QLearning For Videogames

Steven Walton
University of Oregon
Github: stevenwalton/Retro-Learner

CIS 510: Multi-Agent Systems
3 June 2019

What is Q-Learning?

- Markov Decision Processes (MDPs) can represent decision making processes with random outcomes.

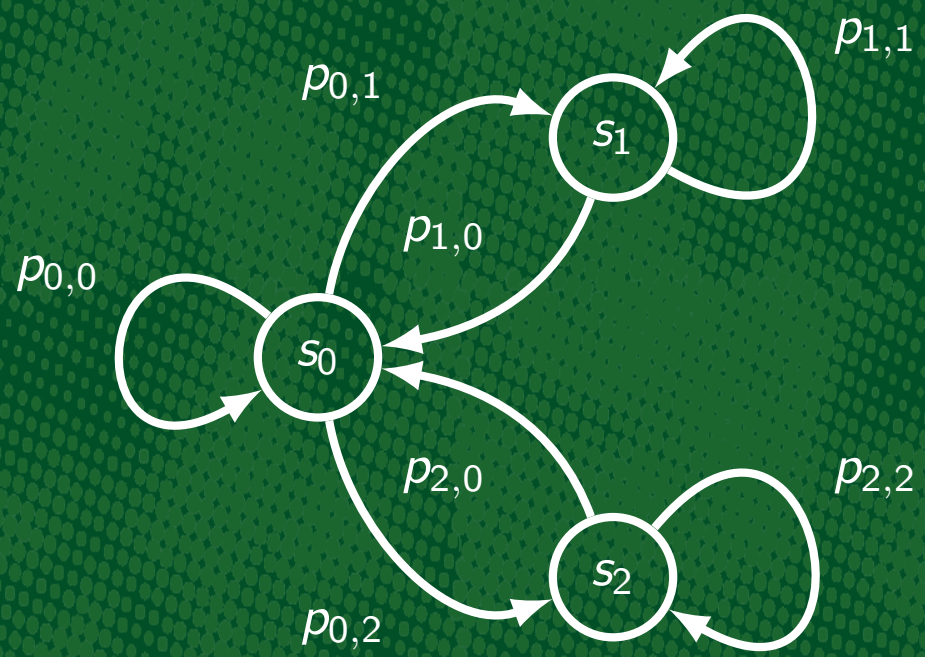


Figure: Sample Markov Decision Process

What is Q-Learning?

- Markov Decision Processes (MDPs) can represent decision making processes with random outcomes.
- We want to find an optimal policy (set of actions)

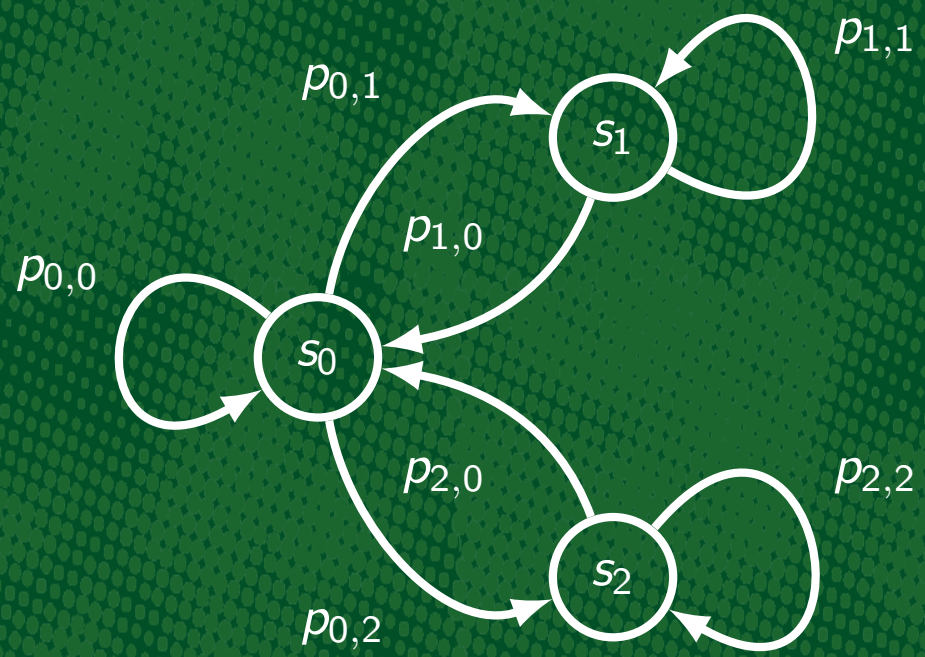


Figure: Sample Markov Decision Process

What is Q-Learning?

