EEE-Bench: A Comprehensive Multimodal Electrical And Electronics Engineering Benchmark

Supplementary Material

Appendix Table of Content

We provide all additional details for our paper in the following sections.

A. Detailed Related Works

A.1. Large Multimodal Models

Following the rapid development and success of Large Language Models (LLMs) [8, 35, 78, 85], Large Multimodal Models (LMMs) have been constructed by integrating LLMs with large vision encoders [9, 15, 64, 68]. Notably, closed-source versatile LMMs such as GPT-4o [62] and Gemini-Pro 1.5 [76] have achieved exceptional visual reasoning capabilities across various domains, including Visual Question Answering [39] and Mathematics [56]. In response to the need for publicly accessible versatile LMMs, several open-source models such as LLaVA [48] have been developed [13, 23, 40, 41, 47, 48, 77, 89, 97]. Moreover, finetuning [32, 43, 104, 105] aims to improve the performance of LMMs using downstream data. Specialized LMMs finetuned on specific subdomain data [32, 79] have been introduced, demonstrating performance on par with closed-source models in specific subject areas like Mathematics [73, 100, 106] and chart [31, 60]. Recent efforts have also focused on leveraging large models for engineering tasks, such as chip design [5, 6, 11, 34] and optimizing operations in the energy sector [57]. However, the reasoning capabilities of existing LMMs in visual contexts have not yet been thoroughly examined for electrical and electronics engineering (EEE) problems. Therefore, this paper proposes EEE-Bench to comprehensively evaluate the EEE multimodal reasoning abilities of current LMMs.

A.2. Multimodal benchmarks

As Large Language Models (LLMs) and Large Multimodal Models (LMMs) continue to evolve rapidly, the creation of comprehensive and challenging benchmarks to assess their capabilities has become increasingly vital [12, 20]. These benchmarks are indispensable for evaluating the strengths and limitations of these models, thereby facilitating their ongoing development and refinement. To measure the abilities of LLMs, various text-only benchmarks have been introduced [29, 30, 36, 95, 102]. For vision and language benchmarks, early efforts primarily concentrated on general-purpose Visual Question Answering (VQA) tasks [3, 27, 28, 33, 58, 59, 59, 74], which may not sufficiently test the reasoning abilities of today's more advanced LMMs. To provide a multi-faceted evaluation, several benchmarks have been introduced, including specialized ones [42, 53, 56, 67, 80, 84, 99] and versatile benchmarks [14, 39, 44, 51, 71, 82, 92]. For instance, MathVista [56] assesses the mathematical reasoning capabilities of LMMs across various visual contexts. However, the visual context in mathematical problems is much less complex compared with real-world engineering problems that appeared in EEE-Bench. MMMU [92] tackles college-level questions that require intricate, domain-specific knowledge across multiple subjects. Although MMMU also includes some EEE problems, its scope in terms of problem variety and subject coverage is limited, which may not thoroughly evaluate the reasoning capabilities of LMMs on EEE tasks. [22] benchmarks LMMs' understanding of engineering requirement documents, yet they are not directly related to solving engineering problems. To the best of our knowledge, we respectively believe EEE-Bench, which comprises 2860 meticulously curated problems from ten pivotal

EEE subjects, is the first in the direction to offer a robust and comprehensive assessment of LMMs' reasoning abilities in complex visual contexts within EEE. **Comparison to E-split in MMMU:** We've addressed the difference between EEE-Bench and the E-split in MMMU before, here we clarify more numeric distinctions. **(1)** EEE-Bench is more **challenging** than MMMU electronics split (E-split). We tested 7 LMMs on both and compared results: GPT-4o: 46.78/53.33 (EEE-Bench/E-split), GPT-4o-mini: 36.99/46.67, Gemini-1.5-Pro: 41.99/56.67, Gemini-1.5-Flash: 33.95/50.00, Qwen-VL-Max: 25.73/43.33, InternVL2-26B: 26.89/36.67, and Phi3V: 24.90/33.33. The LMMs' performances are much lower on EEE-Bench than E-split. **(2)** EEE-Bench is more **comprehensive** than E-split, featuring 2,860 problems across 10 topics in both *electronics and electrical engineering*, whereas E-split includes only 291 problems from 4 topics in *electronics*. **(3)** Evaluations on EEE-Bench are more **robust** compare to E-split. Most LMMs are assessed using MMMU's val set (as test set is NOT open as of now). The E-split val set contains only 30 problems, making it unreliable for evaluating LMMs' EEE capabilities, as a single incorrect answer results in a 3.33% drop in accuracy, inevitably introducing high variance. EEE-Bench contains 2860 problems in contrast. **(4)** EEE-Bench better **differentiates LMM quality** than E-split on EEE. The performance range for closed LMMs on EEE-Bench is 25.73-46.78, indicating a much larger span than that on E-split: 43.33-56.67.

B. Broader Impact

Limitations and future work. While EEE-Bench rigorously covers 10 critical subdomains of EEE (Appendix F), it does not extend to the broader spectrum of engineering (*e.g.*, civil, mechanical, and chemical engineering). Although we believe our results can provide compelling insights that may generalize to other engineering domains, comprehensive and domain-specific studies are needed to thoroughly and accurately evaluate LMMs' capabilities across the full engineering landscape. We plan to carry out such studies and extend EEE-Bench to broader engineering domains in future works.

Ethics statement. We ensure that EEE-Bench complies with legal and ethical guidelines throughout its construction process, with no violations. We provide fair compensation to all annotators and graduate students involved. We collect problems from free-access public open exams [1] and resources under the CC Licenses [2,3]. We guarantee that EEE-Bench is solely for academic research purposes, and we uphold the strict prohibition of any commercial use. The spurious captions generated in section 4 are only related to problem-solving and do not pose any harm to humans.

Reproducibility statement. We are committed to efficient and reproducible research. Our benchmark along with complete source code will be released.

Data source. The majority of our data come from the preparation problems for the GATE (Graduate Aptitude Test in Engineering) exam, related course work at MIT OpenCourseWare, and Chinese university exams (translated to English).

C. Used prompts

C.1. Prompt for Caption Generation

The prompt for caption generation using GPT-4o is as follows:

Here is a diagram related to an Electrical and Electronics Engineering problem. Please describe the diagram in complete detail, ensuring to include all relevant components, symbols, connections, values (such as resistances, voltages, or frequencies), and any other key parameters. Your description should be precise enough to fully replace the diagram and be usable for solving the EEE problem across topics like circuits, signal processing, electromagnetics, power systems, and control systems.

C.2. Spurious caption generation prompt

The prompt for spurious caption generation using GPT-4 on EEE-Bench is as follows:

I will give a caption of an Electrical and Electronics Engineering (EEE) problem image and the associated question. Please help me rewrite the image caption so that it contains incorrect information, which could lead to an incorrect solution. Please only output the revised incorrect caption. Please do not change name of variables, e.g., F, G. Please Do not change too much. Here are some examples of how you can introduce incorrect information into an Electrical and Electronics Engineering (EEE) image caption to potentially mislead a model into generating a wrong solution:

---

[1] https://gate2025.iitr.ac.in/

[2] https://www.allaboutcircuits.com/worksheets/

[3] https://ocw.mit.edu/

Table 4. The task instructions for different question and answer types in answer extraction. Here, Float (1) refers to a floating-point number with one decimal place, Float (2) refers to a floating-point number with two decimal places, and Float (3) refers to a floating-point number with three decimal places.

Question type	Answer type	Task instruction
multiple-choice	Text	Please answer the question and provide the correct option letter, e.g., A, B, C, D, at the end.
Free-form	Integer	Please answer the question requiring an integer answer and provide the final value, e.g., 1, 2, 3, at the end.
Free-form	Float (1)	Please answer the question requiring a floating-point number with one decimal place and provide the final value, e.g., 1.2, 1.3, 1.4, at the end.
Free-form	Float (2)	Please answer the question requiring a floating-point number with two decimal places and provide the final value, e.g., 1.23, 1.34, 1.45, at the end.
Free-form	Float (3)	Please answer the question requiring a floating-point number with three decimal places and provide the final value, e.g., 1.200, 1.300, 1.400, at the end.
Free-form	List	Please answer the question requiring a Python list as an answer and provide the final list, e.g., [1, 2, 3], [1.2, 1.3, 1.4], at the end.

1. Incorrect component identification: Original: "The circuit contains a 10Ω resistor and a 5V voltage source." Misleading: "The circuit contains a 100Ω resistor and a 12V voltage source."

2. Wrong power supply voltage: Original: "The circuit is powered by a 5V DC supply." Misleading: "The circuit is powered by a 15V AC supply."

3. Inverted polarity for components: Original: "The diode is forward biased in this circuit." Misleading: "The diode is reverse biased in this circuit."

4. Incorrect current direction: Original: "The current flows clockwise through the loop." Misleading: "The current flows counterclockwise through the loop."

5. Swapping component values: Original: "The capacitor has a capacitance of 100μF, and the inductor is 10mH." Misleading: "The capacitor has a capacitance of 10μF, and the inductor is 100mH."

6. Incorrect assumptions about the circuit operation: Original: "The transistor is in active mode." Misleading: "The transistor is in saturation mode."

7. Wrong frequency or waveform type: Original: "The AC signal is a 60Hz sine wave." Misleading: "The AC signal is a 50Hz square wave."

8. Incorrect type of circuit: Original: "This is a simple RC low-pass filter."
Misleading: "This is a high-pass filter." 9. Incorrect labeling of nodes or points: Original: "Measure the voltage at node A, which is connected to the positive terminal." Misleading: "Measure the voltage at node A, which is connected to the negative terminal."

10. Misleading behavior of the circuit: Original: "The capacitor charges fully after 5 seconds." Misleading: "The capacitor discharges completely after 5 seconds."

11. Incorrect power consumption: Original: "The total power consumed by the circuit is 50W." Misleading: "The total power consumed by the circuit is 100W."

12. Mislabeling passive vs. active components: Original: "The operational amplifier is used as an active component in this circuit." Misleading: "The operational amplifier is used as a passive component in this circuit."

13. Incorrect reading: Original: "Only when A=1,B=1,C=1, F=1." Misleading: "when A=1,B=1,C=1, F=0, when A=0,B=1,C=1, F=1.

The prompt for spurious caption generation using GPT-4 on Math-Vista is as follows:

I will provide a caption of a Math problem image along with the associated question. Please help me rewrite the image caption so that it contains incorrect information, potentially leading to an incorrect solution. Please only provide the revised incorrect caption. Do not change the names of variables, such as F or G, and avoid making extensive changes.

Please do not change the name of variables, e.g., F, G. Please Do not change too much. Please do not change the essential meaning of original text and only change subtle places like values.

Below are some examples of how you can introduce incorrect information into a Math problem caption to mislead a model:

Table 5. Generating parameters for Closed-Source LMMs.

