AI Project: Evaluating Artificial Intelligence as an Engineering Design Tool using a Transport Phenomena Design Problem

CBE 30355 Transport Phenomena I
D. T. Leighton
Department of Chemical and Biomolecular Engineering
University of Notre Dame

August 16, 2026

1 Introduction

Artificial intelligence is rapidly becoming part of engineering practice. Modern large language models (LLMs) are capable of solving many engineering problems, writing computer programs, producing figures, and generating complete technical reports. These capabilities make AI a potentially valuable engineering assistant. At the same time, AI systems are well known for producing plausible-looking but incorrect results. Unlike a conventional numerical solver, an AI system does not guarantee that every equation, numerical value, or engineering assumption is physically correct.

Consequently, the most important engineering skill is not learning how to obtain an answer from an AI system. Rather, it is learning how to evaluate that answer critically, identify mistakes, and guide the AI toward an improved solution. This project is designed to provide experience with precisely that process.

Throughout this seven-week project your group will use several different AI systems to solve the same engineering design problem. Rather than attempting to determine which AI is "best," you will investigate the strengths and weaknesses of each system, compare their zero-shot responses, iteratively improve one design through a sequence of prompts, and finally evaluate how the design evolved during the refinement process.

The engineering problem selected for this project concerns the design of a simple temperature-measuring device based upon the principles of Galileo's thermometer. Although the underlying physics involves concepts from fluid mechanics, heat transfer, thermodynamics, and hydrostatics, the problem has no known textbook solution. Consequently, each AI system must develop its own design strategy rather than simply reproducing an existing solution.

The goal of this project is therefore *not* simply to obtain a working thermometer design. Instead, the goal is to develop experience using AI as an engineering design tool while simultaneously developing the critical judgment necessary to recognize both the capabilities and limitations of current AI systems.

(a) Float position at 18.0^{o}C (b) Float position at 27.8^{o}C

Figure 1: Continuous chain-loaded thermoscope at two different operating temperatures, demonstrating float displacement along the graduated cylinder.

2 Project Goals

After completing this project, students should be able to

1. Use modern AI systems as engineering assistants rather than as sources of unquestioned answers.
2. Develop effective prompts that communicate technical information clearly and efficiently.
3. Produce professional engineering reports using LaTeX.
4. Critically evaluate AI-generated analyses for physical, mathematical, and engineering correctness.
5. Identify incorrect assumptions, unsupported conclusions, and mathematical errors contained within AI-generated work.
6. Improve an engineering design through an iterative dialogue with an AI system.
7. Compare multiple AI systems when presented with the same engineering design problem.
8. Understand that engineering design is an iterative process requiring both analytical calculations and engineering judgment.

It is expected that by the conclusion of this project students will have developed a much better understanding of both the usefulness and the limitations of modern AI systems. In particular, students should recognize that AI-generated solutions must always be subjected to the same careful engineering scrutiny as work produced by another engineer.

3 General Requirements

This project extends over seven weeks. During each week a specific task will be completed. Except where otherwise noted, all work should be performed using the free versions of the AI systems.

Three different large language models will be investigated:

- Google Gemini (Thinking and Pro modes)
- ChatGPT
- Anthropic Claude

Each AI system will receive exactly the same initial engineering design prompt. Their responses will then be compared, analyzed, and critiqued.

Because the AI models will yield different results (and this would even occur for the same AI every time you do it) there will be different strengths and weaknesses in each. The objective of the refinement phase is to combine the best parts of each of them with human insight and obtain a significantly improved final design.

Finally, the completed design will be compared against each of the three original zero-shot solutions in order to understand how the design evolved and which engineering issues proved to be the most significant.

4 Week 1 – Learning to Communicate with an AI

Before beginning the engineering design problem, students should become comfortable using LaTeX to communicate technical information.

Although modern AI systems can usually interpret ordinary text, complex equations, mathematical notation, figures, and engineering reports are most naturally communicated using LaTeX. Throughout this course all major interactions with the AI systems will therefore be performed using LaTeX whenever practical.

