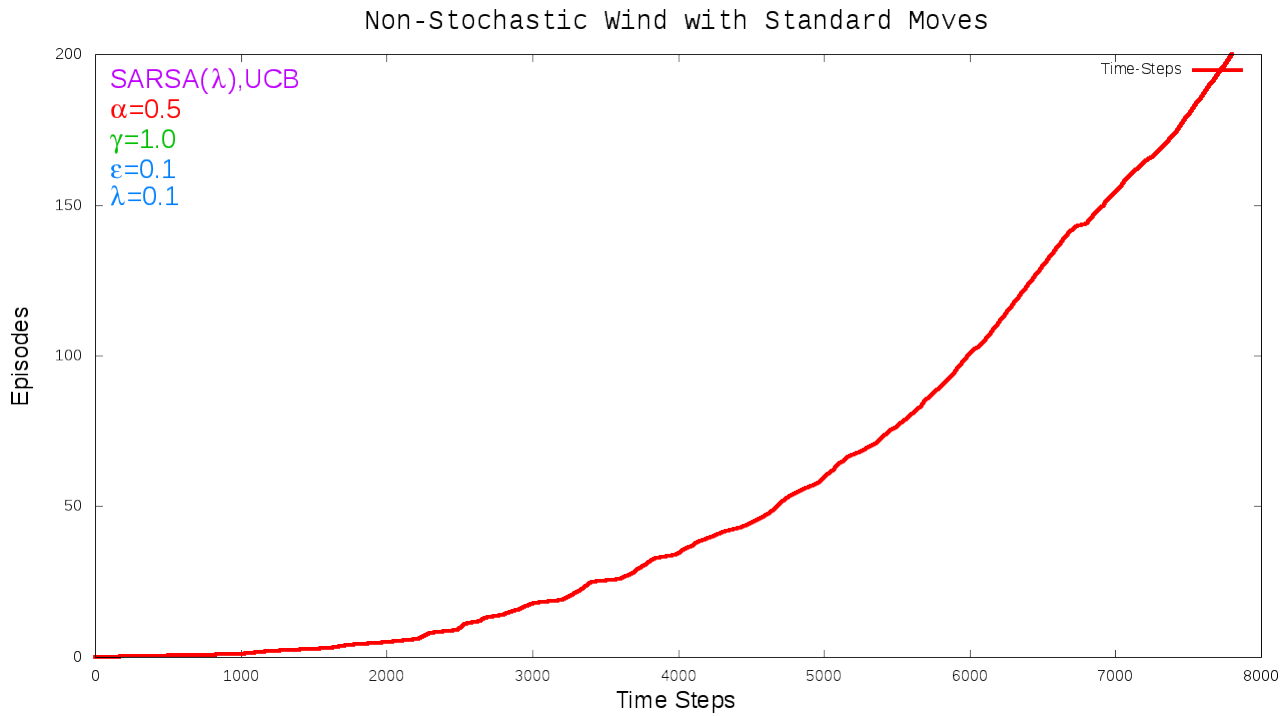   -7.82 -7.85 -7.46 -7.04 -5.44 -2.46 -0.77 -0.56 -4.26 -4.18