Model	Generation Setup
GPT-4	"model" : "gpt-4", "temperature" : 0, "max_tokens" : 1024
GPT-4o	"model" : "gpt-4o-2024-08-06", "temperature" : 0, "max_tokens" : 1024
GPT-4V	"model" : "gpt-4-turbo", "temperature" : 0, "max_tokens" : 1024
GPT-4o-mini	"model" : "gpt-4o-mini", "temperature" : 0, "max_tokens" : 1024
Gemini 1.5 Pro	"model" : "gemini-1.5-pro", "temperature" : 0, "max_tokens" : 1024
Gemini 1.5 Flash	"model" : "gemini-1.5-flash", "temperature" : 0, "max_tokens" : 1024
Qwen-VL-Max	"model" : "qwen-vl-max", "temperature" : 0, "max_tokens" : 1024
Qwen-VL-Plus	"model" : "qwen-vl-plus", "temperature" : 0, "max_tokens" : 1024

Table 6. Generating parameters for Open-Source LMMs.

Model	Generation Setup
Llava-1.6-7B	do_sample = False, temperature = 0, max_new_tokens = 1024
Llava-1.6-13B	do_sample = False, temperature = 0, max_new_tokens = 1024
SPHINX-v2-1k	do_sample = False, temperature = 0, max_new_tokens = 1024
MiniCPM-V-2.6	do_sample = False, temperature = 0.1, max_new_tokens = 1024
DeepSeek-VL-7B	do_sample = False, max_new_tokens = 1024
Phi3-Vision-4.2B	do_sample = False, temperature = 0, max_new_tokens = 1024
InternVL2-26B	do_sample = False, temperature = 0, max_new_tokens = 1024
InternLM-XComposer2.5-7B	do_sample = False
GLM-4V-9B	do_sample = False, max_length = 1024, temperature = 0

Incorrect value of constants: Original: 'The equation has a constant value of 3.' Misleading: 'The equation has a constant value of 5.'
Wrong interpretation of angles: Original: 'Angle ABC is 90°.' Misleading: 'Angle ABC is 60°.'
Incorrect function transformation: Original: 'The function is f(x) = 2x + 3.' Misleading: 'The function is f(x) = 2x - 3.'
Wrong area formula: Original: 'The area of the circle is given by πr^2.' Misleading: 'The area of the circle is given by $2\pi r$.'
Incorrect limit definition: Original: 'The limit of the function as x approaches 2 is 4.' Misleading: 'The limit of the function as x approaches 2 is 5.'
Wrong assumption in calculus: Original: 'The derivative of x^2 is 2x.' Misleading: 'The derivative of x^2 is x.'
Incorrect triangle property: Original: 'In this right triangle, the hypotenuse is 10.' Misleading: 'In this right triangle, the hypotenuse is 8.'
Wrong direction in vector problems: Original: 'The vector points in the positive x-direction.' Misleading: 'The vector points in the negative x-direction.

C.3. Prompts for response generation

We follow MathVista [56] to prepend an instructional prompt header into each question. The prompt can be found in Table 4.

C.4. Prompt for caption generation

Here we provide information regarding the image caption generation. Concretely, we generate for each image a caption X_c using GPT-4o to let it describe detailed key visual information related to potential EEE problem-solving (the prompt is provided in Appendix) and prepend such caption to the original question prompt X_q such that the new query is (X_{ins} +
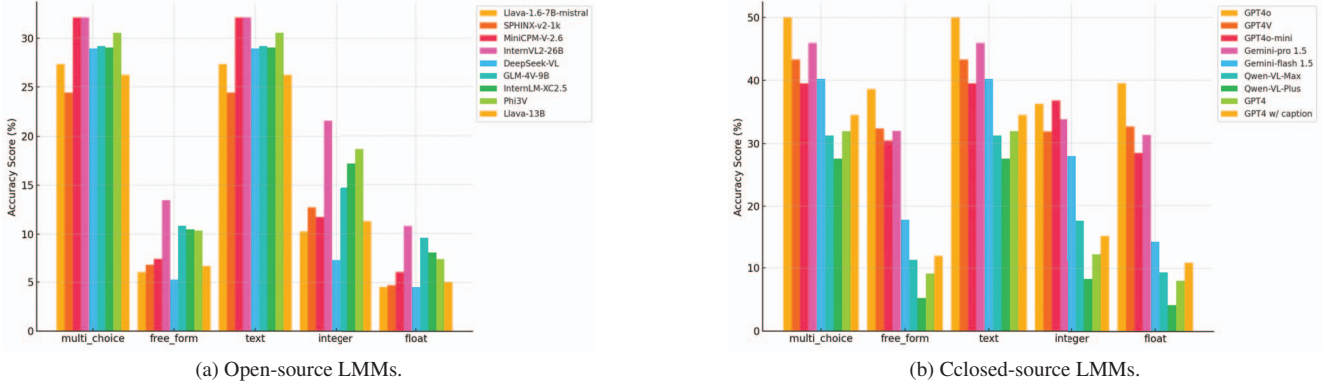
(a) Open-source LMMs.

(b) Cclosed-source LMMs.

Figure 6. Accuracy across different question types and answer types.

Model	ALL	GPS	TQA	FQA	MWP	VQA	SCI	STA	ALG	GEO	ARI	LOG	NUM
					GPT-4o								
Δ	-8.40	-0.48	-12.03	-4.46	-21.51	-6.70	+0.42	-3.20	-7.97	+0.42	-12.18	+2.70	-11.11
					Gemini-Pro 1.5								
Δ	-3.70	-6.20	-8.86	+0.00	-5.38	+0.00	-1.66	-1.66	-7.81	-6.27	-0.89	-2.02	-2.08

Table 7. Performance Delta of using spurious caption as extra text on MathVista. ALL: overall accuracy. Task types: FQA: figure question answering, GPS: geometry problem solving, MWP: math word problem, TQA: textbook question answering, VQA: visual question answering. Mathematical reasoning types: ALG: algebraic reasoning, ARI: arithmetic reasoning, GEO: geometry reasoning, LOG: logical reasoning, NUM: numeric commonsense, SCI: scientific reasoning, STA: statistical reasoning. Positive changes are in blue and negative changes in red.

$X_c) + X_q$ where X_{ins} is a system instruction prompt which reads "*Image caption context:*".

D. More Experimental Results

In this section, we provide more experimental results.

Results on different question types and answer types. We first discuss the performance of LMMs on different question types: multi-choice and free-form, and answer types: text, integer, and float (since the number of list-type questions is too small, we ignore them here). The results are shown in Figure 6. We can find that the open-source models perform very poorly on free-form problems. Since free-form problems need accurate answers, the computational ability of open-source models can not support them in obtaining accurate final answers. Compared with open-source models, closed-source models show a better balance of accuracy between multi-choice problems and free-form problems.

Results on MathVista with spurious captions. We provide the results on MathVista under the setting of spurious captions. The results are shown in Table 7. Similar to experiments done on EEE-Bench, adding spurious captions into text will significantly degrade LMMs' performance.

Examples of response from LMMs. In Figure 40 and Figure 41, we provide qualitative examples of responses for LMMs for reference.

Wrong responses from GPT-4o augmented with captions. We discuss examples of wrong responses from GPT-4o with augmented captions. Figure 28 and Figure 29 show examples. We found that these errors stem from incorrectly generated captions, which aligns with the concept of LMM laziness.

E. Model, Hyperparameter, and Additional Prompting Methods

Model and hyperparameter. The release time and source of open-source and closed-source models can be found in Table 8. For closed-sourced models, we access them with APIs and adopt the generation scheme shown in Table 5. and then run the inference with CPUs, which typically completes within several hours. For all open-source models, we utilize a cluster

with 8 NVIDIA 40GB A100 GPUs to run the inference, and we follow the hyper-parameter settings specified in Table 6.

Additional prompting methods. Some of the method selections and descriptions are based on [72].
- **0-Shot CoT** [37]: A variant of CoT where the exemplar is omitted and instead an instruction "Let's think step by step" is added.
- **1-Shot CoT** [86]: Introduces a single example to guide models in producing step-by-step reasoning for enhanced response accuracy.
- **LTM** [103]: A stepwise prompting approach that gradually increases complexity, letting the model break down the problem into sub-problems where the solution to the final sub-problem is adopted.
- **PROGRAM** [19]: Prompting the model to write a Python program to solve the problem.
- **Retrieval** [25]: Including a relevant retrieved document to mimic the case where a RAG [38] is used.
- **Self-Correction** [65]: Prompting the model to reexamine its own answer and make necessary adjustments.
- **Self-Consistency** [83]: Sampling several solutions (from both same method or different methods) from the LMM and taking the majority vote. Note that self-consistency is orthogonal to all previous methods.
- **DDCot** [101]: Prompting the model to break the task down to sub-problems and then solving them individually.
- **Cantor** [24]: Breaking the problem into sub-problems and then utilizing different model instances to detail with different kinds of sub-tasks.

Table 8. **The Release Time and Model Source of LMMs Used in EEE-Bench.**

Model	Release Time	Source
GPT-4o	2024-05	https://gpt4o.ai/
GPT-4V	2024-04	https://openai.com/index/gpt-4v-system-card/
GPT-4o-mini	2024-07	https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence
Gemini 1.5 Pro	2024-05	https://deepmind.google/technologies/gemini/pro/
Gemini 1.5 Flash	2024-05	https://deepmind.google/technologies/gemini/flash/
Qwen-VL-Plus	2023-11	https://help.aliyun.com/zh/dashscope/developer-reference/vl-plus-quick-start
Qwen-VL-Max	2024-01	https://help.aliyun.com/zh/dashscope/developer-reference/vl-plus-quick-start
LLaVA-1.6-7B	2024-03	https://huggingface.co/liuhaotian/llava-v1.6-mistral-7b
LLaVA-1.6-13B	2024-03	https://huggingface.co/liuhaotian/llava-v1.6-vicuna-13b
SPHINX-Plus	2023-11	https://huggingface.co/Alpha-VLLM/LLaMA2-Accessory/tree/main/finetune/mm/SPHINX/SPHINX-v2-1k
MiniCPM-V-2.6	2024-08	https://github.com/OpenBMB/MiniCPM-V
DeepSeek-VL-7B	2024-03	https://huggingface.co/deepseek-ai/deepseek-vl-7b-chat
Phi3-Vision-4.2B	2024-05	https://huggingface.co/microsoft/Phi-3-vision-128k-instruct
InternVL2-26B	2024-07	https://huggingface.co/OpenGVLab/InternVL2-26B
InternLM-XComposer2.5-7B	2024-07	https://huggingface.co/internlm/internlm-xcomposer2d5-7b
GLM-4V-9B	2024-06	https://huggingface.co/THUDM/glm-4v-9b

Table 9. **Key Statistics of EEE-Bench.**

Statistic	Number
Total questions	2,860
- Multiple-choice questions	2,059 (72.0%)
- Free-form questions	801 (28.0%)
- **Newly collected questions**	**2,860(100%)**
Number of unique images	2813
Number of unique questions	2826
Number of unique answers	329
Maximum question length	466
Minimum question length	7
Maximum answer length	8
Average question length	57.4
Average answer length	1.4

Figure 7. **Subject Distribution of EEE-Bench.** DLC&MPD: Digital Logic Circuits and Microprocessor Design, PS: power systems. NA: network analysis
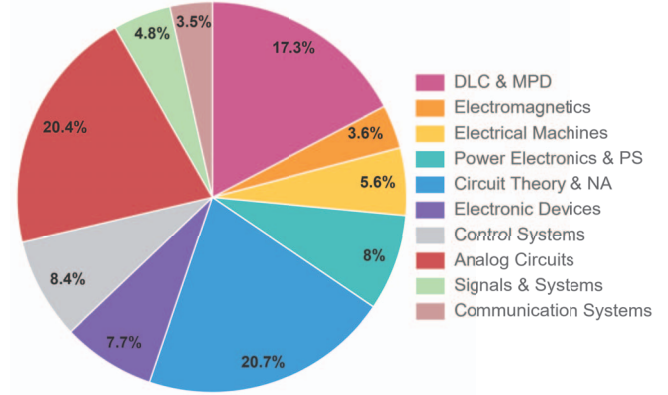
F. More About EEE-Bench

In this section, we provide more details about EEE-Bench.