Each student should install a LaTeX distribution on his or her computer. Suitable distributions include TeX Live, MacTeX, and MiKTeX. A convenient editor such as TeXstudio, TeXShop (what I use on my Mac), Overleaf, or Visual Studio Code with a LaTeX extension may also be used. AI is a *great* LaTeX coach. Have it walk you through the various installation options and to deal with the inevitable "LaTeX errors". The latter is done by simply pasting in the error message from the compiler console and asking the AI to tell you how to fix it. I have never had to read the LaTeX manual or help guides!

To gain familiarity with using LaTeX in conjunction with AI, students will complete a short introductory exercise based upon the "Miranda Quote" prompt. The complete prompt is provided in Appendix A.

Upload the prompt into Gemini (Thinking mode) and compile the resulting script in your LaTex editor (and read it). After you have done so, open up a new chat window in Claude, upload the .tex file which you just generated, and feed in the following prompt:

> *the result of the miranda assignment is uploaded. To what extent is the "how this was generated" response a hallucination?*

Ask Claude to provide the response it generates as a latex section, and append it to your Gemini Miranda file (e.g., a prompt such as that below):

> *please provide this response as a latex script for insertion into the tex file.*

Add a final section to the tex file with your (brief) comments on the outputs of the different models (including references, confabulation, and hallucination).

The objectives of this preliminary exercise are

1. to verify that the LaTeX installation is working correctly,
2. to gain experience copying and compiling LaTeX documents generated by an AI system,
3. to observe how accurately (or not) the AI deals with references and citations, and
4. to become familiar with the conversational style that will be used throughout the remainder of the project.
5. to grasp some of the issues with AI usage in writing and analysis.

4.1 Deliverables

The report for Week 1 is simply the composite PDF generated from the LaTeX produced by Gemini and Claude, together with your brief comments.

5 Week 2 – Design Proposal A: Zero-Shot Engineering Design Using Google Gemini

The first engineering design exercise will be performed using **Google Gemini in Thinking mode**. Gemini was selected as the first AI system because it generally has the largest available context window among the freely available models and is capable of producing an entire engineering report in a single conversation.

The purpose of this week's assignment is *not* to obtain a perfect engineering design. Rather, the objective is to observe what an AI system can produce when presented with a sufficiently detailed design problem without any additional guidance or corrective feedback. This is commonly referred to as a *zero-shot prompt*.

The Galileo Thermometer design prompt is provided in Appendix B. The prompt should be copied exactly as written. No additional instructions, corrections, or follow-up questions (beyond identification of LaTeX errors which must be corrected to allow compilation) should be provided during this stage of the project.

The AI should therefore receive only the original prompt and be allowed to generate its complete solution without assistance. Note, however, that you may need to prompt the AI to actually give you the full LaTeX. Sometimes it partially interprets it in error. Such a prompt might be:

> *This latex was partially interpreted. Please provide as a latex script only.*

5.1 Engineering Review

Once the report has been generated, you now need to analyze it! You will do this two ways. First, open a new window in Gemini (so that you start fresh) and ask it to do it for you. Upload the

`.tex` file of the report and then give it the following prompt:

> Please analyze the report just uploaded, returning the analysis as a LaTeX script. The analysis should consist of two parts:
>
> 1. A brief summary (approximately one page) describing the overall quality of the response.
> 2. A critical engineering evaluation identifying both the strengths and weaknesses of the design.
>
> In your engineering review, the following questions should be considered.
>
> 1. Is the overall design physically reasonable?
> 2. Are all assumptions clearly stated?
> 3. Are the governing equations appropriate?
> 4. Are the numerical calculations internally consistent?
> 5. Does the proposed thermometer actually appear capable of being constructed?
> 6. Are all dimensions physically realistic?
> 7. Are all figures dimensionally consistent with the calculations?
> 8. Does the AI distinguish between measured quantities, assumed quantities, and calculated quantities?
> 9. Does the report reach conclusions that are justified by the analysis?
>
> **Common Issues to Consider**
>
> The following checklist is intended only as a guide. Not every report will contain every issue, and additional problems not listed here may also be present.
>
> - Failure to distinguish between mass and weight.
> - Incorrect application of Archimedes' principle.
> - Incorrect use of the ideal gas law.
> - Ignoring the thermal expansion of water.
> - Ignoring atmospheric pressure effects.
> - Ignoring hydrostatic pressure changes as the float moves.
> - Incorrect treatment of the submerged weight of the chain.
> - Failure to account for the U-shaped chain kinematics.
> - Failure to satisfy static force equilibrium.
> - Ignoring phase equilibrium, relative humidity, or water vapor pressure inside the balloon.
> - Unrealistic geometric dimensions.
> - Predicted motions larger than the available height of the graduated cylinder.

