CHOW, Hau Cheung Jasper
Student ID: 20589533
hcjchow@connect.ust.hk
COMP4211 Assignment 3
=====================
Initial state:

	facebook	quit	study	pub	sleep
FACEBOOK	-1	-1	-	-	-
CLASS1	-1.9	-	-2	-	-
CLASS2	-	-	-1.1	-	0
CLASS3	-	-	10	1	-

After observing state CLASS3, and taking action pub, we obtain immediate reward $r = 1$ and observe state CLASS3.
Observe $max_{a'}\widehat{Q}(CLASS3, a') = 10$ *when a' = study*

$\widehat{Q}(s, a) \leftarrow r + \gamma max_{a'}\widehat{Q}(s', a') = r + \gamma max_{a'}\widehat{Q}(CLASS3, a') = 1 + \gamma(10) = 10$
$\widehat{Q}(CLASS3, pub) \leftarrow r + \gamma max_{a'}\widehat{Q}(CLASS3, a') = 1 + \gamma(10) = 10$

New table:

	facebook	quit	study	pub	sleep
FACEBOOK	-1	-1	-	-	-
CLASS1	-1.9	-	-2	-	-
CLASS2	-	-	-1.1	-	0
CLASS3	-	-	10	**10**	-

After observing state CLASS3, and taking action pub, we obtain immediate reward $r = 1$ and observe state CLASS2.
Observe $max_{a'}\widehat{Q}(CLASS2, a') = 0$ *when a' = sleep*

$\widehat{Q}(s, a) \leftarrow r + \gamma max_{a'}\widehat{Q}(s', a')$
$\widehat{Q}(CLASS3, pub) \leftarrow r + \gamma max_{a'}\widehat{Q}(CLASS2, a') = 1 + \gamma(0) = 1$

New table:

=====================

	facebook	quit	study	pub	sleep
FACEBOOK	-1	-1	-	-	-
CLASS1	-1.9	-	-2	-	-
CLASS2	-	-	-1.1	-	0
CLASS3	-	-	10	**1**	-

After observing state CLASS2, and taking action study, we obtain immediate reward $r = -2$ and observe state CLASS3.

Observe $max_{a'}\widehat{Q}(CLASS3, a') = 10$ when $a' = study$

$\widehat{Q}(s, a) \leftarrow r + \gamma max_{a'}\widehat{Q}(s', a')$
$\widehat{Q}(CLASS2, study) \leftarrow r + \gamma max_{a'}\widehat{Q}(CLASS3, a') = -2 + \gamma(10) = 7$

New table:

	facebook	quit	study	pub	sleep
FACEBOOK	-1	-1	-	-	-
CLASS1	-1.9	-	-2	-	-
CLASS2	-	-	**7**	-	0
CLASS3	-	-	10	1	-

After observing state CLASS3, and taking action pub, we obtain immediate reward $r = 1$ and observe state CLASS1.

Observe $max_{a'}\widehat{Q}(CLASS1, a') = -1.9$ when $a' = facebook$

$\widehat{Q}(s, a) \leftarrow r + \gamma max_{a'}\widehat{Q}(s', a')$
$\widehat{Q}(CLASS3, pub) \leftarrow r + \gamma max_{a'}\widehat{Q}(CLASS1, a') = 1 + \gamma(-1.9) = -0.71$

New table:

	facebook	quit	study	pub	sleep
FACEBOOK	-1	-1	-	-	-

CLASS1	-1.9	-	-2	-	-
CLASS2	-	-	7	-	0
CLASS3	-	-	10	**-0.71**	-

After observing state CLASS1, and taking action study, we obtain immediate reward $r = -2$ and observe state CLASS2.

Observe $max_{a'}\widehat{Q}(CLASS2, a') = 7$ *when a' = study*

$\widehat{Q}(s, a) \leftarrow r + \gamma max_{a'}\widehat{Q}(s', a')$
$\widehat{Q}(CLASS1, study) \leftarrow r + \gamma max_{a'}\widehat{Q}(CLASS2, a') = -2 + \gamma(7) = 4.3$

New table:

	facebook	quit	study	pub	sleep
FACEBOOK	-1	-1	-	-	-
CLASS1	-1.9	-	**4.3**	-	-
CLASS2	-	-	7	-	0
CLASS3	-	-	10	-0.71	-

After observing state CLASS2, and taking action sleep, we obtain immediate reward $r = 0$ and observe state SLEEP.

Observe $max_{a'}\widehat{Q}(SLEEP, a') = 0$

$\widehat{Q}(s, a) \leftarrow r + \gamma max_{a'}\widehat{Q}(s', a')$
$\widehat{Q}(CLASS2, sleep) \leftarrow r + \gamma max_{a'}\widehat{Q}(SLEEP, a') = 0$

New table:

	facebook	quit	study	pub	sleep
FACEBOOK	-1	-1	-	-	-
CLASS1	-1.9	-	4.3	-	-
CLASS2	-	-	7	-	**0**

CHOW, Hau Cheung Jasper
Student ID: 20589533
hcjchow@connect.ust.hk
COMP4211 Assignment 3
=====================

CLASS3	-	-	10	-0.71	-