Release. We release the benchmark as organized in two formats: 1) one complete set and 2) sub-domains. We also release metadata annotations denoting question title (stored in markdown format), image path, answer, question type, image ID, source, source image ID, topic, query, answer type, and choices in JSON format similar to [56]. For the question title, we also prepend an instructional prompt, see Appendix C for details. We also release the complete evaluation source code.

More details on benchmark summary. The key statistics of EEE-Bench are shown in Figure 9 and subject distribution is shown in Figure 7. Next, we provide a breakdown of the 10 sub-domains. *Digital Logic Circuits and Microprocessor Design* (496 examples) assess the design and functionality of circuits central to computing systems like processors and memory. *Electronic Devices* (219 examples) evaluate the understanding of components such as diodes and transistors, foundational to all electronics. *Communication Systems* (101 examples) assess signal transmission and processing, crucial for wireless networks and satellite communications. *Signals and Systems* (138 examples) evaluate the mathematical representation and manipulation of signals, fundamental for control systems and digital signal processing. *Electrical Machines* (161 examples) assess knowledge of devices like motors and generators, essential for power generation and industrial applications. *Electromagnetics* (102 examples) evaluate the principles of electromagnetic fields and waves, vital for RF communication and antenna design. *Control Systems* (239 examples) assess feedback mechanisms and system stability, critical for automation and robotics. *Power Electronics and Power Systems* (229 examples) evaluate the conversion and control of electrical power, key to renewable energy systems and power grids. *Circuit Theory and Network Analysis* (592 examples) assess both AC and DC circuit analysis, forming the foundation of all electrical applications. Lastly, *Analog Circuits* (583 examples) evaluate continuous-time circuit design, essential for applications like audio, RF, and sensor interfaces. We leave more detailed descriptions of these 10 subdomains to Appendix G.

More details on benchmark quality. Benchmarks catered toward LMM evaluations should provide unbiased, fair, and exhaustive evaluations. EEE-Bench achieves these qualities by ensuring diversity in terms of both type and sub-domain. The dynamic question settings reduce the potential for bias, and the wide coverage of sub-domains captures the diverse nature of engineering problems while ensuring a comprehensive and fair representation of the EEE domain. More importantly, the diverse nature of EEE-bench also enables granular evaluation of model performance, thus providing a framework for more nuanced by-type and by-sub-domain analysis (subsection 3.2), readily unveiling the underlying strengths and weaknesses of different models. Varied difficulty levels are also introduced to best mimic real-world situations and enhance the interpretability of the assessment. We note that the best-performing proprietary model GPT-4o [63] and open-source model

InternVL2-26B [17] set at 46.78% and 26.89% overall accuracy respectively, while the averages of all proprietary models and open-sourced models are 35.28% and 23.19% respectively. For more detailed results and analysis, please refer to subsection 3.2.

Significance of vision in EEE-Bench. Meaningful multimodal benchmarks should place largely equal weights on both visual and language inputs. Recent studies have unveiled that some previous multimodal benchmarks lack visual depth and complexity thus making their visual input ad-hoc [77, 93]. In EEE-Bench, we ensure that vision plays an indispensable role. **First**, as shown in Table 1, the text-only baseline with GPT-4 results in a 21.26% and 14.69% drop in accuracy compared to GPT-4o and GPT-4V respectively with complete input. **Second**, we found that the improved text-only-with-caption baseline merely improves accuracy by 2.7% with GPT-4 (Table 1). These results highlight the significance and indispensability of vision in EEE-Bench.

Comprehensive evaluation of LMM's various abilities. EEE-Bench encompasses 10 pivotal subjects within the field of electrical and electronics engineering, extending from circuit analysis to control systems. Each subject demands specialized capabilities for addressing unique challenges. More detailed discussions are provided in Appendix F and subsection 3.2. For instance, as shown in Figure 18, tackling problems within the Digital Logic Circuits and Microprocessor Design discipline requires models to deeply understand intricate logic relationships in digital circuits and reason effectively. These problems demand a thorough comprehension of logical structuring and timing analysis, yet they do not heavily rely on computational power or advanced mathematical computations. In contrast, many challenges from Circuit Theory and Network Analysis may present less complexity regarding logic understanding but require extensive computational processing and the use of sophisticated mathematical tools. This involves solving complex equations and applying numerical methods to analyze and interpret circuit behaviors and responses under various conditions. A detailed introduction about the characteristics of each subject's problems and the required ability for LMMs to solve them are discussed in Appendix H. The juxtaposition of these disciplines within EEE-Bench highlights the diverse skill sets needed to address the spectrum of issues prevalent in the field of electrical and electronics engineering, therefore effectively examining various abilities of existing LMMs.

Image examples. Figure 8 presents images from Digital Logic Circuits and Microprocessor Design. Figure 9 presents images from Circuit Theory and Network Analysis. Figure 10 presents images from Analog Circuits. Figure 11 presents images from Power Electronics and Power Systems. Figure 12 presents images from Signals and Systems. Figure 13 presents images from Communication Systems. Figure 14 presents images from Control Systems. Figure 15 presents images from Electronic Devices. Figure 16 presents images from Electrical Machines. Figure 17 presents images from Electromagnetics. EEE-Bench covers a wide range of images and thus can comprehensively evaluate the visual reasoning abilities of LMMs.

Table 11. Required Abilities for LMM among EEE Subjects

Subject	Required LMM Abilities
(1) Digital Logic Circuits and Microprocessor Design	Solving problems in this domain requires models to deeply understand intricate logic relationships in digital circuits and reason effectively. While these problems demand strong logical structuring and timing analysis capabilities, they do not heavily rely on computational power or advanced mathematical computations. The ability to recognize logic patterns and reason through circuit operations is essential.
(2) Circuit Theory and Network Analysis	LMMs tackling problems in this subject require extensive computational processing abilities. The models must be proficient in solving complex equations and applying numerical methods to analyze and interpret circuit behaviors under varying conditions. Although less complex in terms of logic, these problems require strong numerical and computational skills to manage transient and steady-state responses.