2.10   Data "D10.txt"
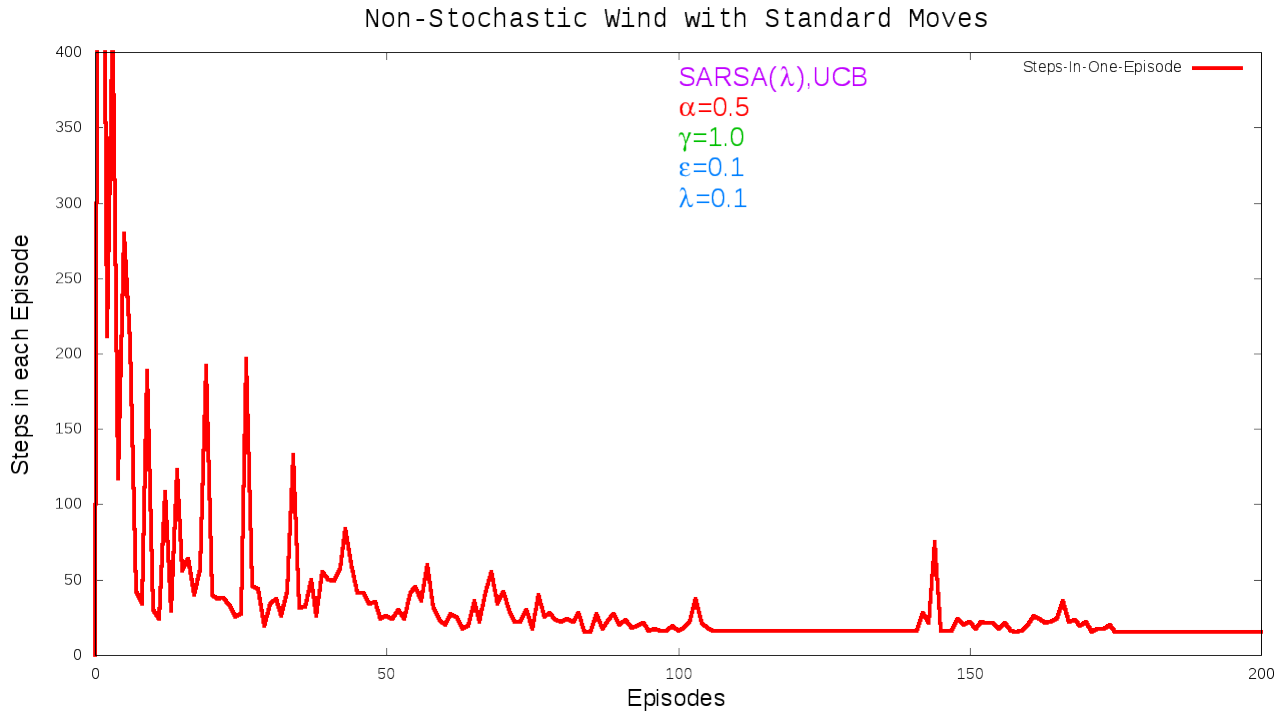

Figure 20: Episodes Vs Time-Step : Standard move

15

Figure 21: Episodes Vs Steps-in-each-episode : Standard move

State Values
-14.29 -14.27 -13.49 -12.32 -11.09 -10.02 -09.0 -08.0 -07.0 -06.0
-14.5 -14.37 -13.48 -12.37 -11.46 -10.0 -9.83 -8.67 -7.48 -05.0
-14.69 -14.99 -13.81 -12.71 -11.0 -10.87 -9.13 -6.15 -6.36 -04.0
-15.04 -14.02 -13.01 -12.0 -11.76 -10.27 -8.44 000.0 -5.69 -03.0
-14.93 -14.19 -13.36 -12.7 -10.84 -8.92 000.0 -0.88 -01.0 -02.0
-14.02 -13.5 -12.56 -11.86 -9.68 000.0 000.0 000.0 -1.67 -2.68
-13.53 -12.91 -11.94 -10.75 000.0 000.0 000.0 000.0 -1.15 -1.93
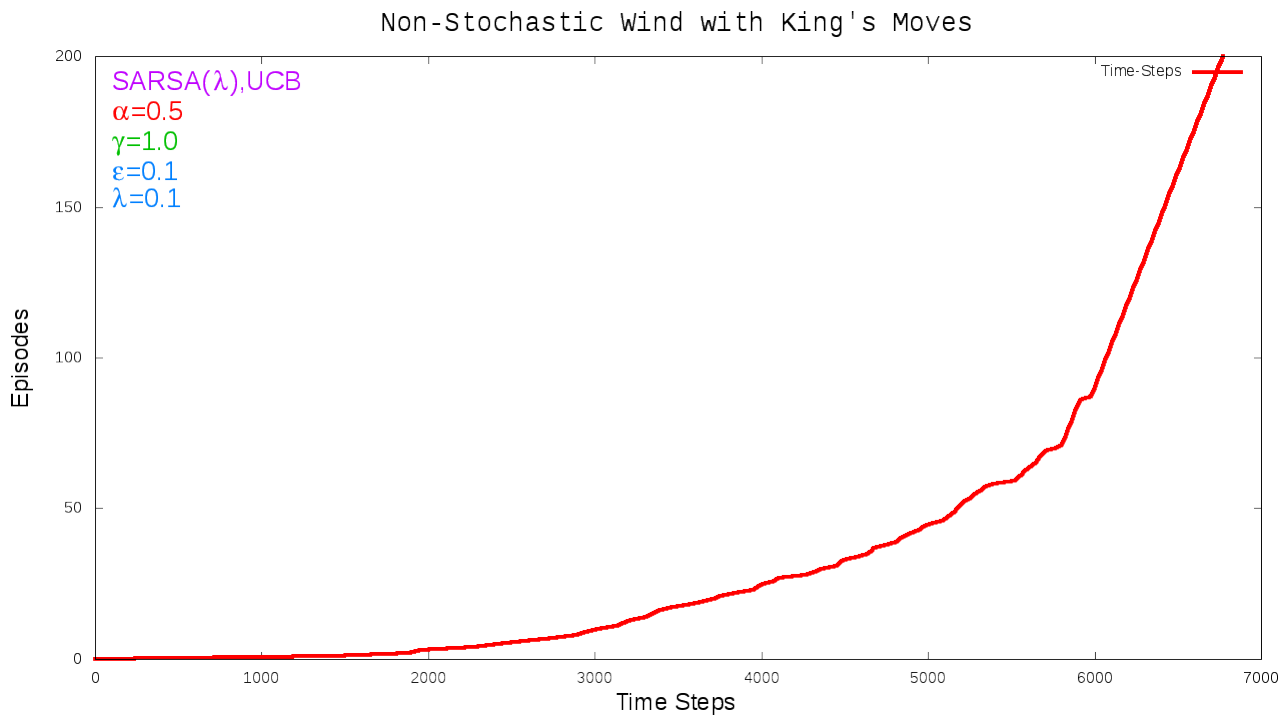
2.11 Data "D11.txt"