- Markov Decision Processes (MDPs) can represent decision making processes with random outcomes.
- We want to find an optimal policy (set of actions)
- We learn with the Bellman Equation, using an iterative process

$Q(s,a) = Q(s,a) + \alpha(r + \gamma \max_{a'} Q(s',a') - Q(s,a))$

$\alpha =$ learning rate

$r =$ reward

$\gamma =$ discount

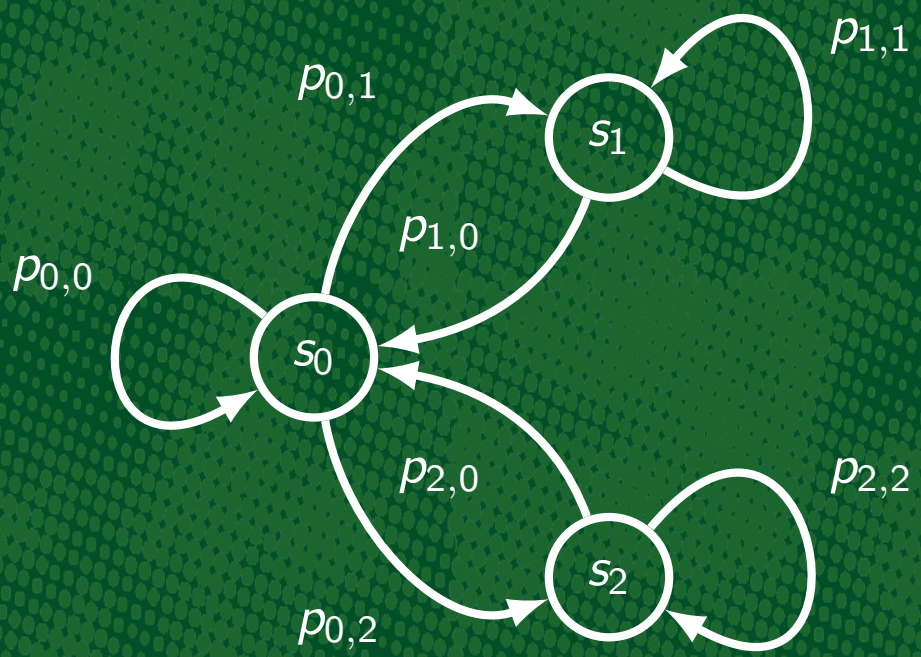


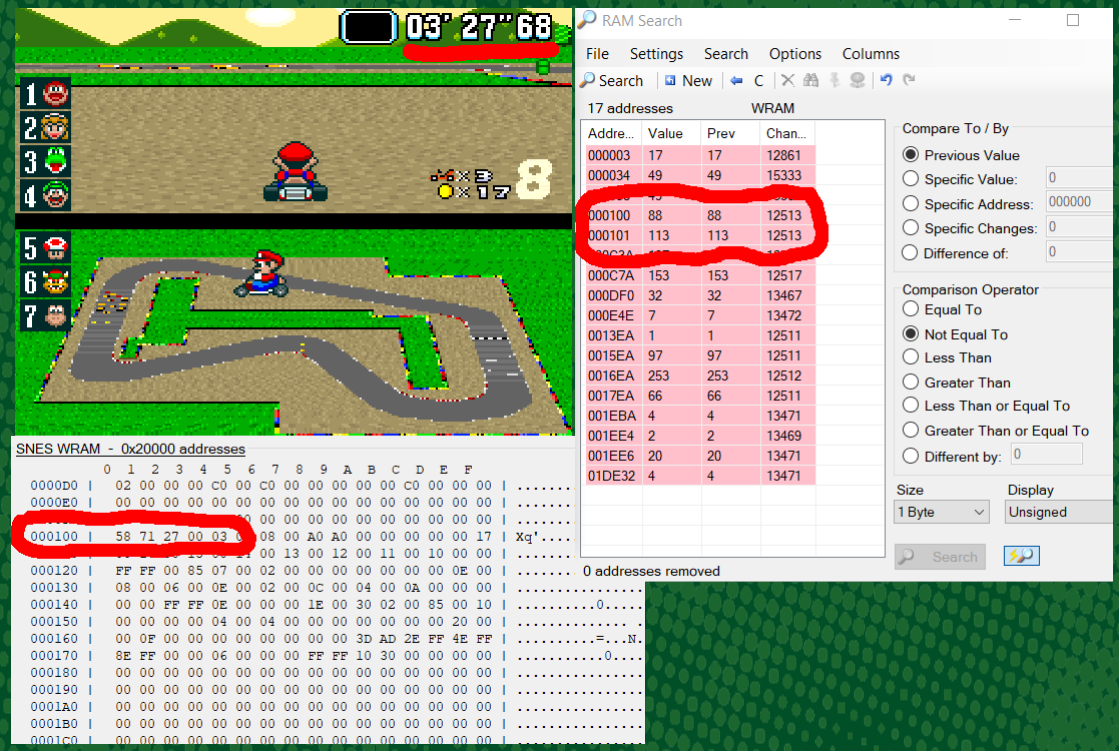
Figure: Sample Markov Decision Process

