TASK BRIEF

Full-scale Autonomous Driving: Simulation-Based Policy Comparison

Submission details:

There are two assessments for this task, as detailed by the following tables. The whole task, and hence both assessments, is **to be completed in a team of 2 or 3**.

Type:	Progress update: verbal description and discussion of team progress.
Length:	Approximately 10 minutes per team.
Weighting:	4% of the overall subject grade.
Due Date:	Monday 16 October 2023, 16:15-18:15, i.e., during the workshop session.
Instructions:	• You must be present in the workshop session to complete this assessment. • This assessment is conducted with each team separately, i.e., it is **NOT** a presentation to the whole class.

Type:	Written report submission. One document is to be submitted per team.
Length:	**Maximum** page length of content: • 20 pages for a **team of 2**. • 30 pages for a **team of 3**. This page length: **includes** figures and tables; **excludes** appendices.
Weighting:	21% of the overall subject grade.
Due Date:	Monday 13 November 2023, 23:59.
Instructions:	• Upload report document to LMS. One report upload per team.

The following learning outcomes are demonstrated through this assessment activity:

- **ILO2)** Analyse and implement model-based methods for providing stability and performance, such as: PID, MPC, system identification, or adaptive control schemes.
- **ILO3)** Analyse and implement reinforcement learning (RL) methods for discrete-time dynamical systems with continuous spaces.
- **ILO4)** Compare and contrast the complexities of implementing decision-making and learning algorithms on real-world systems.

PURPOSE STATEMENT

The primary **technical purpose** of this task is to synthesize, verify, and compare multiple policies for a **full-sized** passenger vehicle **autonomously driving** along a **high-speed** road with **tight corners**.

The primary **professional purpose** of this task is that you take on the responsibility for all aspects of synthesis, verification, and analytical comparison. In other words, you are the student engineers who are tasked with **performing a rigorous comparison** and **drawing rigorous conclusions** for what autonomous decision-making policy your company should progress to the next (more expensive) stage of real-world development and testing. This means that:

- **You are responsible** for defining the requirements and performance metrics that together define the competing objectives of successful autonomous driving, i.e., the rewards.
- **You are responsible** for carrying out policy synthesis of multiple methods.
- **You are responsible** for the design of verification tests, for example:
  - Choosing how many simulations runs to perform for verifying policy performance.
  - Selecting the range of parameters values that you simulate for the purpose of verifying the robustness of the policies in-the-face-of uncertainty.

- **You are responsible** for performing and presenting a meaningful comparison between the policies that you synthesize.

This task brief is **high-level** and **open-ended** (as opposed to being highly detailed and specific about every step to perform). Hence, there are **many plausible solution paths** that are worthy of full marks for this task.

TASK DESCRIPTION

TASK SUMMARY:
Your task, as a team, is to design, synthesize, and compare multiple policies that enable a **full-sized** passenger vehicle to **drive autonomously** along a **high-speed** road with **tight corners**.
In order to achieve this goal, you need to:

(A) **Create** one or more roads in simulation that sufficiently capture the difficulties of driving on high-speed roads with tight corners.

(B) **Specify, explain, and briefly justify** the performance metrics that you use for assessing the quality of a policy for autonomous driving on such roads.

(C) **Design and synthesize** one advanced policy per team member. Advanced policies include:
  - MPC, which includes: Linear MPC, RTI-MPC, non-linear MPC.
  - RL using deep neural networks, which includes: DDPG, SAC, VPG, TRPO, PPO, TD3 (see Spinning Up algorithm docs for a description of each)

(D) **Perform a comparison** of the policies from Part (C) using the performance metrics from Part (B) for the road(s) from Part (A).

(E) **Reflect** on your experiences in this task and/or in the subject as a whole (**or lack of** such experiences), with-respect-to **specific topics** covered by the industry guest lectures; 2 topics per team member, 200-600 words per topic.

To support you with the simulation, we provide an Autonomous Driving Gymnasium class that adheres to the Farama Gymnasium framework. Details of the Autonomous Driving Gymnasium class are provided in a separate box at the end of this task brief document. **Note** that this code is similar but different to what was provided for the Sim-2-Real task.

TASK GUIDANCE:
The following are suggestions for various parts of this task. As described in the purpose statement above, this is an **open-ended task**, hence, there are many plausible solution paths, hence, you should interpret the following suggestions as guidance rather than as being fully prescriptive.