Figure 22: Episodes Vs Time-Step : King's move
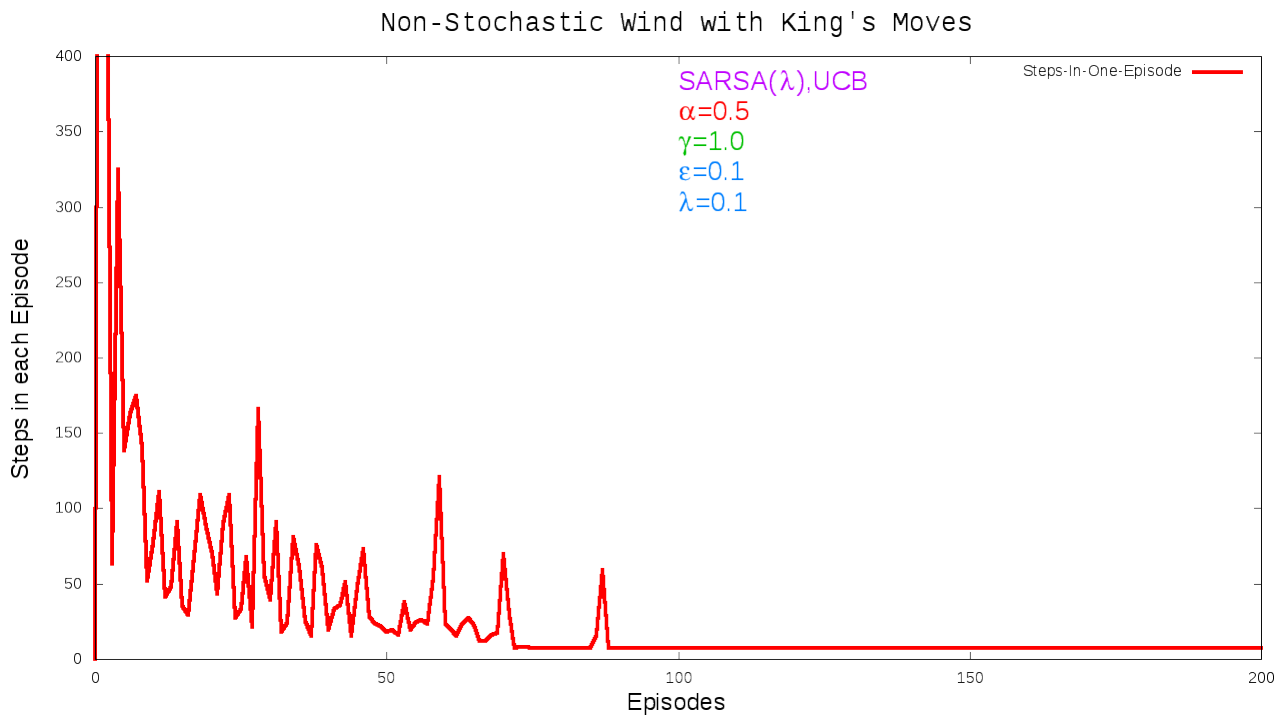