Framework

- OpenAI Gym: Allows for abstracts out environments and allows researchers to focus on creating algorithms.
- OpenAI Retro: Wrapps Gym framework to focus on retro videogames.
- Allows same code to be run on any game that is integrated.
- Creates an environment to (supposedly) allow for easy integration of new games

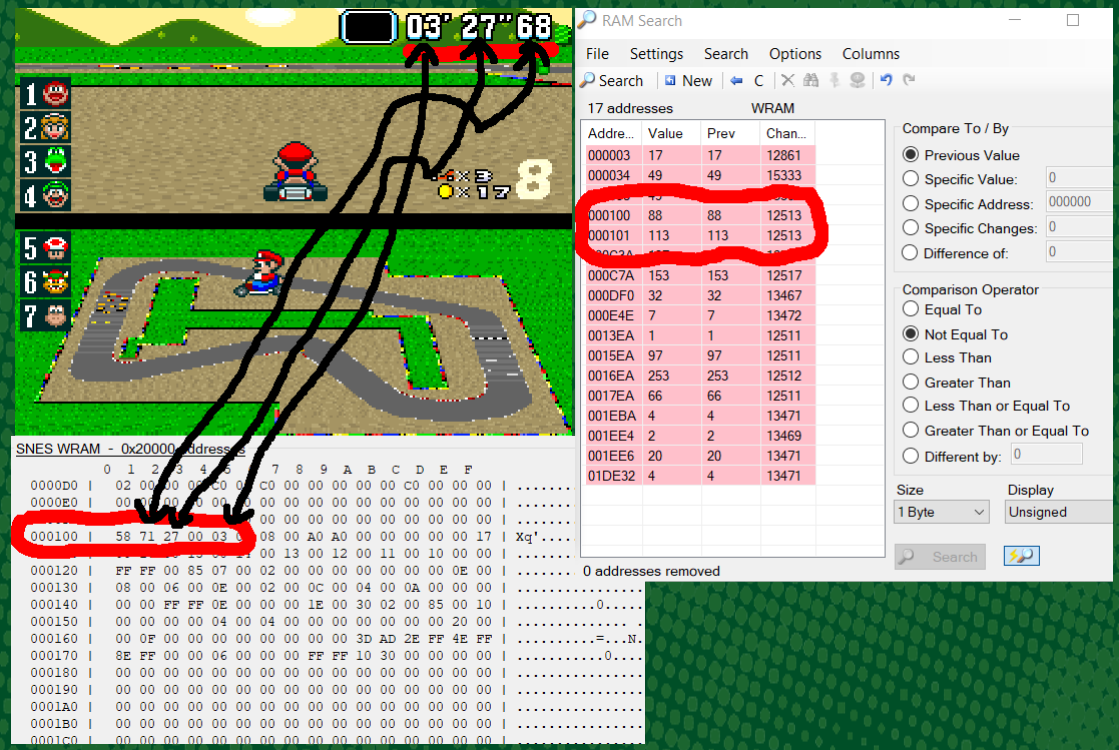
Integration

- Need to collect RAM values (used BizHawk).