Table 11 – continued from previous page

| Subject	Required LMM Abilities
(3) Analog Circuits	For analog circuits, LMMs must handle complex visual contexts, including intricate circuit diagrams. These problems demand both sophisticated vision recognition abilities to interpret circuit structures and strong computational skills to predict continuous signal behavior. Advanced mathematical tools, such as Laplace transforms and frequency domain analysis, are required for solving circuit dynamics and performance analysis.
(4) Power Electronics and Power Systems	LMMs solving power electronics and power systems problems must manage large-scale computational tasks related to energy conversion, grid dynamics, and efficiency optimization. Although the visual elements might not be as complex as other subjects, the computational effort needed for modeling and simulating power systems is significant. Strong numerical capabilities and reasoning about energy flow are crucial.
(5) Signals and Systems	This domain requires LMMs to process and analyze signals in both time and frequency domains. Vision recognition skills are needed to understand the visual representations of signals, such as waveforms and spectrograms, while computational tools like Fourier and Laplace transforms must be applied to analyze system responses. The models need to handle both symbolic reasoning and numerical calculations.
(6) Communication Systems	Although the visual information in communication systems tends to be straightforward, solving these problems demands considerable computational effort. LMMs must decode and simulate communication protocols, analyze noise, and optimize transmission systems. Numerical accuracy and simulation capabilities are key for optimizing communication channels and analyzing signal integrity.
(7) Control Systems	LMMs addressing control system problems must excel in both reasoning and computation. Feedback systems require sophisticated reasoning to design and stabilize control loops, while mathematical modeling and simulation of dynamic systems call for strong computational abilities. Complex visual representations of system dynamics, such as block diagrams, require models to recognize system structures and respond accordingly.
(8) Electronic Devices	In this subject, LMMs must understand the physical principles of semiconductor devices, but the problems primarily involve numerical computation to predict device behaviors. Visual contexts, such as device schematics, are generally simple, but the computational tasks can be intensive, requiring models to handle material properties and electrical characteristics with high precision.
(9) Electrical Machines	Problems in this domain involve electromechanical energy conversion, and LMMs must reason about mechanical and electrical interactions. The visual complexity is moderate, but models must be proficient in numerical simulations to predict machine performance under different conditions. Computational tools are essential for solving mechanical equations and analyzing energy transfer processes.
(10) Electromagnetics	LMMs need to possess strong reasoning skills to interpret complex visual contexts like field distributions and wave propagation diagrams. Additionally, these problems demand high computational power to solve Maxwell's equations and simulate electromagnetic fields. Numerical accuracy and proficiency in solving partial differential equations are key to successfully addressing challenges in this domain.
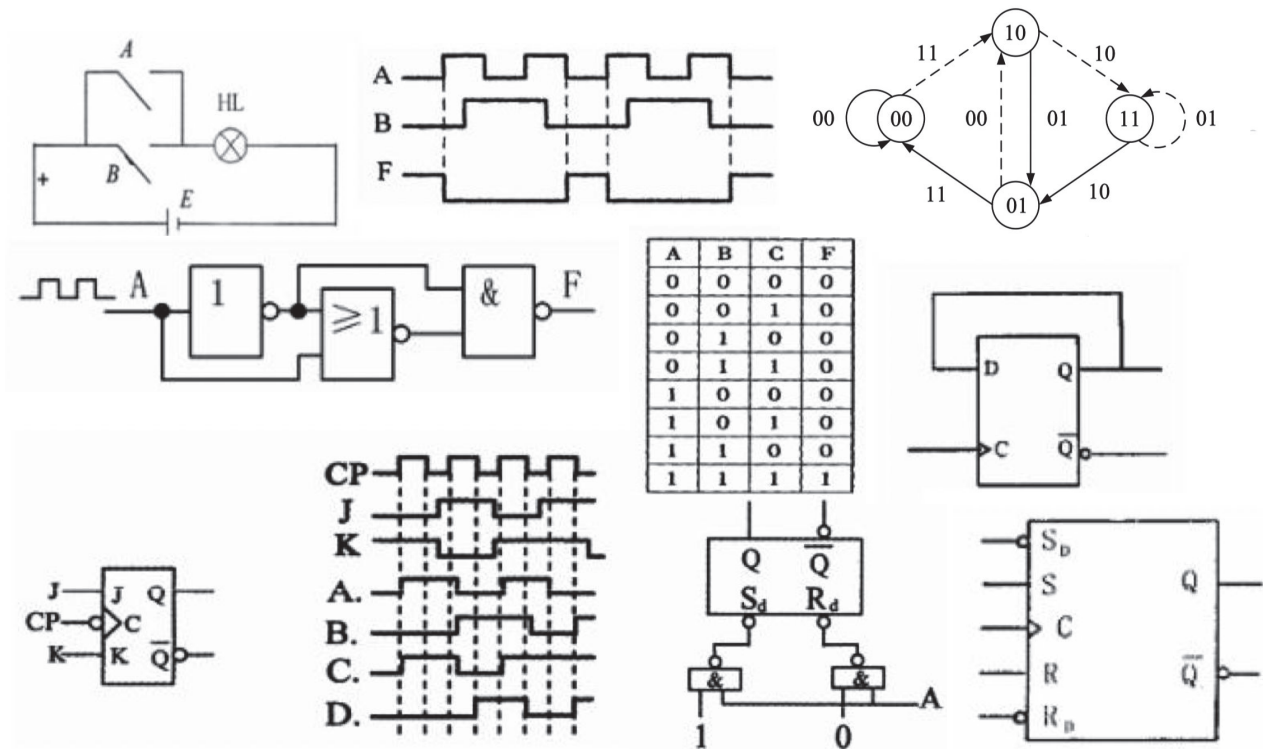
Figure 8. Examples of images from Digital Logic Circuits and Microprocessor Design
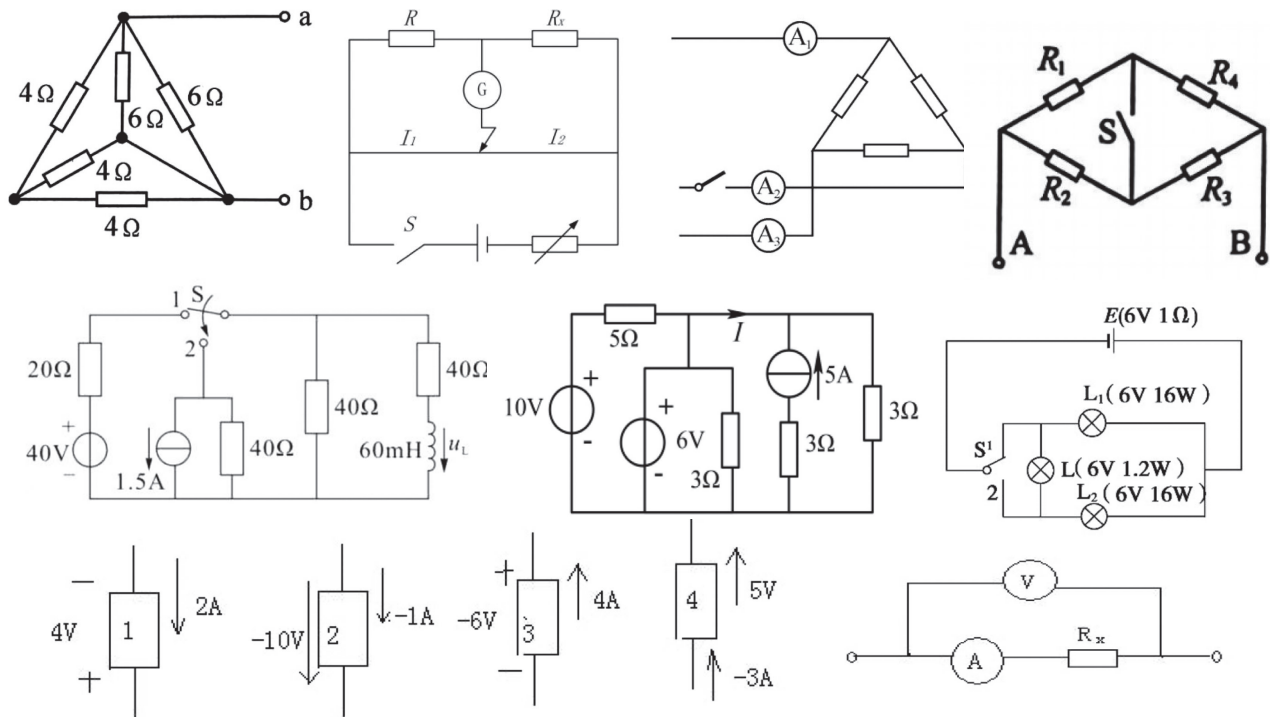

Figure 9. Examples of images from Circuit Theory and Network Analysis

G. Detailed Introduction to Ten Subjects

To better help readers understand the key points of these ten subjects, we provide a detailed introduction for each in this section. Further information on each subject is available in Table 10.
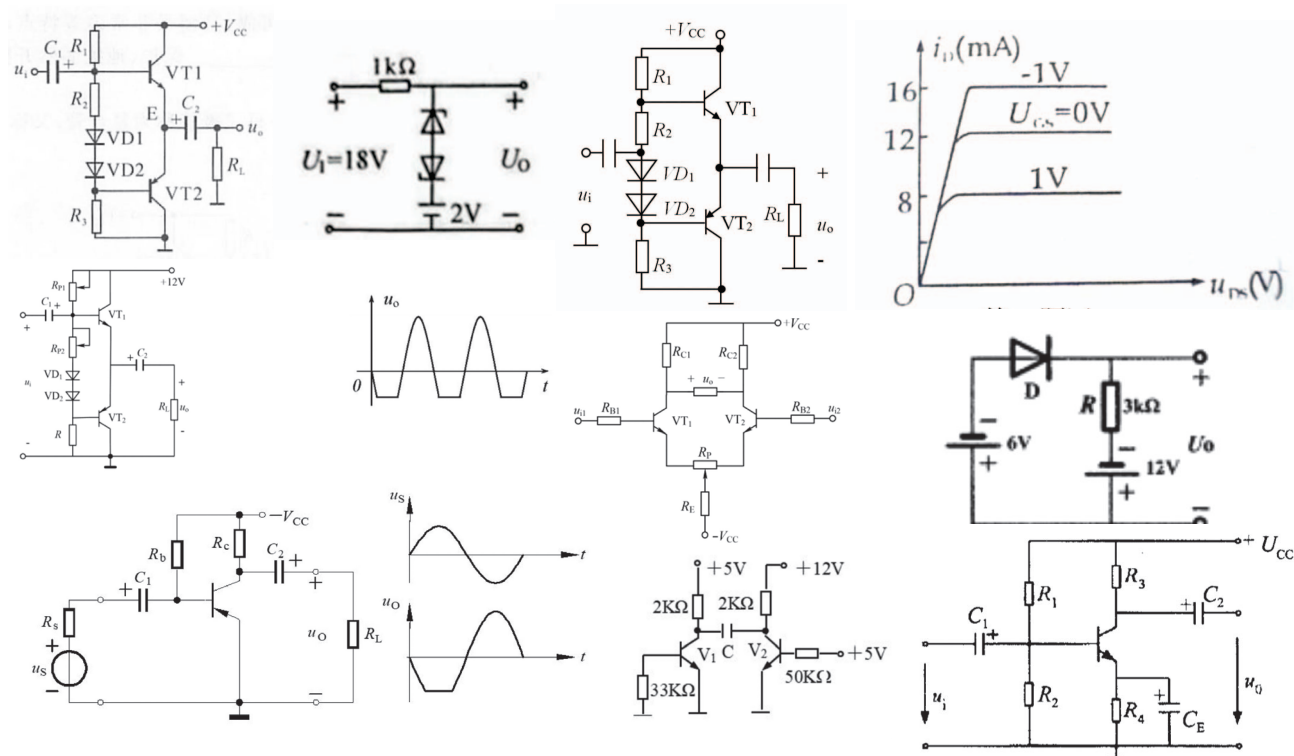
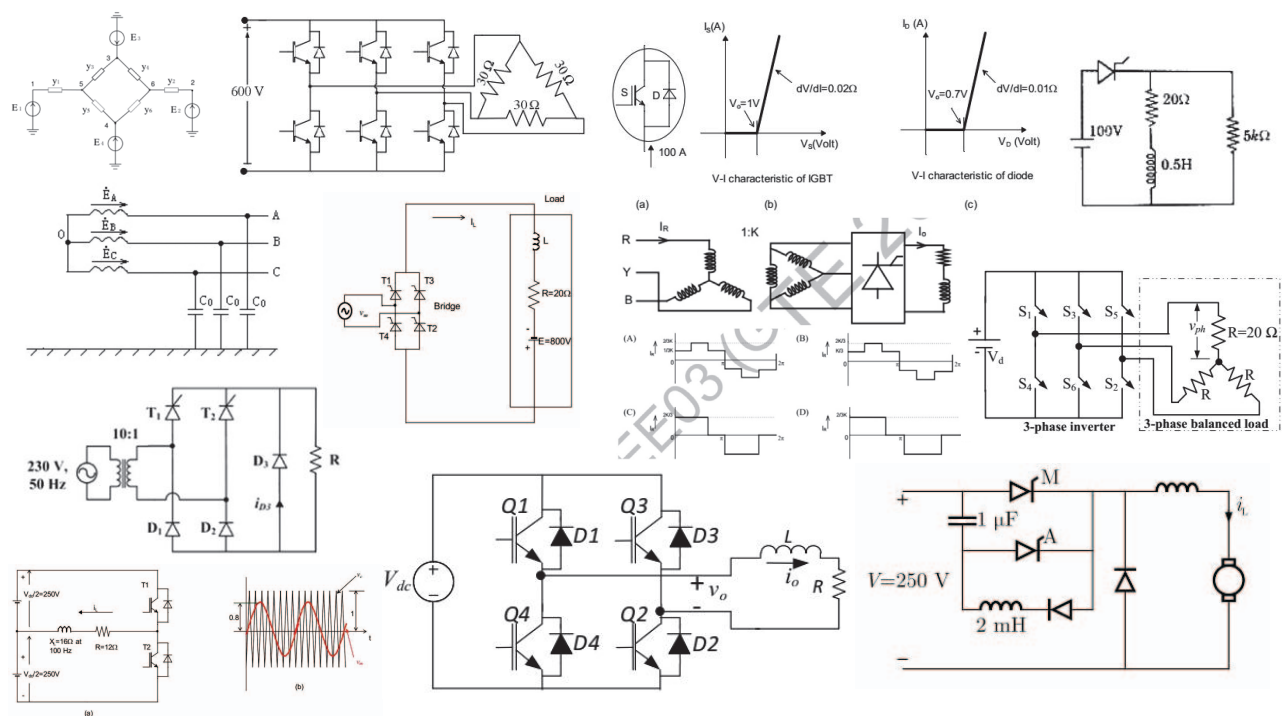Figure 10. Examples of images from Analog Circuits