- Response-time estimates that are unsupported or physically unreasonable.
- Figures that do not agree with the numerical design.
- References that are incomplete, broken, or fabricated.
- Equations that are dimensionally inconsistent.
- Numerical values presented with no derivation.

5.2 Your Analysis

Carefully review both the report, the list of things to think about, and Gemini's analysis. Then append to the review *your* analysis of both: what errors did the report make, what errors did Gemini catch, what errors did it miss. This is actually quite revealing. It will be different every time it is done, but I have found that while an AI can often find mistakes it will both miss some and hallucinate others. That is why you can't blindly rely on AI output!

5.3 Deliverables

The report for Week 2 shall contain the following items.

1. The complete AI-generated report.
2. The Gemini analysis of the report.
3. Your (brief) commentary on both.

You should *not* attempt to correct the design during this week. The objective is simply to document the quality of the initial proposal.

6 Week 3 – Design Proposal B: Zero-Shot Engineering Design Using ChatGPT

The second zero-shot design proposal will be generated using **ChatGPT**. The identical engineering prompt given to Gemini should again be used without modification.

Unlike Gemini, the free version of ChatGPT frequently cannot produce a complete engineering report of this size in a single response. Consequently, you may need to generate the report in several successive sections.

A typical interaction is

1. Request the complete report.
2. Allow ChatGPT to generate as much of the report as possible.
3. Request the next section.
4. Continue this process until the report has been completed.
5. Combine the individual sections into a single LaTeX document.
6. Compile the document and correct any transcription errors that may have occurred while assembling the individual sections.

This limitation is primarily a consequence of response-length restrictions rather than engineering ability. The resulting report should therefore be judged on its technical content rather than on the fact that it was produced in multiple pieces.

6.1 Deliverables

The deliverables are identical to those of Week 2.

1. The complete AI-generated report.
2. The Gemini analysis of the report.
3. Your (brief) commentary on both.

7 Week 4 – Design Proposal C: Zero-Shot Engineering Design Using Claude

The third zero-shot design proposal will be generated using **Anthropic Claude**. As in the previous two weeks, the Galileo thermometer design prompt provided in Appendix B should be copied exactly as written. No follow-up prompts, corrections, or suggestions should be provided until the initial design proposal has been completed.

If you are using the free version of Claude you should be aware that the conversation will likely exceed the usage limits before the report is finished. This is expected and is not an indication that the AI has failed.

If the conversation times out, simply wait until the usage limit resets (currently approximately five hours for the free tier, although this is pretty random), reopen the same conversation, and ask Claude to continue generating the report from the point at which it stopped. Continue this process until the report is complete, although generally it only takes one extra try.

As with ChatGPT, the final report should be assembled into a single LaTeX document if necessary before it is compiled.

7.1 Deliverables

The deliverables for Week 4 are identical to those of Weeks 2 and 3.

1. The complete AI-generated report.
2. The Gemini analysis of the report.
3. Your (brief) commentary on both.

8 Week 5 – Iterative Engineering Refinement

Engineering design is almost never completed in a single attempt. Instead, an initial concept is repeatedly analyzed, modified, and improved until it satisfies the design objectives.

The purpose of Week 5 is to investigate whether an AI system can participate effectively in this iterative engineering process and to generate a correct, validated design report.

The first part of the iteration is to get a good starting point. There are many ways of doing this, and it depends on what the zero-shot design proposals produced. First, you need to identify what was good and bad about each. If two were "pretty good" then feed these into Gemini and ask it to "meld them" into a single document, describing in your prompt what elements of each to preserve and to correct (e.g., TikZ diagram from one, analysis from the other, etc.). This should yield a single latex script which gets you closer to the desired result.