Integration

- Need to collect RAM values (used BizHawk).



Integration

- Need to collect RAM values (used BizHawk).
- Convert Hex to decimal and add rambase number and integrate into scenario and data files

```
{
    "info": {
        "lap": {
            "address": 8261825,
            "type": "|u1"
        },
        "lapsize": {
            "address": 8257864,
            "type": "|u1"
        },
        "checkpoint": {
            "address": 8261852,
            "type": "|u1"
        },
        "rank": {
            "address": 8261696,
            "type": "|u1"
        },
        "coins": {
            "address": 8261120,
            "type": "|u1"
        },
        "minute": {
            "address": 8257796,
            "type": "|u1"
        },
        "second": {
            "address": 8257794,
            "type": "|u1"
        },
        "milisecond": {
            "address": 8257793,
            "type": "|u1"
        }
    }
}
```

```
{
    "done": {
        "variables": {
            "finish": {
                "reference": "coins",
                "op": "ex"
            }
        }
    },
    "reward": {
        "variables": {
            "rank": {
                "reward": 1.0
            }
        }
    },
    "actions": [
        [[], ["UP"], ["DOWN"]],
        [[], ["LEFT"], ["RIGHT"]],
        [[], ["A"], ["B"], ["Y"], ["START"], ["R"], ["R", "A"], ["R", "B"], ["R", "Y"], ["R", "A", "B"], ["R", "A", "Y"], ["A", "B"], ["A", "Y"]]
}
```

Integration

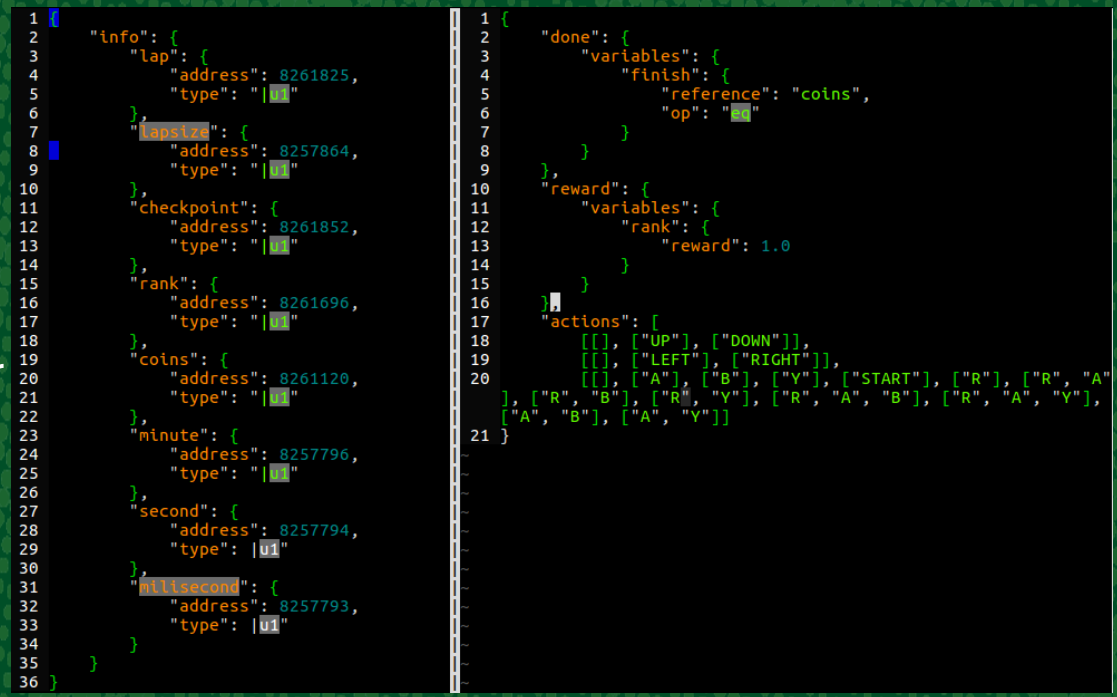
- Need to collect RAM values (used BizHawk).
- Convert Hex to decimal and add rambase number and integrate into scenario and data files
- Define actions and rewards for the game.

```
{
    "info": {
        "lap": {
            "address": 8261825,
            "type": "|u1"
        },
        "lapsize": {
            "address": 8257864,
            "type": "|u1"
        },
        "checkpoint": {
            "address": 8261852,
            "type": "|u1"
        },
        "rank": {
            "address": 8261696,
            "type": "|u1"
        },
        "coins": {
            "address": 8261120,
            "type": "|u1"
        },
        "minute": {
            "address": 8257796,
            "type": "|u1"
        },
        "second": {
            "address": 8257794,
            "type": |u1"
        },
        "milisecond": {
            "address": 8257793,
            "type": |u1"
        }
    }
}
```

```
{
    "done": {
        "variables": {
            "finish": {
                "reference": "coins",
                "op": "e"
            }
        }
    },
    "reward": {
        "variables": {
            "rank": {
                "reward": 1.0
            }
        }
    },
    "actions": [
        [[], ["UP"], ["DOWN"]],
        [[], ["LEFT"], ["RIGHT"]],
        [[], ["A"], ["B"], ["Y"], ["START"], ["R"], ["R", "A"], ["R", "B"], ["R", "Y"], ["R", "A", "B"], ["R", "A", "Y"], ["A", "B"], ["A", "Y"]]
}
```