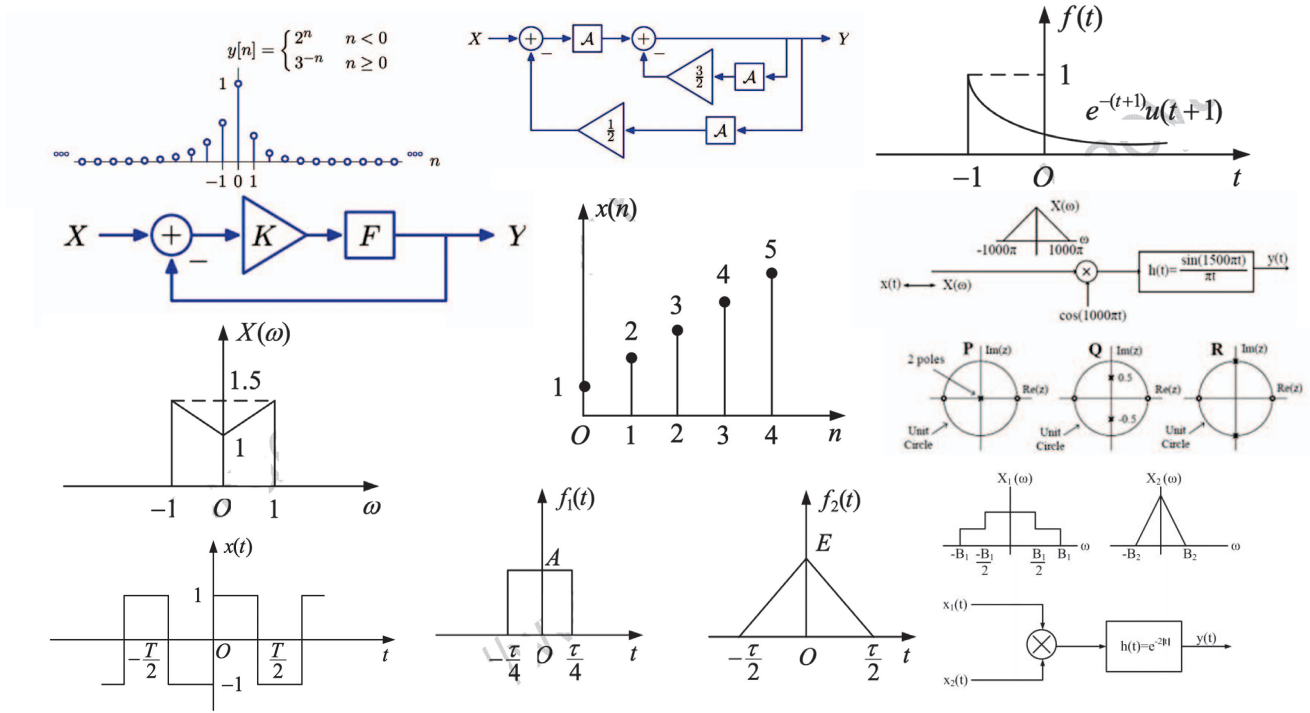Figure 11. Examples of images from Power Electronics and Power Systems

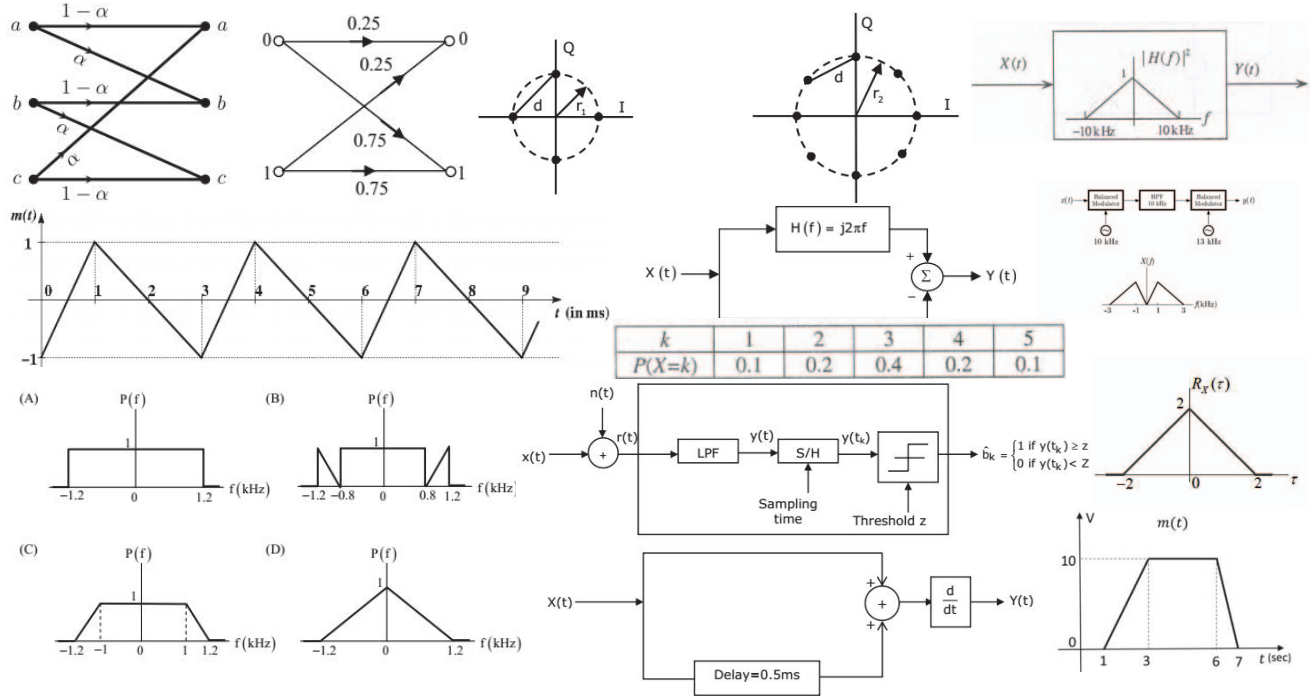Figure 12. Examples of images from Signals and Systems

Figure 13. Examples of images from Communication Systems

H. Required Ability for LMMs to Solve EEE Problems

In this section, we discuss the characteristics of each subject's problems and the required ability for LMMs to solve them. The discussion can be found in Table 11. Some examples of problems and detailed solutions in EEE-Bench are shown in
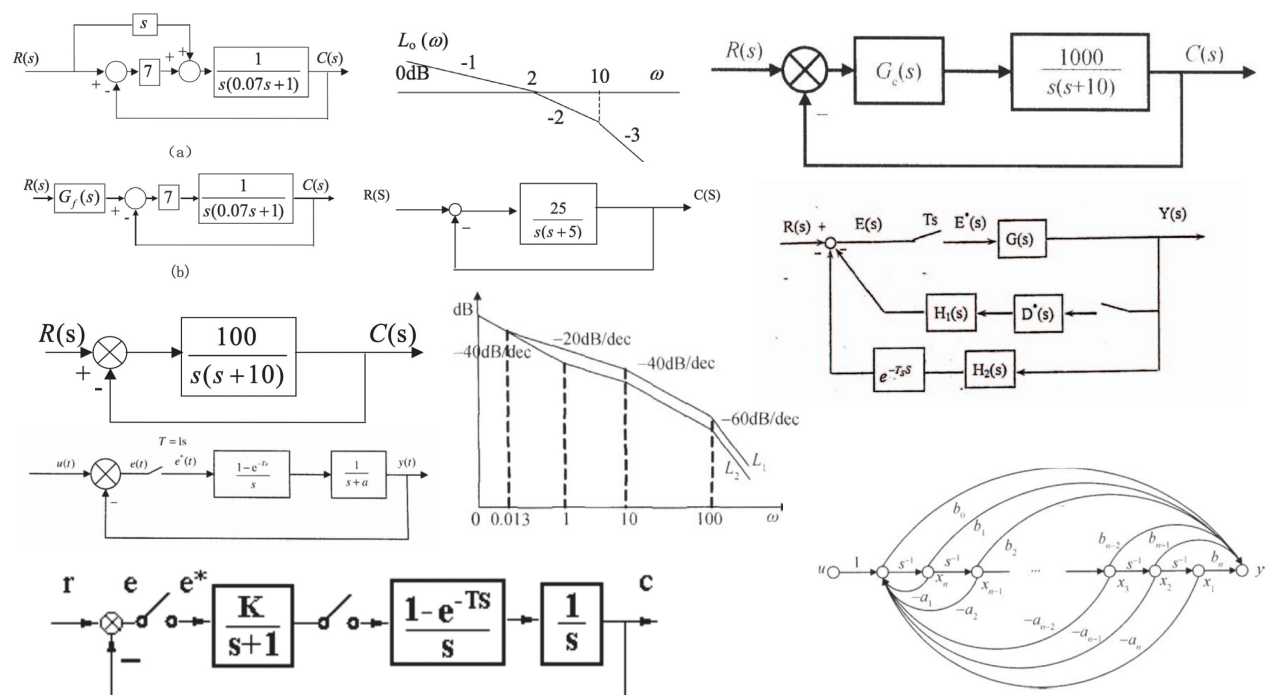
Figure 14. Examples of images from Control Systems
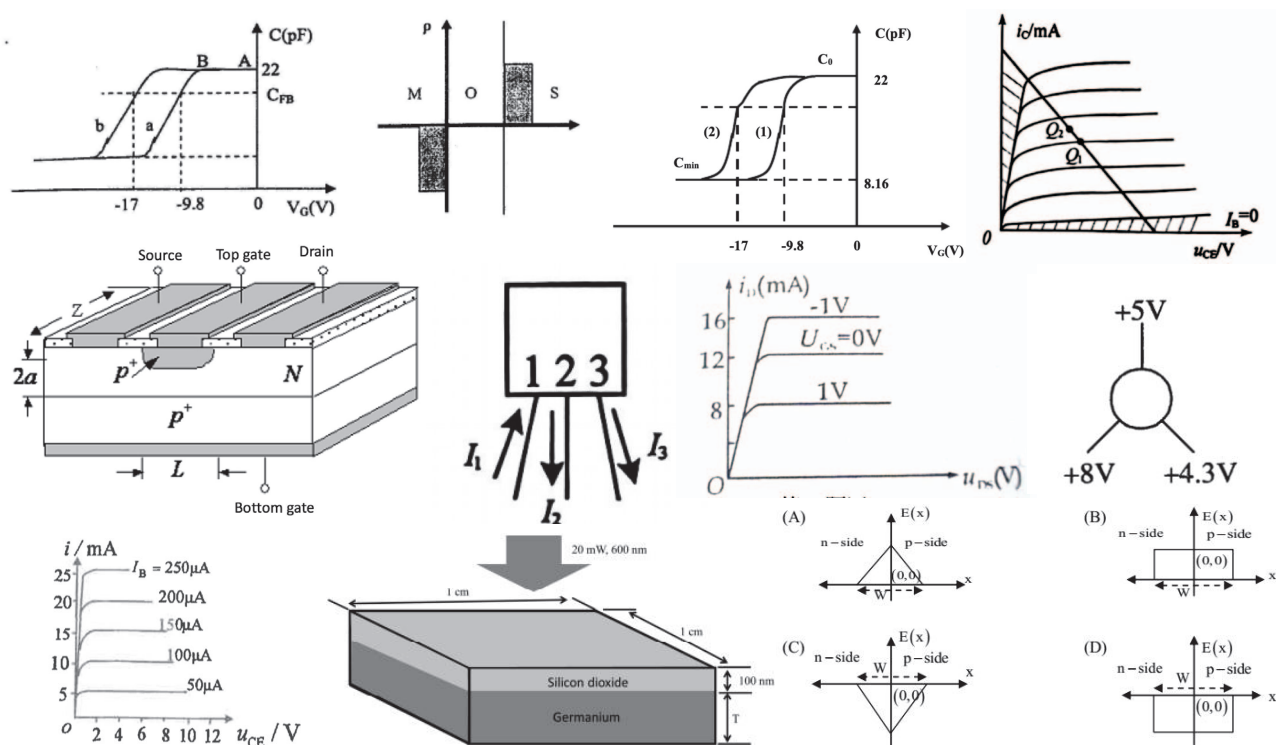
Figure 15. Examples of images from Electronic Devices