The next (and vital) step is to iterate the report section by section, pointing out where things could be improved (and references corrected). This can be done in either Gemini or Claude (if your free version doesn't time out on you). Make sure that the sources of all the material properties are provided (Claude is usually better at this) and check all the references! At the end have it generate an abstract (if it isn't already there), making sure it is congruent with the final report, as well as a conclusions section.

Once it is all done, append a section describing the method you used to get the final design report.

8.1 Engineering Refinement

In improving the report you need to treat the AI as though it were a junior engineer working with you on the project. Rather than asking the AI to "improve the report," identify specific engineering issues and request that they be addressed one at a time.

Examples include

- getting the physics correct,
- correctly calculate the hydrostatic amplification factor,
- correctly handling the effects of water vapor,
- identifying the tradeoffs between balloon volume and chain weight on sensitivity and range,
- showing how atmospheric-pressure sensitivity is related to temperature sensitivity,
- improving the transient thermal model,
- correcting numerical errors,
- producing a dimensionally consistent TikZ figure,
- improving the bibliography and references.

Each refinement should have a clear engineering objective.

Whenever possible, verify that the proposed correction actually improves the design before moving on to the next refinement.

Continue the iterative process until you believe that no significant engineering deficiencies remain.

8.2 Deliverables

At the conclusion of the refinement process, submit your final design report!

This report will be referred to as the *Final Design*. Don't forget to include your (brief) summary of how you obtained it from the zero shot AI designs!

9 Week 6 – Comparative Engineering Evaluation

The final week of analysis examines how the engineering design changed during the refinement process and examines the strengths and weaknesses of using an AI for comparative evaluations.

Take your three zero-shot prompt reports (tex files!) and duplicate them, labeling them report A, B, and C. Do the same for your final design report (stripping off your comments at the end) and label it report D. To perform the engineering evaluation, open a *new* session in Google Gemini so that the evaluation is conducted independently of prior conversation history and upload the four reports. Note that the AI may not keep the file names straight - instead it will likely number them based on the order of upload, so you will need to tell it which one is which afterwards. Note that the total amount of text in these four reports is stretching the capacity of the AI to keep it straight - better results could be obtained if you strip out everything except the analysis and design sections. This would certainly be required for longer reports.

Stage the Evaluation: Rather than requesting a full document review in a single prompt, evaluate the proposals through sequential, targeted turns using structured outputs.

Instruct Gemini to evaluate the four reports as though they were anonymous engineering design proposals, prioritizing mathematical and physical analysis over writing style.

9.1 Comparison of Physical Models

In your initial evaluation turn, direct the AI to analyze the physical foundations of each proposal.

Prompt the AI to construct a comparative table addressing:

- Was Archimedes' principle applied correctly using submerged fluid density differences?
- Was the force balance complete, accounting for the buoyant submerged weight of the chain rather than its dry weight?
- Was the trapped air volume (V_{air}) modeled dynamically as a function of temperature and local pressure?
- Were atmospheric pressure variations (P_{atm}) and depth-dependent hydrostatic pressure ($\rho_w g h$) integrated into the balloon's equilibrium volume?
- Was the transient response modeled using appropriate conductive and convective heat transfer parameters?
- Were the governing equations dimensionally consistent across all terms?

9.2 Engineering Feasibility

In the subsequent prompt turn, evaluate the practical mechanical operation and physical feasibility of each design.

Prompt the AI to generate a feasibility matrix assessing:

- **Physical Dimensions:** Compatibility between balloon expansion volume, sinker mass, and cylinder dimensions.
- **Mechanical Stability:** Adequacy of the chain's weight gradient to prevent hydrostatic tipping or rapid uncontrolled surfacing.

- **Sensitivity vs. Range:** Predicted vertical height sensitivity ($\mathrm{d}x/\mathrm{d}T$) and whether the float remains within the physical bounds of the cylinder.
- **Response Time:** Physical plausibility of the calculated thermal time constant (τ).
- **Completeness:** Overall practical viability of the proposal as an actionable laboratory specification.

9.3 Evolution of the Design

Finally, ask Gemini to compare Proposal D (the Final Design) directly against each of the three original zero-shot proposals (A, B, and C).