(A) **For creating roads**, you can consider:
  - **Be pragmatic!** For example, quickly construct a road that is about 10km long, has a few long straight sections (0.5-1km) and has many corners with radii in the range 40-200m. Then revisit the road design once you have things "working".
    - An alternative pragmatic approach is to randomly generate the road, e.g., by randomly sampling the lengths and curvatures from a reasonable range.
  - You do **NOT** need to re-create a real-life road in simulation.
  - You **do need to** create a road in simulation that has lengths and curvatures that are representative of a high-speed road with tight corners (for example, a road with a speed limit of 80-120 km/h and whose corners recommend to reduce speed by 30km/h or more below the speed limit).

- As **inspiration** for manually designing a road, you can think of your "favourite" road that has a speed limit of around 80 km/h or 120 km/h and which has many corners that recommend a significant reduction in speed.
  - If you are unfamiliar with such roads, one example is the section of the Great Ocean Road between "Memorial Arch at Eastern View" and "The Springs" (i.e., between (https://maps.app.goo.gl/ypLuSkj85KDRhUt68 and https://maps.app.goo.gl/AkZpgDWqTyaxTh986).

(B) For performance metrics, you can consider:

- **Be pragmatic!** For example, quickly define a few metrics to get your team moving with policy synthesis, then revisit these metrics once you have things "working". For example: a linear penalty for time taken to complete the road; a quadratic penalty for exceeding the speed limit; and a quadratic penalty for deviating from the road line by more than 0.5 meters.
- Some **metrics** to consider are:
  - Time to complete a road (or some part of the road).
  - Going over the speed limit (by how much, and for how long).
  - Acceleration forces experienced by the passenger, i.e., g-force. This metric could be different for linear acceleration, linear braking, and centripetal acceleration during cornering (i.e., "v^2/r")
  - Maximum deviation from the center-line of the lane (i.e., deviation from the road description that you create in Part (A)).
  - Maximum slip angle of the wheels.
  - Steering (or vehicle) oscillations that are high frequency and "noticeable" in magnitude.
  - The "amount" of information available to the policy.
  - Other metrics that you propose.
- For **quantifying robustness** (i.e., transfer-ability), you could define a metric for the change-of-performance when specific parameter values change, for example:
  - changes of vehicle mass;
  - changes of road surface condition (i.e., dry or wet);
  - changes of steering alignment offset;
  - changes of the amount of information available to the policy about the upcoming road.

(C) For policy design and synthesis, you can consider:

- This part is expected to be the **bulk** of the work-load for this assignment.
- As described in the Autonomous Driving Gymnasium additional info at the end of this task brief document, you are able to **adjust the observations** that are provided to the policy. Hence this is an additional parameter that you can adjust and it is expected to change the policy performance.
- When presenting this part in the report it is **important that you are clear about**:
  - The formulation and algorithm of the policy and policy synthesis that you implement.
  - The observations available to the policy.
  - The parameters of the policy and policy synthesis algorithms that you need to select, and a discussion of how you adjusted and selected these parameters.
- You are allowed to use **external libraries and toolboxes** for assisting your policy synthesis and implementation. For completeness and clarity, the use of such libraries and/or toolboxes should be briefly and explicitly stated in the report.
- You are likely to encounter many **"bugs"** in getting a policy up and running, which is likely to be a combination of code bugs and algorithm bugs and parameter selection bugs.
- Design and synthesis of **"less advanced" policies** (instead of the "advanced" policies listed above) will also be awarded marks, albeit less than the full marks available for this part. Examples of "less advanced" policies are: pure pursuit controller; Stanley controller; ruled-based switching between PID controllers.

(D) For comparing policy performance, you can consider:

- This analysis should consider that there are **multiple performance metrics**:
  - hence, the “true” autonomous driving task is multi-objective;
  - hence, there is no clear way to use a single-objective function for assessing whether one policy is “better” than another;
  - hence, you need to discuss the trade-offs between competing objectives that you observe in your results;
  - hence, the concepts and terminology from Pareto efficiency / Pareto optimality are relevant, most notably Pareto dominance and the Pareto front.
- The comparison should include **discussion** about **why** one policy performs better than another for a particular performance metric, and **why** this is to be expected (or not).

(E) For reflection on industry guest lectures, refer to the marking guide below.

MARKING GUIDE AND RUBRIC

For the progress update, 4% of the subject grade is awarded if your team:

- Shows and describes an initial road that your team is using.
- Shows and describes an initial set of performance metrics that your team is using.
- Describes the current status of the policy and policy synthesis method that each team member is working on.