Integration

- Need to collect RAM values (used BizHawk).
- Convert Hex to decimal and add rambase number and integrate into scenario and data files
- Define actions and rewards for the game.
- Not all RAM values are so easy to get nor define. (Lua scripts)

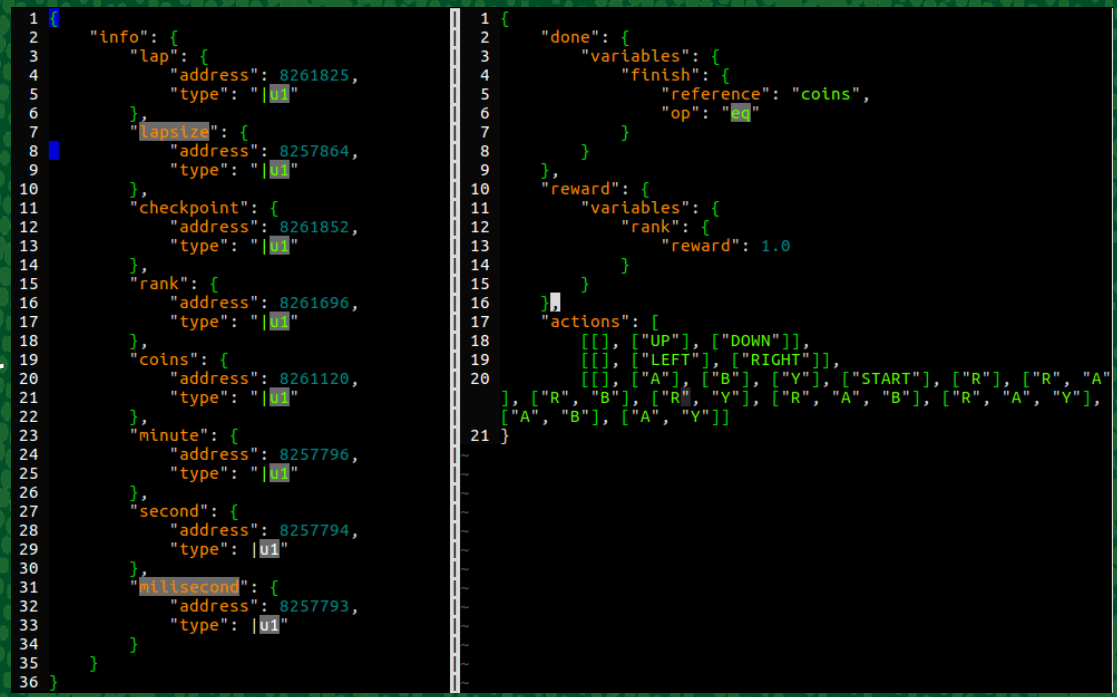


```
{
    "info": {
        "lap": {
            "address": 8261825,
            "type": "|u1"
        },
        "lapsize": {
            "address": 8257864,
            "type": "|u1"
        },
        "checkpoint": {
            "address": 8261852,
            "type": "|u1"
        },
        "rank": {
            "address": 8261696,
            "type": "|u1"
        },
        "coins": {
            "address": 8261120,
            "type": "|u1"
        },
        "minute": {
            "address": 8257796,
            "type": "|u1"
        },
        "second": {
            "address": 8257794,
            "type": |u1"
        },
        "milisecond": {
            "address": 8257793,
            "type": |u1"
        }
    }
}
```

```
{
    "done": {
        "variables": {
            "finish": {
                "reference": "coins",
                "op": "ne"
            }
        }
    },
    "reward": {
        "variables": {
            "rank": {
                "reward": 1.0
            }
        }
    },
    "actions": [
        [[], ["UP"], ["DOWN"]],
        [[], ["LEFT"], ["RIGHT"]],
        [[], ["A"], ["B"], ["Y"], ["START"], ["R"], ["R", "A"], ["R", "B"], ["R", "Y"], ["R", "A", "B"], ["R", "A", "Y"], ["A", "B"], ["A", "Y"]]
}
```