The evaluation should explicitly document:

1. The principal engineering deficiencies and physical hallucinations present in each original proposal.
2. The specific mathematical and physical corrections introduced during the refinement process.
3. The engineering justification for each significant modification.
4. Which specific corrections had the greatest impact on overall system performance.
5. Whether the Final Design represents a genuine engineering improvement grounded in transport phenomena principles or merely a more detailed document.

Note that while this process may *sound* complicated, you can simply feed the text of this section into Gemini and ask it to generate the required sequence of prompts for you! Instruct the AI to return each section of its comparative analysis formatted as a LaTeX snippet so that you can easily copy, paste, and compile the final deliverable document.

Once you have compiled these comparative sections (or had the AI do it for you), submit the complete draft back to Gemini and instruct it to generate an executive abstract for the beginning of the report and a synthesizing conclusion for the end.

You should recognize that engineering progress is rarely the result of a single brilliant idea. Instead, it typically emerges through rigorous analysis, critical evaluation, and systematic refinement of an initial concept.

9.4 Your Assessment of the Comparison

Once Gemini has produced its comparative tables and its account of how the design evolved, do not simply accept this analysis at face value. You have something Gemini does not: direct memory of actually performing the Week 5 refinement, including which corrections were straightforward, which were contentious, and which the AI initially resisted or got wrong before you intervened.

Read through Gemini's comparison critically and address the following:

- Does Gemini's account of *why* the Final Design improved on Proposals A, B, and C match your own recollection of the refinement process? Note any places where its explanation seems retrospectively tidied up, oversimplified, or disconnected from what actually happened during Week 5.

- Did Gemini overstate or understate the severity of any deficiency in Proposals A, B, or C? In particular, check any claim you cannot independently verify against the original reports.
- Did the comparative tables miss an engineering issue that you know was present in one of the proposals, or attribute a fix to the wrong cause?
- Are there any claims in the "Evolution of the Design" discussion that are simply assertions, unsupported by the equations or numbers in the reports themselves?

Append a brief (roughly half-page) commentary addressing these questions at the end of the analysis. In particular look at whether the AI's judgment shifts suspiciously once it learns which proposal is the "final" one. Sycophantic context-anchoring is a known failure mode — check whether its stated reasoning actually changed or just its conclusion. As in previous weeks, this is not a formality: the objective is to demonstrate that you evaluated the AI's comparative judgment with the same scrutiny you applied to its engineering analysis, rather than allowing the most complex synthesis task in the project to go unchecked.

10 Week 7 – Physical Fabrication & Experimental Validation

The ultimate test of any engineering model is physical reality. In this phase (the fun part) you will construct a balloon thermoscope in the laboratory and compare its real-world performance against your theoretical predictions.

This task may be completed at any point during the semester, and does not need to wait until your Final Design (Week 5) is complete. You are free to build using whatever set of hardware parameters seems reasonable at the time – one of your zero-shot AI designs, an intermediate refinement, or simply your own sensible estimate based on the raw measurements given in Appendix B. What matters is that you accurately record the parameters you actually used (air volume, sinker mass, chain length, etc.), since these – not the numbers originally proposed by any particular AI design – are what your Final Design model will ultimately be compared against. Groups are strongly encouraged to build and collect data early, both to spread out demand on the lab and to leave time for a second attempt if the first prototype turns out to be unstable or otherwise unworkable (see below).

The experimental work will be done in the A68 projects lab, generally available during normal lab hours (Mondays through Thursdays, 2–5pm). Other times are possible by arrangement. Because lab space and hardware are limited, groups should sign up for a time slot in advance rather than waiting until the last available week; a sign-up sheet will be provided. **All fabrication and data collection must be completed before the last lab day of the semester, which is Thursday 12/3**, but there is no other deadline attached to this portion of the task – only the written comparison report (below) has a fixed due date.