Note: You do **NOT** need to prepare slides for the discussion part, though you may if you wish. You **MUST** be prepared to succinctly describe your current status.

For the written report, 21% of the subject grade is awarded as per the following:

- **For Part (A), 1% is awarded** for a clear and complete description.
- **For Part (B), 2% is awarded** for a clear and complete description, with sensible justification statements.
- **For Part (C), 11% is awarded** according to the rubric below.
- **For Part (D), 4% is awarded** according to the rubric below.
- **For Part (E), 3% is awarded** for reflections that make sensible connections between: experiences within this subject (or lack of such experiences within this subject); and **specific topics** from the industry guest lectures. **Specific topics** means that you are expected to reference specific slides from one or more industry guest lectures, or give specific quotes from one or more industry guest lectures.

Parts (C) and (D) of your team report are assessed as per the following rubric:

Criteria	Weighting	Aspects assessed by each criterion
Rigorous approach	**85%**	• Analytic approach and structure for all aspects of the policy design, policy synthesis, policy robustness evaluation, and policy comparison. • Supporting evidence in the form of calculations and/or simulations. • Discussion and justification throughout that builds upon and links specifications, theory, algorithms, verification, and comparisons. • Limitations identified and discussed throughout.
Presentation style and clarity	**15%**	• Clarity of writing, with a focus on being intelligible and unambiguous. • Use of tables, figures, and appendices that support of the main text, and vice versa. • Document structure, sectioning, layout, appearance, and formatting.

CRITERIA GUIDANCE

- The "task description" box above provides guidance that you are **strongly recommended** to read in full.
- **It is strongly recommended to AVOID a chronological flow within the sections of your report.**
  - An example of a chronological flow: "**We first** defined some simple performance metrics, (...more details). **Then**, after some policy synthesis we realised the performance metrics were not adequate (... more details). So we changed the performance metrics **and then** we re-ran the analysis again (... more details). **Finally**, the performance metrics we used were (... more details)."
  - This example of a chronological flow is **NOT** desirable and is **NOT** typical of science and engineering reports.
  - Science and engineering methods do indeed follow a chronological flow when you carry them out. **However**, recounting this chronological flow in a report often causes the reader to misinterpret the key details of the report.
  - We **strongly recommend a results-focused flow** for the sections of your report. This is a flow where you:
    * **First** describe the final design (or equivalent) that you reached, thus making the key details clear to the reader.
    * **Then** discuss the reasons and justifications for that final design, which usually draws upon the steps you worked through in your chronological methodology.
  - An example of a results-focused flow (for the same topic as the example above): "The performance metrics we use are (... more details from what came at the end of the chronological flow example). The key reasons for selecting these performance metrics is that we tested (... more details of the simple performance metrics and discussion of why they are insufficient)."
  - It may feel like that this "just" writing the chronological flow backwards in time. **However**, you should find that the results-focused flow has an impact on your phrasing, your presentation style, and your thinking.
  - This does **NOT** mean that you should discard the chronological flow all together. The chronological flow can be the basis of your notes for planning the report, and then you convert this to a results-focused flow when you write the report that is submitted.

ADDITIONAL INFO - AUTONOMOUS DRIVING GYMNASIUM

- The Gymnasium environment that we provide for simulating driving of a full-scale passenger vehicle on a high-speed road with tight corners can be downloaded / cloned from here:
  `https://gitlab.unimelb.edu.au/asclinic/ai4r-gym`
- Instructions for installing the Gymnasium environment are given in the `README.md` file.
- The "Autonomous Driving Env" class adheres to the Farama Gymnasium framework, which means that it should be compatible with existing libraries and toolboxes for training RL policies.
- The Road class is essentially identical to what was provided for the Sim-2-Real task.
- The "Bicycle Model Dynamic" class implements the dynamic bicycle model presented in Lecture 7, and hence includes a model for tyre slip. Further details below.
- The examples scripts provide you with a starting point for how to use the Gymnasium environment.