Figure 23: Episodes Vs Steps-in-each-episode : King's move

State Values
-7.35 -7.27 -8.06 -8.72 -8.83 -8.08 -7.05 -6.16 -5.05 -4.05
-6.93 -6.82 -6.96 -7.46 -07.8 -7.92 -7.27 -6.04 -5.12 -3.05
-6.74 -06.0 -6.51 -6.74 -6.54 -6.78 -06.2 -4.23 -3.65 -2.03
-06.0 -6.18 -05.0 -6.06 -5.46 -5.16 -4.33 000.0 -1.01 -1.97
-6.11 -5.97 -5.34 -04.0 -03.0 -02.0 -01.0 000.0 -0.97 -1.07
-5.73 -5.31 -4.87 -4.09 -3.06 -1.53 -0.52 000.0 000.0 -0.56
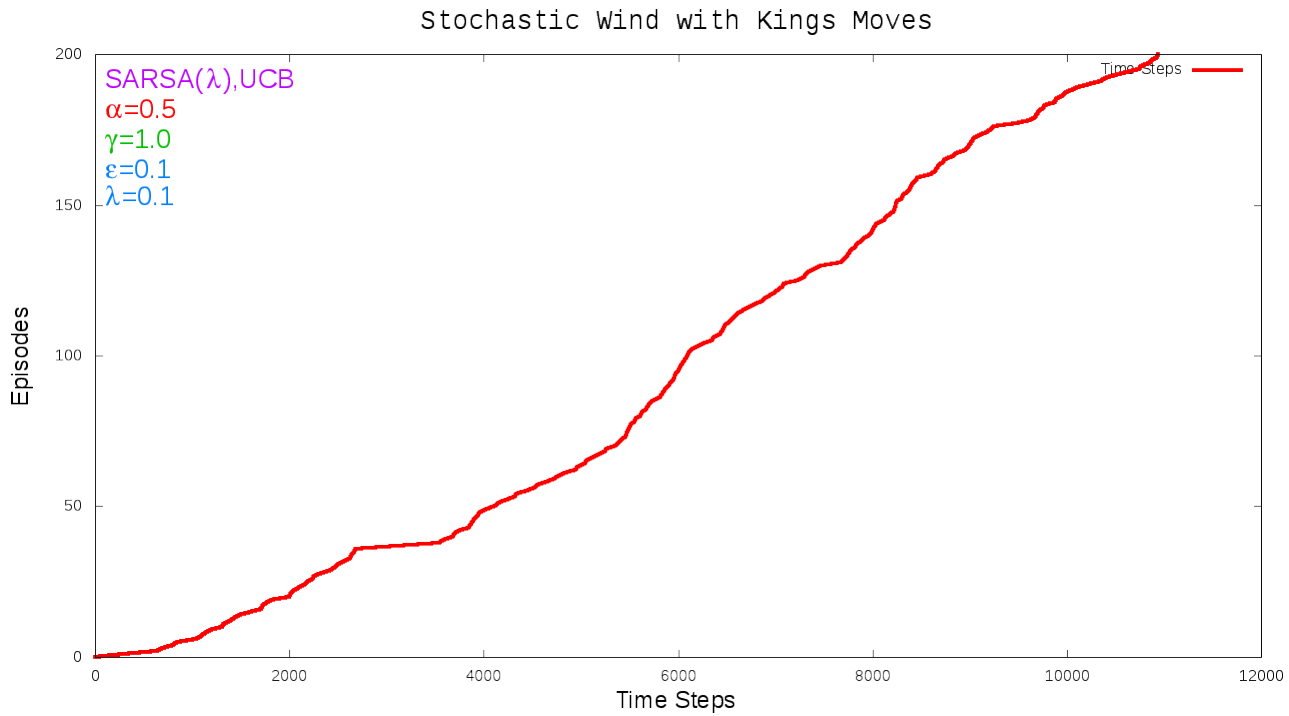-5.78 -5.47 -04.8 -4.09 000.0 000.0 000.0 000.0 000.0 000.0