10.1 Apparatus & Fabrication

Assemble the thermoscope using the provided laboratory hardware:

- 500 mL Graduated Cylinder
- 9-inch Latex Party Balloons
- Thread & Paperclip

- Metal Chain
- Steel Washers for Counterweight
- Lab Thermometer
- Ruler & Binder Clip

It is actually a little tricky to get a precise volume into the balloon by inflating it. One method that works is to simply capture air in the deflated balloon and tie it off tightly with a bit of thread. Tie on a paperclip as well, and hang the desired mass of weights and washers onto the clip, adding one extra small washer that would account for the (later) addition of the chain. Then drop the balloon into a 2 liter beaker filled with water at the desired temperature. If it floats, as is likely, squeeze a bit of air out past the knotted thread and try it again. Eventually you will reach a point where it sinks with the extra washer and floats without it - and you are then ready to go! You can also use the weight of all of your washers (and paperclip - even small pieces of a paperclip) at neutral buoyancy to get a very good measurement of the actual volume of air in the balloon. Attach the chain to the paperclip, removing the extra washer, and insert into the graduated cylinder.

10.2 Experimental Testing Protocol

1. **System Zeroing & Equilibrium:** Fill the cylinder with water from the beaker that you balanced the system with so that it is at the correct temperature. Allow the balloon float to come to complete static equilibrium, paying in the chain until it is properly centered. Note that this requires you to have the correct amount of air and counterweight! Once you are satisfied, pin the other end of the chain to the top of the cylinder with a binder clip. Record the vertical position z_0 (read from the graduated cylinder scale, but don't forget to use a ruler to get the conversion!) and the ambient water temperature T_0.
2. **Thermal Sensitivity Measurement ($\mathrm{d}z/\mathrm{d}T$):** By far the most reproducible way of measuring the thermal sensitivity is to carefully pour out the water and replace it with water of a different temperature, making sure that the final water level is the same (which would have been recorded or marked on the cylinder). Each time you do this make sure that there are no kinks in the chain, as that would affect the weight/length. Measure the temperature of the water at the location of the float using a thermocouple probe as best possible (not always easy to do, but if it is properly mixed it shouldn't be stratified). Record float height z as a function of temperature once the float reaches static equilibrium. Do this for 5 or more different temperatures, randomly obtained by adding hot water or ice water to the water you have poured out (and then back in).
3. **Parameter Check:** The behavior of the system is critically dependent on the effective chain weight. Thus, this requires extra investigation. The design assumes a particular density for the chain - but you know all about the risk of assumptions! Measure the effective density using the same Archimedes Law approach used in the class demonstration. Take a pencil and lay some thread along it. Wrap a coil of the (dry) chain around the pencil and thread, then tie up the thread and pull out the pencil! You can now dangle it so that it is completely submerged in water on a scale, then let it rest on the bottom. The ratio of the weights is the density you need in your calculations. Do the same for your collection of weights and washers.
4. **Parameter Sensitivity:** A simple way to vary the sensitivity of the system to chain weight

is to simply double it! Attach a second chain to the paperclip (you will likely need to remove a small washer) and measure the change in height for at least three different water temperatures.

10.3 Model vs. Data Comparison Report

Once your Week 5 Final Design is complete, use AI to help you prepare a comparison of your measurements to the expectations of the final design. To start, have the AI recalculate the behavior for the actual parameters used in your experiment. Check the calculations to make sure it hasn't made any errors! Then feed in all of your experimental observations and have it do the analysis of the experimental results, again checking to make sure it hasn't hallucinated things! Your report should include the following:

- A theoretical description of the design equations (likely extracted from the design report).
- A description of your experimental procedure.
- Measurements of all key parameters and ranges of observations.
- A plot of the predicted $z(T)$ vs. measured $z(T)$. Was the response linear as expected?
- Contrast the measured and predicted slopes. Are they significantly different?
- If the measured sensitivity is much higher than expected, did the model leave out the effect of humidity?
- Did the doubling of the chain weight by adding a second chain behave as expected?
- Discuss possible sources of discrepancy and how they might be evaluated.
- An abstract and conclusion, with a (validated) list of references.

This PDF is the Week 7 deliverable, due at the date specified in the course schedule even if your data collection was completed earlier in the semester.

11 Submission Requirements

At the conclusion of each week the group should submit a PDF with the requested materials (e.g., week 2 would be the Gemini report with analysis, week 3 would be the ChatGPT report with analysis, etc.). The final design (week 5) and the AI comparison (week 6) and experimental results and comparison (week 7) should also be returned as a PDF.

