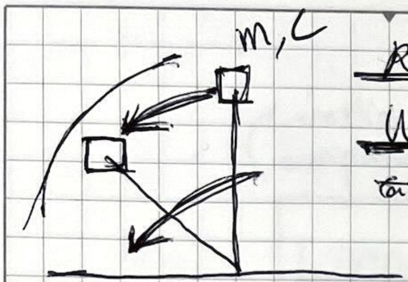
Title ~~MPRL DDA~~ Game Design (RL)　　Date / Time 24/7/10　　Page No.



RoM.

WL : arrival time diff. → distance btw target curs.

~~car racing following~~ X　　Vicon data → target ?

? chromosome (difficulty) — (inertia / damping / rom) — 1 DOF ? 높음

① W/L

②

③

~~now~~ present vs. past

trial n-1 trajectory vs. n trajectory

24 FPS = 24 Hz

W : 초록색
L : 빨간색

~~Score = time left (−∞, 10]~~

고거시 나중 앞지르면 score − |W − L|

1st trial. ~~[illegible]~~

앞지른 시간　　뒤 따라간 시간

reward = score = scala

g_1, g_2 satisfied　　~~$r(t) = \frac{1}{|W-L|}$~~

g_3 ?　　연승 보너스

by SW guidance

k_C　k_n　　~~$r = \frac{1}{|W-L|}$~~

$r = -|W-L|$ reward

winning time (green)　　losing time (red)

Inertia　Damping　ROM
↑+ ↓−　↑+ ↓−　↑+ ↓−

chromosome (q-table)

P_{1-}	P_{1+}	P_{2-}	P_{2+}	P_{3-}	P_{3+}
−10%	+10%	〃	〃	〃	〃

M　C　Δ

$$-1 \le r = -\frac{|W-L|}{|W+L|} \le 0$$

W = L !

L = 0 or W = 0

$$-1 \le -\frac{2|W-L|}{|W+L|} + 1 \le 1$$

$$-0.5 \le -\frac{|W-L|}{|W+L|} + \frac{1}{2} \le 0.5$$

① $p\%2==0$ $c_p = (1-\alpha)\,c_p + \alpha(sign(p) + rw)$

$==1$ $c_p = (1-\alpha)\,c_p - \alpha(sign(p) + rw)$) ?

reward ~~of learning~~ ... RL algorithm validation with no state definition.

$p=1$ speed up ?

$$-1 \le r = -\frac{2|W-L|}{|W+L|} + 1 \le 1$$

~~$-1 \le r =$~~

$$\boxed{-1 \le 1\left(-\frac{2|W-L|}{|W+L|} + 1\right) \le 1}$$

break ind.

$s = \langle$ inertia, damping, rpm $\rangle$

$+ \; - \quad + \; - \quad + \; -$

기존기준 sec vs. 리그기준 sec.

$-|W-L| +$ $\quad \frac{1}{e^{|W-L|}}$.

~~$-|W+L|$~~ $\cdot e^{-|W-L|}$

$\gamma \cdot e \cdot e^{-|W-L|}$

$\gamma \cdot e^{-|\Delta t|} \cdot e^{-|\Delta s|}$

$\gamma \ge 1 + 0.1$

$+ 0.1$

$\vdots$

연속보상 progress