Further details for the "Bicycle Model Dynamic" class

- This class implements a **kinematic** bicycle model at low speeds, a **dynamic** bicycle model at high speeds, and **transitions** between the two model over the speed range $v_x \in [3, 5]$ [meters/second] (i.e., the range $v_x \in [10.8, 18]$ [km/h]), implemented as per Section 5.1 of the AMZ Driverless paper.
- The task of autonomous driving on high-speed roads with tight corners is most interesting when the dynamic bicycle model is "active". Hence it is **strongly recommended** that you use an initial condition speed of $v_x > 5$ [meters/second] and always keep the vehicle speed above this.
  - For example, create the road with a long straight at the start, use an initial condition of $v = 100$ [km/h] (i.e., the speed limit), and create the road so that the sharpest corners can be drive with a speed of greater than 40 [km/h].
- The **slip angles** of the front and rear tyres, denoted α_f and α_r respectively, are the two most important quantities for determining the behaviour of the dynamic bicycle model. In the bicycle model class, the slip angles are computed as:

$$\alpha_f = -\arctan\left(\frac{v_y + l_f\,\omega}{v_x}\right) + \delta\,, \qquad \alpha_r = -\arctan\left(\frac{v_y - l_r\,\omega}{v_x}\right), \tag{1}$$

  where all notation is the same as defined in Lecture 7 (noting that $\omega = \dot{\theta}$).
- The **lateral forces** at the front and rear tyre, denoted $F_{f,y}$ and $F_{r,y}$ respectively, are computed using the following Pacejka tyre Formula:

$$F_{\{f,r\},y} = F_{\{f,r\},z}\,D_p\,\sin\left(C_p\,\arctan\left(B_p\,\alpha_{\{f,r\}} - E_p\,\left(B_p\,\alpha_{\{f,r\}} - \arctan\left(B_p\,\alpha_{\{f,r\}}\right)\right)\right)\right)$$

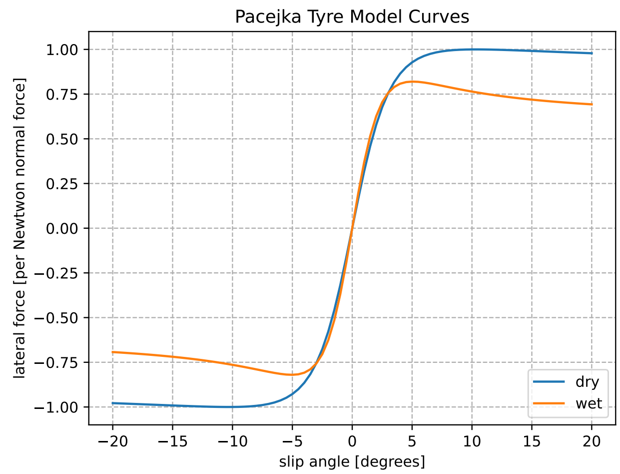
  where $F_{f,z}$ and $F_{r,z}$ are the normal force on the front and rear tyre respectively, computed for simplicity as the split of the total mass relative to the l_f and l_r distances, i.e.:

$$F_{f,z} = mg\left(\frac{l_r}{l_f + l_r}\right), \qquad F_{r,z} = mg\left(\frac{l_f}{l_f + l_r}\right).$$

- The coefficient values for the Pacejka Tyre Formula (i.e., D_p, C_p, B_p, E_p) are taken from this source as:

Symbol	Name	Dry Tarmac	Wet Tarmac
D_p	Peak	1.0	0.82
C_p	Shape	1.9	2.3
B_p	Stiffness	10.0	12.0
E_p	Curvature	0.97	1.0

Hence, for one unit of normal force on the tyre (i.e., $F_{\{f,r\},z} = 1$), the following figure is a plot of the Pacejka Tyre Formula.



- The most important feature to interpret from the figure of the Pacejka Tyre Formula is that for slip angle magnitude greater than approximately 5 degrees, the lateral force curve flattens, hence, significant slipping (e.g., out-of-control rotation) of the dynamic bicycle model is expected for such slip angles.

> **NOTE:** the equations above are already implemented in the "Bicycle Model Dynamic" class and hence there is nothing that you need to change about that class. The important information is that:
> – You can use equation (1) for computing slip angles α_f and α_r.
> – You can use the the figure above of the Pacejka Tyre Formula to decide a limit on the slip angles that you consider to be unacceptable for safe driving.

Further details for the "Autonomous Driving Env" class

- You should review the doc strings of the `_get_observation` and `_get_info` functions (e.g., by searching for those function here) so that you are fully aware of what information is available to the policy.
- As mentioned above, the examples scripts provide you with some guidance on how to use this class.
- To use existing RL libraries, you likely need to adjust the observation that this class implements. In order to do this, you will need to edit the `_get_observation` function or create an observation wrapper.