Students are encouraged to retain the complete conversation histories used during the project. Although these conversations need not be submitted unless requested, they provide a useful record of the design process and may assist in understanding how the final design evolved.

12 Concluding Remarks

Large language models are becoming increasingly capable engineering assistants. They can derive equations, generate computer programs, construct engineering figures, summarize technical literature, and produce complete engineering reports.

Nevertheless, current AI systems remain imperfect. They may make incorrect assumptions, overlook important physical effects, fabricate references, produce internally inconsistent calculations, or reach unsupported engineering conclusions while presenting their work with a high degree of confidence.

For these reasons, engineers should view AI-generated work in much the same way that they would evaluate the preliminary design prepared by a junior colleague. Every assumption should be questioned, every calculation should be examined, and every important conclusion should be supported by sound physical reasoning.

A successful engineer will not necessarily be the one who uses AI most frequently. Instead, it will be the engineer who best understands when the AI is correct, when it is wrong, and how to guide it toward a better solution. That is the skill that this assignment is intended to help develop.

A Miranda Quote Exercise

This appendix contains the introductory exercise used during Week 1.

The complete prompt should be copied exactly as provided below.

> In the Tempest, Miranda exclaims "O Brave New World that has such people in't". Return a 400 word latex script reflecting on this quote and how it applies to AI, both good and bad. After a `\newpage` command, return a second 400 word essay on how the original reflection was generated. At the end, include four validated references (including in-line hyperlinks) for further investigation of the way in which the reflection and essay are generated. The references must be freely accessible on-line.

B Galileo Thermometer Design Prompt

This appendix contains the engineering design prompt used throughout the project.

The identical prompt should be supplied to each AI system during the zero-shot portion of the project. No modifications should be made to the prompt.

> AI Design Project: A variant on Galileo's Thermometer
>
> Many years ago in a local shop I saw a variant on Galileo's Thermometer (you have probably seen those glass tubes with the colored balls that float or sink depending on the temperature). In this case there was only a single disk which was attached to the wall of a narrow rectangular container with a dangling chain. As the temperature of the fluid changed, so did the buoyant force on the disk - resulting in more or less of the chain being pulled upwards, supported by the disk. The chain was in the form of a U shaped loop, with one end attached to the bottom of the disk and one to the wall of the container.
>
> In this project you are to use an AI to design a device based on a similar principle. The float in this case is what we had down in the lab: a deflated 9 inch party balloon partially filled with air. It was measured to have a dry weight of 2.09g, and when filled with water (to measure its maximum deflated volume) weighed 39.27g. To make it nearly neutrally buoyant (and make it stable) steel washers are attached to the bottom, as is one end of a metal chain. A yard of the 5mm curb chain we have down in the lab weighed 24.4g. The balloon is placed in a 500ml graduated cylinder, with the other end of the U chain pinned to the top of the cylinder. By allowing more or less of the chain to be in the cylinder and supported by the float the location of the balloon (determined by neutral buoyancy of the system) can be adjusted.
>
> Using AI, do the following:
>
> 1) For some choice of balloon inflation determine the other system parameters and the sensitivity of the vertical motion of the float to the water temperature.
>
> 2) Based on this calculation determine a reasonable choice for parameters which lead to a "good" height sensitivity.

3) Determine if the performance could be improved if a different chain was used, or if the one we have on hand is close to optimal.

4) Estimate the response time of the thermometer to changes in temperature.

5) Provide a TikZ diagram of the final device to scale.

....

Based on this analysis, produce the following report:

The design report will be returned as a LaTeX document.

1) Begin your document with a historical introduction to the Galileo's Thermometer.

2) In the second section provide an overview of the physics behind this new design and the various factors which affect the performance.

3) In the third section provide the detailed calculations on which the new design is based.

4) In the fourth section conclude with an evaluation of the limitations of this device, in particular considering the relative effect of temperature and air pressure from both atmospheric variation and hydrostatics.

5) Provide a list of appropriate on-line references with verified urls. Journal articles are preferred, but wikipedia is acceptable. Use the `\bibitem` format and include citations in the document.