Integration

- Need to collect RAM values (used BizHawk).
- Convert Hex to decimal and add rambase number and integrate into scenario and data files
- Define actions and rewards for the game.
- Not all RAM values are so easy to get nor define. (Lua scripts)
- Thanks SethBling



```
{
    "info": {
        "lap": {
            "address": 8261825,
            "type": "|u1"
        },
        "lapsize": {
            "address": 8257864,
            "type": "|u1"
        },
        "checkpoint": {
            "address": 8261852,
            "type": "|u1"
        },
        "rank": {
            "address": 8261696,
            "type": "|u1"
        },
        "coins": {
            "address": 8261120,
            "type": "|u1"
        },
        "minute": {
            "address": 8257796,
            "type": "|u1"
        },
        "second": {
            "address": 8257794,
            "type": |u1"
        },
        "milisecond": {
            "address": 8257793,
            "type": |u1"
        }
    }
}
```

```
{
    "done": {
        "variables": {
            "finish": {
                "reference": "coins",
                "op": "ez"
            }
        }
    },
    "reward": {
        "variables": {
            "rank": {
                "reward": 1.0
            }
        }
    },
    "actions": [
        [[], ["UP"], ["DOWN"]],
        [[], ["LEFT"], ["RIGHT"]],
        [[], ["A"], ["B"], ["Y"], ["START"], ["R"], ["R", "A"], ["R", "B"], ["R", "Y"], ["R", "A", "B"], ["R", "A", "Y"], ["A", "B"], ["A", "Y"]]
}
```