Figure 16. Examples of images from Electrical Machines

Figure 17. Examples of images from Electromagnetics

I. More examples of error analysis

In this section, we give more examples of error analysis. Figure 24 and Figure 25 are examples of image perception errors. Figure 26 and Figure 27 are examples of reasoning errors.

Table 10. Description and Key Knowledge in EEE Subjects

Subject	Description and Key Knowledge
(1) Digital Logic Circuits and Microprocessor Design	Focuses on boolean logic, logic gates design, and the sequential circuits fundamental to digital systems. Key challenges include microprocessor architecture analysis, timing issues, and peripheral interfacing, requiring a deep understanding of logic operations and system integration.
(2) Circuit Theory and Network Analysis	Involves detailed study of Ohm's and Kirchhoff's laws, network theorems, and electrical components. Emphasizes solving complex networks through techniques like mesh and nodal analysis, with problems often revolving around transient and steady-state behaviors in electrical circuits.
(3) Analog Circuits	Covers operational amplifiers, filters, and oscillators, focusing on the design and analysis of circuits for continuous signal processing. Challenges include designing robust power amplifiers and ensuring accurate response analysis in varying operational conditions.
(4) Power Electronics and Power Systems	Examines electrical power conversion and control using semiconductor devices. Key areas include designing efficient converters and inverters, and managing complex power systems with an emphasis on optimizing grid dynamics and energy distribution.
(5) Signals and Systems	Deals with the representation and analysis of signals and the behavior of systems, employing tools like Fourier and Laplace transforms. Challenges include analyzing both time-domain and frequency-domain data to understand system properties and behaviors.
(6) Communication Systems	Focuses on techniques for efficient signal transmission and reception. Involves studying modulation schemes, coding techniques, and noise analysis with problems often related to designing and optimizing communication links and networks for reliability and performance.
(7) Control Systems	Addresses the design and analysis of dynamic systems using feedback. Involves stability analysis and controller design, with practical challenges in tuning controllers like PID for desired system responses and ensuring long-term stability and reliability.
(8) Electronic Devices	Explores semiconductor physics and the application of devices like diodes and transistors. Problems focus on device operation under different conditions and the integration of special purpose devices such as LEDs and photodetectors in practical applications.
(9) Electrical Machines	Studies the principles of operation and design of electric machines such as transformers and motors. Key challenges include analyzing performance, designing for efficiency, and understanding the electromechanical interactions in different operating conditions.
(10) Electromagnetics	Explores Maxwell's equations and their applications in electromagnetic field theory and wave propagation. Challenges include designing transmission lines, waveguides, and antennas, and ensuring electromagnetic compatibility in various applications.

J. More examples of LMM laziness

In this section, we provide more examples of the laziness phenomenon of existing LMMs. Figure 30, Figure 31, Figure 32, Figure 33, Figure 34, Figure 35, Figure 36 and Figure 37 show examples of LMM laziness phenomenon on EEE-Bench.

Logical circuits

5 V

J Q_A

K $\bar{Q}_A$

J Q_B

K

CLK

Question: The current state Q_A, Q_B of a two JK flip-flop system is 00 . Assume that the clock rise-time is much smaller than the delay of the JK flip-flop. The next state of the system is

A. 00
B. 01
C. 11
D. 10

Solution: From the figure we get $J_A = K_A = A$ and $J_B = K_B = \bar{Q}_A$. So we have:

Clock	$J_A K_A$	$J_B K_B$	$Q_A Q_B$
0	11	11	00
1	-	-	11

So, next state will be 11 .

Circuit theory

2 kΩ   2 kΩ

I_D D

V_{in}

V_{out} 2 kΩ

D

V_D

I_D (mA)

20

10

0

0   0.7 V_D (V)

Question: A circuit and the characteristics of the diode (D) in it are shown. The ratio of the minimum to the maximum small signal voltage gain

$\frac{\partial V_{out}}{\partial V_{in}}$ is ___(rounded off to two decimal places).

Solution: When diode is ON, as $r_d(ON)=0$, the 2kΩ resistor in parallel to the diode becomes open circuit.

So $V_{out} = V_{in} \times \frac{2}{4} = \frac{V_{in}}{2}$,

then $\frac{\partial V_{out}}{\partial V_{in}}|_{max} = 0.5 \ldots (i)$.

When diode is off, $r_d(OFF) = \infty$.

So total $R_{eq} = 2 + 2 + 2 = 6k\Omega$.

Then $V_{out} = \frac{4V_{in}}{6} = \frac{2V_{in}}{3}$.

So $\frac{\partial V_{out}}{\partial V_{in}}|_{min} = \frac{1}{3} \ldots (ii)$.

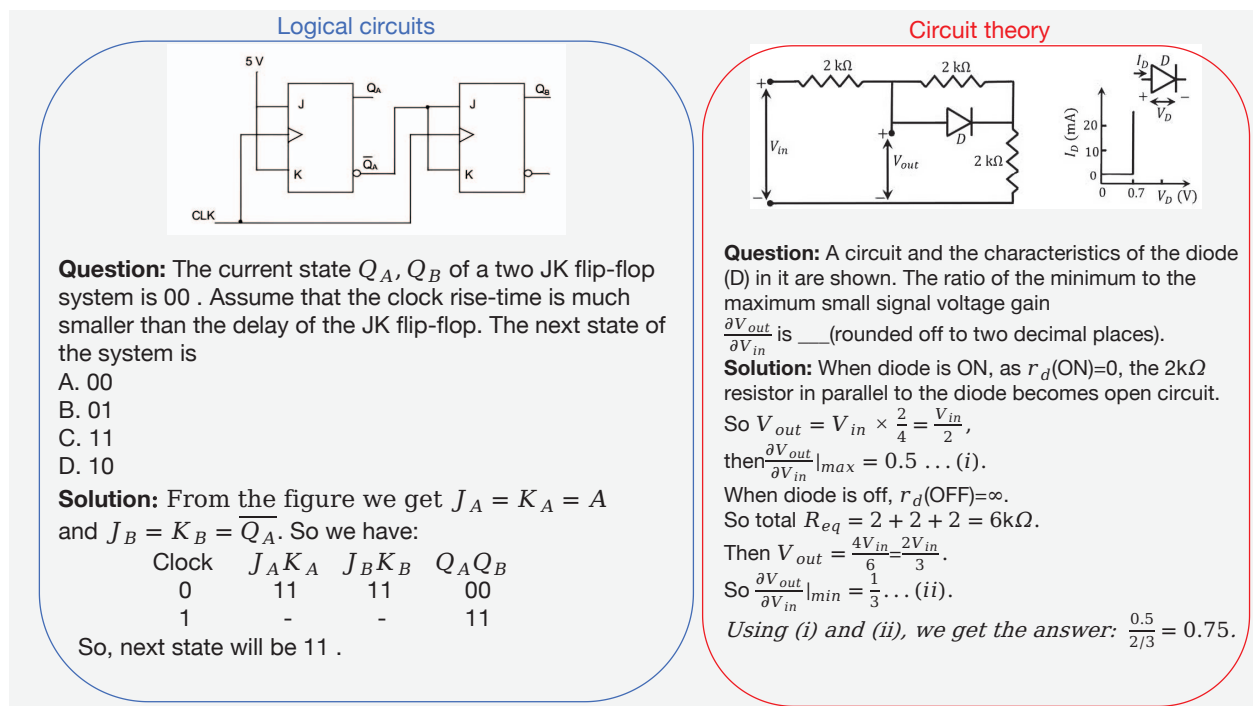*Using (i) and (ii), we get the answer:* $\frac{0.5}{2/3} = 0.75.$

Figure 18. Comparison between solutions of logical circuits and circuit theory. Logic Circuits demand a thorough comprehension of logical structuring and timing analysis, yet they do not heavily rely on computation. In contrast, problems of circuit theory involve solving complex equations and applying numerical methods to analyze circuits.