2.12 Data "D12.txt"

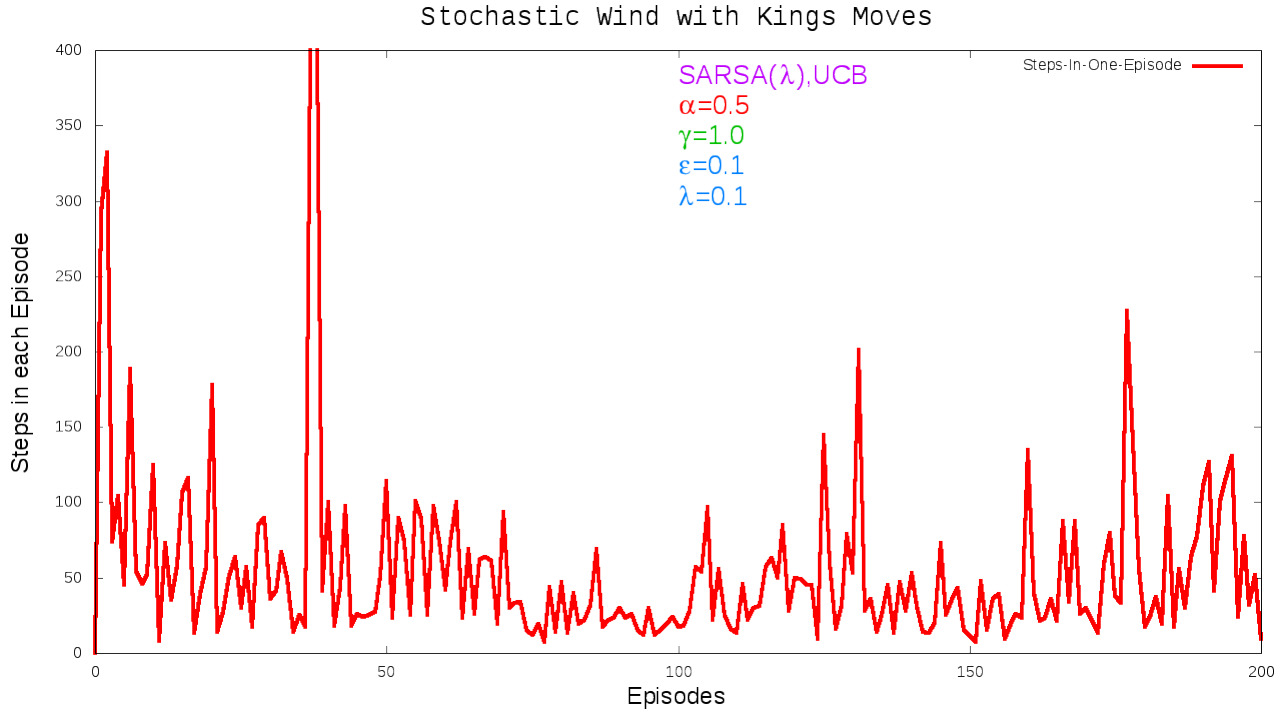Figure 24: Episodes Vs Time-Step : Stochastic King's move

18

Figure 25: Episodes Vs Steps-in-each-episode : Stochastic King's move

State Values
-10.38 -10.42 -10.91 -11.01 -11.3 -11.33 -10.54 -10.29 -10.36 -10.04
-10.36 -10.29 -10.52 -10.93 -11.32 -10.35 -10.51 -10.05 -10.28 -8.87
-10.35 -10.13 -10.14 -10.04 -10.06 -9.39 -3.13 -9.91 -7.98 -6.88
-10.32 -09.0 -8.77 -09.9 -09.2 -8.95 -8.73 000.0 -8.27 -4.59
-9.93 -9.57 -7.78 -9.16 -7.54 -6.89 -5.44 -7.15 -5.35 -5.84
-09.5 -8.94 -7.99 -6.73 -4.55 -03.7 -4.67 -5.06 -4.95 -4.85
-09.1 -08.8 -7.58 -7.79 -3.86 -02.8 -1.27 -0.64 -3.15 -4.35

3   Conclusion

SARSA(λ) performs better than SARSA(0) at the cost of higher complexity.

References

[1] Sarsa: On-policy td control.