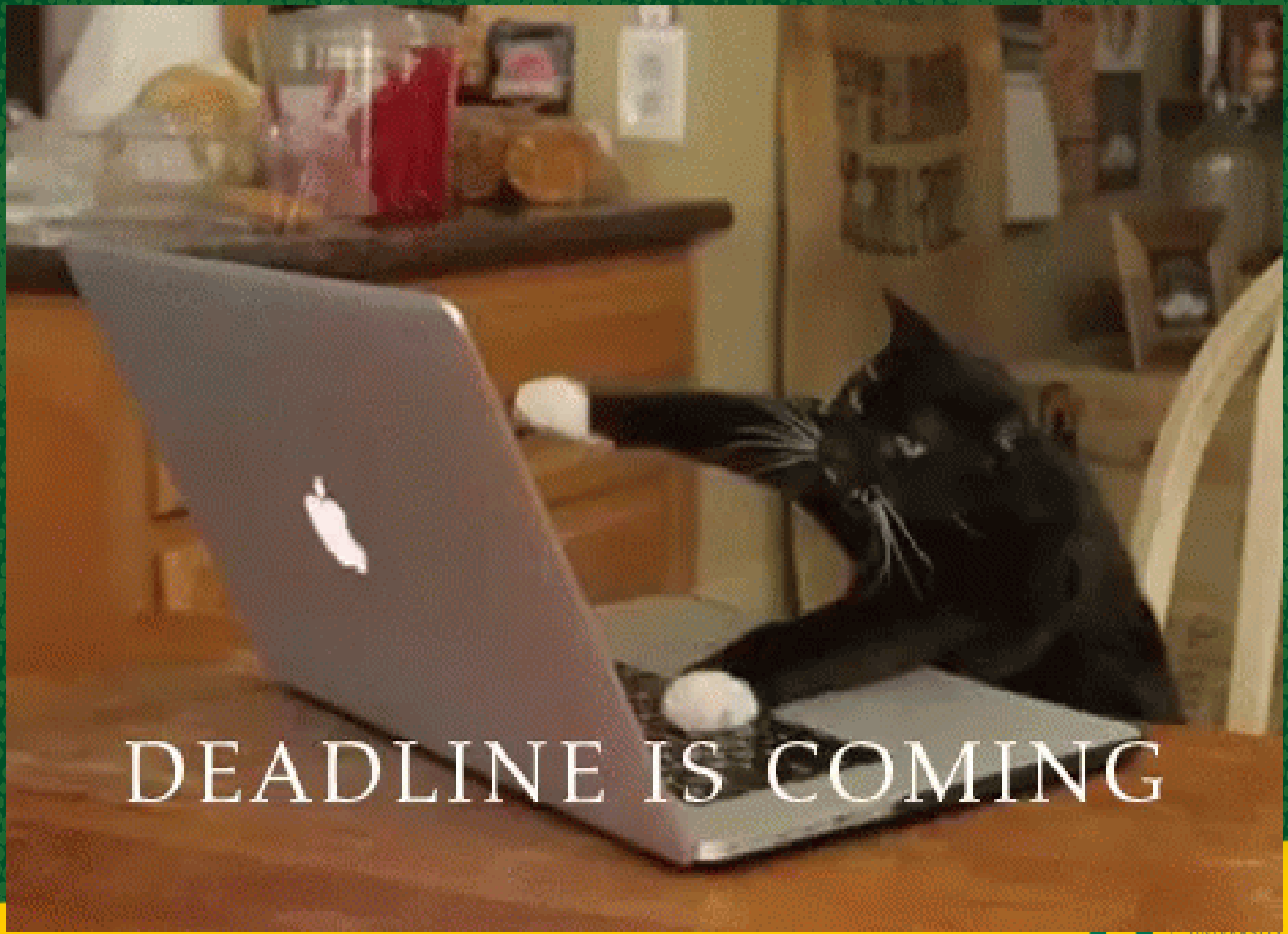
- Need to collect
(used BizHawk).
- Convert Hex to
add rambase nu
integrate into sc
data files
- Define actions a
the game.
- Not all RAM va
easy to get nor
scripts)
- Thanks SethBlin


```
": {
variables": {
    "finish": {
        "reference": "coins",
        "op": "
    }

ard": {
variables": {
    "rank": {
        "reward": 1.0
    }

ions": [
[], ["UP"], ["DOWN"]],
[], ["LEFT"], ["RIGHT"]],
[], ["A"], ["B"], ["Y"], ["START"], ["R"], ["R", "A"
"B"], ["R", "Y"], ["R", "A", "B"], ["R", "A", "Y"],
"], ["A", "Y"]]
```

So Let's Make Something Work



So Let's Make Something Work

Goals

- Get a program to work nicely with retro.

So Let's Make Something Work

Goals

- Get a program to work nicely with retro.
- Enable easy parameterisation of variables.

So Let's Make Something Work

Goals

- Get a program to work nicely with retro.
- Enable easy parameterisation of variables.
- Test test test

Let's see some games!

(Airstriker Genesis)

Figure: Airstriker Genesis Random: Max 160

Figure: Airstriker Genesis Longer Run: Max 940

Let's see some games!

(Donkey Kong Country)

Figure: Donkey Kong Country with lookahead 100, 10x

Let's see some games!

(Super Mario Brothers)

Figure: Super Mario Bros, low training at 10x

Figure: Super Mario Bros, a weekend of training at 10x