Figure 38 and Figure 39 show examples of LMM laziness phenomenon on MathVista. We can find that when there is (misleading) information for problem-solving in text, the LMMs will not try to read the relevant information in the images, even if this information is spurious. Consequently, LMMs generate wrong answers with this spurious information.
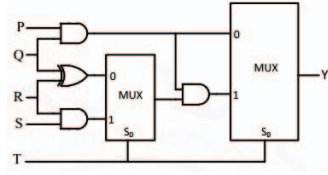
Subject: Digital Logic Circuits and Microprocessor Design
Question: The propagation delays of the XOR gate, AND gate and multiplexer (MUX) in the circuit shown in the figure
are $4\,ns$, $2\,ns$ and $1\,ns$, respectively.
If all the inputs P, Q, R, S and T are applied simultaneously and held constant, the maximum propagation delay of the
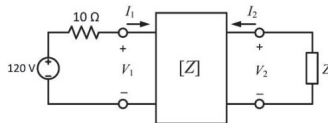circuit is ( )
(A) 3 ns (B) 5 ns (C) 6 ns (D) 7 ns
Solution: Case -1 : when T=0
Propagation delay = $t_{AND1} + t_{MUX2} = 2 + 1 = 3ns$
Case -1 : when T=1
Propagation delay = $t_{AND2} + t_{MUX1}t_{AND3} + t_{MUX2} = 2 + 1 + 2 + 1 = 6ns$.
Answer: C

Subject: Circuit Theory and Network Analysis

Question: In the given circuit shown in figure, the two-port network has the impedance matrix [Z] = $\begin{bmatrix} 40 & 60 \\ 60 & 120 \end{bmatrix}$. The

value of Z_L for which maximum power is transferred to the load is ( ) Ω.
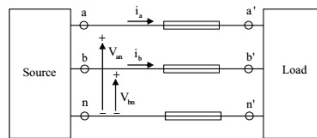A. 48 B. 38 C. 40 D. 54
Solution: From maximum power transfer theorem

$$Z_L = Z_{th}, \ Z_{th} = Z_{22} - \frac{Z_{12} \times Z_{21}}{R_s + Z_{11}}$$

For given data,

$$Z_{th} = 120 - \frac{60 \times 60}{10 + 40} = 48\Omega, \ Z_L = 48\Omega.$$

Answer: A

Subject: Power Electronics and Power Systems
Question: A source is supplying a load through a 2-phase, 3-wire transmission system as shown in the figure. The
instantaneous voltage and current in phase-a are

$$V_{an} = 220\sin(100\pi t)\,V \quad \text{and} \quad i_a = 10\sin(100\pi t)\,A,$$

respectively. Similarly, for phase-b the instantaneous voltage and current are

$$V_{bn} = 220\cos(100\pi t)\,V \quad \text{and} \quad i_b = 10\cos(100\pi t)\,A,$$

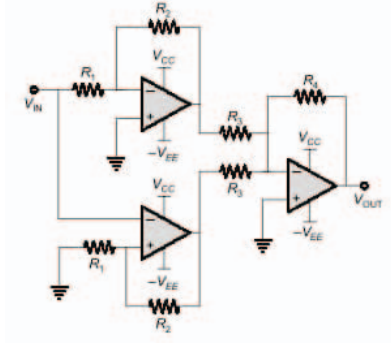respectively. The total instantaneous power flowing from the source to the load is ( ).
(A): 2200 W (B): $2200\sin^2(100\pi t)$ W (C): 440W (D): $2200\sin(100\pi t)\cos(100\pi t)W$
Solution: $V_{an} = 220\sin(100\pi t)V$, $i_a = 10\sin(100\pi t)A$,
$V_{bn} = 220\cos(100\pi t)V$, $i_b = 10\cos(100\pi t)A$
$p = V_{an}i_a + V_{bn}i_b = 2200W$
Answer: A

Figure 19. Examples of problems in EEE-Bench.

Subject: Analog Circuits

Question: The $\frac{V_{OUT}}{V_{IN}}$ of the circuit shown in figure is ( ).

$$A: -\frac{R_4}{R_3} \quad B: \frac{R_4}{R_3} \quad C: 1+\frac{R_4}{R_3} \quad D: 1-\frac{R_4}{R_3}$$

Solution: Here, A_1 is an inverting amplifier and A_2 is a non-inverting amplifier.

$$V_{01} = \frac{-R_2}{R_1}V_{in}, \ V_{02} = \left(1+\frac{R_2}{R_1}\right)V_{in}$$
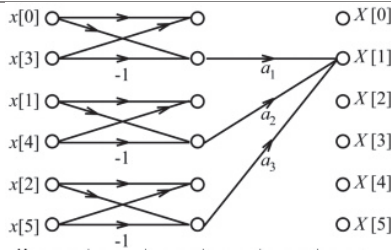
Also, A_3 is an inverting summing amplifier,

$$V_{out} = \frac{-R_4}{R_3}V_{01} + \frac{-R_4}{R_3}V_{02} = \frac{-R_4}{R_3}\left[\frac{R_2}{R_1}V_{in} + \left(1+\frac{R_2}{R_1}\right)V_{in}\right]$$

$$V_{out} = \frac{-R_4}{R_3}V_{in}$$

Gain,

$$\frac{V_{out}}{V_{in}} = \frac{-R_4}{R_3}$$

Answer: A

Subject: Signals and Systems

Question: Consider a six-point decimation-in-time Fast Fourier Transform (FFT) algorithm, for which the signal-flow graph corresponding to $X[1]$ is shown in the figure. Let $W_6 = \exp\left(-\frac{j2\pi}{6}\right)$. In the figure, what should be the values of the coefficients a_1, a_2, a_3 in terms of W_6 so that $X[1]$ is obtained correctly?

$A:$ $a_1 = -1, a_2 = W_6, a_3 = W_6^2$ $B:$ $a_1 = 1, a_2 = W_6^2, a_3 = W_6$

$C:$ $a_1 = 1, a_2 = W_6, a_3 = W_6^2$ $D:$ $a_1 = -1, a_2 = W_6^2, a_3 = W_6$

Solution: $X(k) = \sum_{n=0}^{N-1} x(n)e^{-j\frac{2\pi}{N}kn}$.

$X(1) = \sum_{n=0}^{5} x(n)W_6^n = x(0) + x(1)W_6 + x(2)W_6^2 + x(3)W_6^3 + x(4)W_6^4 + x(5)W_6^5 \ldots (i)$
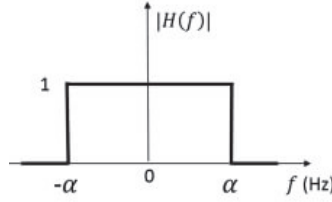
From the given flow graph, we have $X(k) = [x(0) - x(3)]a_1 + [x(1) - x(4)]a_2 + [x(2) - x(5)]a_3 \ldots (ii)$.

By comparing equations (i) and (ii), we get $a_1 = 1, \quad a_2 = W_6, \quad a_3 = W_6^2$.

Answer: C

Figure 20. Examples of problems in EEE-Bench.

Subject: Communication Systems
Question: The frequency response $H(f)$ of a linear time-invariant system has magnitude as shown in the figure.
Statement I: The system is necessarily a pure delay system for inputs which are bandlimited to $-\alpha \leq f \leq \alpha$.
Statement II: For any wide-sense stationary input process with power spectral density $S_X(f)$, the output power spectral density $S_Y(f)$ obeys $S_Y(f) = S_X(f)$ for $-\alpha \leq f \leq \alpha$. Which one of the following combinations is true?
(A): Statement I is correct, Statement II is correct (B): Statement I is correct, Statement II is incorrect
(C): Statement I is incorrect, Statement II is correct (D): Statement I is incorrect, Statement II is incorrect
Solution: For the system to be a delay system

$$y(t) = x(t - t_d), \ y(F) = e^{-j\omega t_d} \times F$$

$$\Rightarrow H(F) = \frac{Y(F)}{X(F)} = e^{-j\omega t_d}$$

Here given system is constant, hence this is not a delay system, therefore statement I is incorrect.
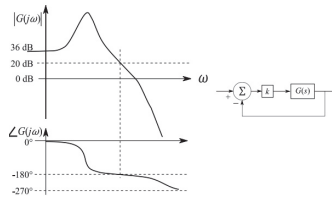
$$S_y(f) = S_x(f)|H(f)|^2$$

and $|H(f)| = 1$(given). Hence,

$$S_y(f) = S_x(f) \text{ for } -\alpha \leq f \leq \alpha$$

Statement II is correct.
Answer: C

Subject: Control Systems
Question: The figure below shows the Bode magnitude and phase plots of a stable transfer function $G(s) = \frac{n_0}{s^2 + d_2 s^2 + d_1 s + d_0}$.
Consider the negative unity feedback configuration with gain k in the feedforward path. The closed loop is stable for $k < k_o$. The maximum value of k_o is( ).
A. 0.1 B. 0.2 C. 0.3 D. 0.4
Solution: For $G(s)$

$$M_{dB}(\omega_{pc}) = 20dB$$

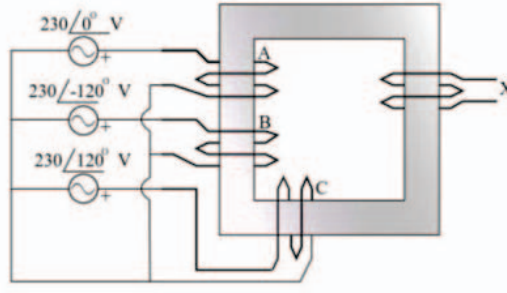When cascaded with k,

$$GM_{dB} = -20dB - 20\log_{10}(k) > 0dB$$

$$20 + 20\log_{10}(k) < 0$$

$$20\log_{10}(k) < -20$$

$$k < 10^{-1} = 0.10$$

So, $k_0 = 0.10$
Answer: A

Figure 21. Examples of problems in EEE-Bench.

Subject: Electrical Machines

Question: Windings 'A', 'B' and 'C' have 20 turns each and are wound on the same iron core as shown, along with winding 'X' which has 2 turns. The figure shows the sense (clockwise/anti-clockwise) of each of the windings only and does not reflect the exact number of turns. If windings 'A', 'B' and 'C' are supplied with balanced 3-phase voltages at 50 Hz and there is no core saturation, the no-load RMS voltage (in V, rounded off to 2 decimal places) across winding 'X' is ( ).

A. 36 B. 46 C. 12 D. 58

Solution: $V_X = \frac{2}{20}(230\angle 0° - 230\angle 120° - 230\angle -120°) = 46\angle 0°$ V

Answer: B
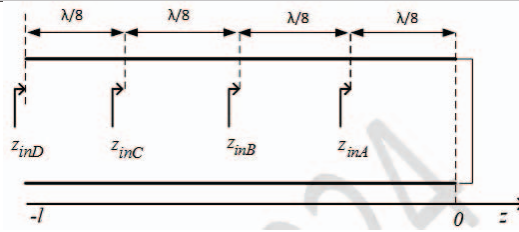
Subject: Electromagnetics

Question: Consider a lossless transmission line terminated with a short circuit as shown in the figure. As one moves towards the generator from the load, the normalized impedances $z_{\text{in }A}, z_{\text{in }B}, z_{\text{in }C},$ and $z_{\text{in }D}$ (indicated in the figure) are ( ).

A: $z_{\text{in }A} = +0.4j\,\Omega, z_{\text{in }B} = \infty, z_{\text{in }C} = 0, z_{\text{in }D} = -0.4j\,\Omega$

B: $z_{\text{in }A} = \infty, z_{\text{in }B} = +0.4j\,\Omega, z_{\text{in }C} = 0, z_{\text{in }D} = +0.4j\,\Omega$

C: $z_{\text{in }A} = -1j\,\Omega, z_{\text{in }B} = 0, z_{\text{in }C} = +1j\,\Omega, z_{\text{in }D} = \infty$

D: $z_{\text{in }A} = +1j\,\Omega, z_{\text{in }B} = \infty, z_{\text{in }C} = -1j\,\Omega, z_{\text{in }D} = 0$

Solution: $Z_{s/c} = jZ_0 \tan \beta l.$

$\Rightarrow$ Normalized impedance, $\bar{Z}_{S/C} = \frac{Z_{S/C}}{Z_0} = j \tan \beta l$

$\bar{Z}_{in\,A} = j \tan \left(\frac{2\pi}{\lambda} \cdot \frac{\lambda}{8} \right) = j \tan \left(\frac{\pi}{4} \right) = j1\,\Omega$
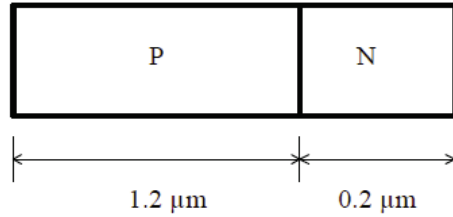
$\bar{Z}_{in\,B} = j \tan \left(\frac{2\pi}{\lambda} \cdot \frac{\lambda}{4} \right) = \infty$

$\bar{Z}_{in\,C} = j \tan \left(\frac{2\pi}{\lambda} \cdot \frac{3\lambda}{8} \right) = -j1\,\Omega$

$\bar{Z}_{in\,D} = j \tan \left(\frac{2\pi}{\lambda} \cdot \frac{\lambda}{2} \right) = 0$

Answer: D

Figure 22. Examples of problems in EEE-Bench.

| P | N |

1.2 μm   0.2 μm

Subject: Electronic Devices

Question: A silicon P-N junction is shown in the figure. The doping in the P region is $5 \times 10^{16}\, cm^{-3}$ and the doping in the N region is $10 \times 10^{16}\, cm^{-3}$. The parameters given are:

$$\text{Built-in voltage } (\Phi_{bi}) = 0.8\,V$$

$$\text{Electron charge } (q) = 1.6 \times 10^{-19}\,C$$

$$\text{Vacuum permittivity } (\varepsilon_0) = 8.85 \times 10^{-12}\,F/m$$

Relative permittivity of silicon $(\varepsilon_{si}) = 12$

The magnitude of reverse bias voltage that would completely deplete one of the two regions (P or N) prior to the other (rounded off to one decimal place) is ( ) V (round to one decimal).

Solution: Given that $N_A = 5 \times 10^{16}\,\text{cm}^{-3}$; $N_D = 10 \times 10^{16}\,\text{cm}^{-3}$

Built-in potential: $\Phi_{bi} = 0.8\,\text{V}$

Electron charge: $q = 1.6 \times 10^{-19}\,\text{C}$

Vacuum permittivity, $\varepsilon_0 = 8.85 \times 10^{-12}\,\text{F/m} = 8.85 \times 10^{-14}\,\text{F/cm}$

Relative permittivity of silicon, $\varepsilon_{si} = 12$

$\Rightarrow$ Doping on both sides is comparable, so the smaller region would deplete first.

So, depletion region width on the N-side $= x_n = 0.2\,\mu m$

$\Rightarrow x_n = 0.2 \times 10^{-4}\,cm$

$$x_n = \sqrt{\frac{2\varepsilon_{si}}{q}\left(\frac{N_A}{N_D}\right)\left(\frac{1}{N_A + N_D}\right)(\phi_{bj} + V_R)}$$

where, $V_R \to$ Magnitude of reverse bias potential

$$\Rightarrow 0.2 \times 10^{-4} = \sqrt{\frac{2 \times 12 \times 8.85 \times 10^{-14}}{1.6 \times 10^{-19}} \cdot \frac{5 \times 10^{16}}{10 \times 10^{16}} \cdot \frac{1}{(15 \times 10^{16})}(\phi_{bj} + V_R)}$$

$\Rightarrow \phi_{bj} + V_R = 9.039$

$\Rightarrow V_R = 9.039 - 0.8$

$\Rightarrow V_R = 8.239\,V \approx 8.2\,V$

Answer: 8.2

Figure 23. Examples of problems in EEE-Bench.

| A	B	C	F
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	1

Question: The logical expression corresponding to the logic state table shown in figure is ( )
A. $F = A + B + C$
B. $F = ABC$
C. $F = A + B + \overline{C}$
D. $F = AB\overline{C}$

Answer: B.
Response from GPT-4o: To determine the logical expression, we need to analyze the truth table and identify when the output F is 1.
From the table: F = 1 when A=1,B = 1,C=0. This corresponds to the logical expression$F = A \cdot B \cdot \overline{C}$.Therefore, the correct option is:
D. $F = AB\overline{C}$
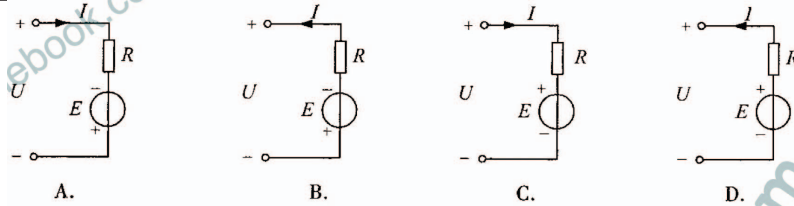
**Question:** Among the four circuits shown in the figure, the circuit that satisfies the equation $U = -IR + E$ is ( ).
A. circuit A
B. circuit B
C. circuit C
D. circuit D

Answer: C.
Response from GPT-4o: The equation $U = -IR + E$ suggests that the potential difference U is equal to the electromotive force E minus the voltage drop across the resistor IR. This implies that the current I should flow from the positive terminal of the battery through the resistor.
In circuit B, the current I flows from the positive terminal of the battery through the resistor, which matches the equation $U = -IR+E$.
Therefore, the correct answer is:
B. circuit B
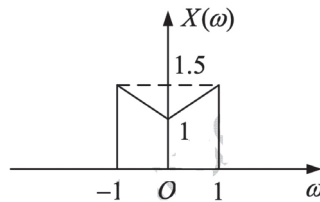
Question: Given that the Fourier transform $X(\omega)$ of the signal $x(t)$ is as shown in the figure, determine the value of $\int_{-\infty}^{+\infty} x(t)\,dt = ($
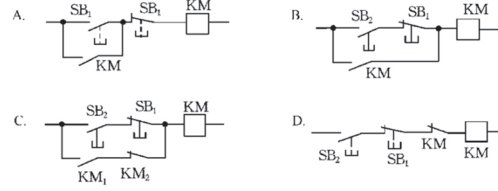).
Answer: 1.
Response from GPT-4o: To find the value of $\int_{-\infty}^{+\infty} x(t)\,dt$, we need to evaluate $X(0)$, the Fourier transform of $x(t)$ at $\omega = 0$.
From the given figure,$X(\omega)$ is a triangular function with a peak at $\omega = 0$ and a height of 1.5. Therefore, $X(0) = 1.5$.
Thus, the value of $\int_{-\infty}^{+\infty} x(t)\,dt$ is 1.5.
Final value: 1.5

Figure 24. Image Perception Error examples of the response from GPT-4o. The error part in response is marked as red color

.

Question: In the following circuits, which circuit ( ) exhibits the following phenomenon: After pressing the start button SB2, the contactor KM is engaged, but once the start button is released, the contactor KM immediately disengages?
A. graph A B. graph B
C. graph C D. graph D
Answer: C.
Response from GPT-4o: The circuit that exhibits the phenomenon where the contactor KM is engaged after pressing the start button SB2, but immediately disengages once the button is released, is circuit B. This is because circuit B lacks a holding contact to maintain the circuit after the button is released.
Correct option: B. graph B
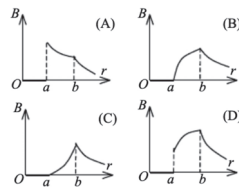
Question: The inner and outer radii of an infinitely long hollow cylindrical conductor carrying current are a and b, respectively. The current is uniformly distributed over the cross-section of the conductor. The qualitative relationship between the magnitude of the magnetic field $\vec{B}$ at various points in space and the distance r from the center axis of the cylinder is shown in the figure. The correct graph is:
A. graph A B. graph B
C. graph C D. graph D
Answer: B.
Response from GPT-4o: To determine the correct graph, we need to consider the behavior of the magnetic field $\vec{B}$ in different regions:
1. **Inside the hollow part $(r < a)$**: The magnetic field is zero because there is no current enclosed.
2. **Within the conductor $(a \leq r \leq b)$**: The magnetic field increases linearly with r because the current enclosed increases with r.
3. **Outside the conductor $(r > b)$**: The magnetic field decreases with r as the total current is enclosed and the field behaves like that of a long straight wire.
The correct graph that represents this behavior is graph D.
Answer: D
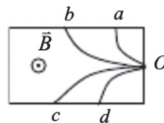
Question: The image shows the deflection trajectories of four charged particles after being injected into a uniform magnetic field at point O in the same direction, perpendicular to the magnetic field lines. The magnetic field direction is out of the plane of the paper. The masses of the four particles are equal, and their charges are also equal. The trajectory of the negatively charged particle with the greatest kinetic energy is:
(A) Oa (B) Ob
(C) Oc (D) Od
Answer: C.
Response from GPT-4o: The trajectory of a charged particle in a magnetic field is determined by the Lorentz force, which causes the particle to move in a circular path. The radius of this path is given by the formula:
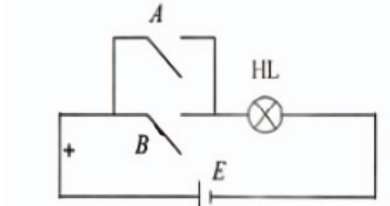
$$r = \frac{mv}{|q|B}$$

where m is the mass, v is the velocity, q is the charge, and B is the magnetic field strength. Since the masses and charges are equal, the radius is directly proportional to the velocity, and hence the kinetic energy (since $KE = \frac{1}{2}mv^2$).
The particle with the greatest kinetic energy will have the largest radius of curvature. In the image, the trajectory with the largest radius is Ob.
Therefore, the correct answer is: (B) Ob

Figure 25. Image Perception Error examples of the response from GPT-4o. The error part in response is marked as red color

Question: In the logic circuit composed of switches shown in figure, set switches A and B to "1" (connected), and set to "0" (disconnected), if the light is on "1" and off "0", then this circuit is ( )
A. "AND" gate
B. "OR" gate
C. "NOT" gate
D. "NAND
Answer: B.
Response from GPT-4o: The circuit shown is an "AND" gate. For the light to be on "1", both switches A and B must be closed (set to "1"). If either switch is open, the circuit is broken, and the light will be off ("0").Correct option: A. "AND" gate.
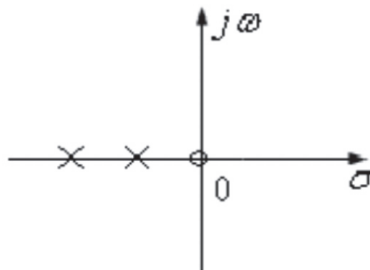

Question: If the pole-zero distribution of the system H(s) is as shown in Figure 1, determine which type of filter network it is:
A. Low-pass
B. High-pass
C. Band-pass
D. Band-stop
Answer: C.
Response from GPT-4o: The pole-zero plot shows zeros on the left side of the imaginary axis and a pole at the origin. This configuration typically represents a **high-pass filter**, as it attenuates low frequencies and allows high frequencies to pass.
Answer: B. High-pass
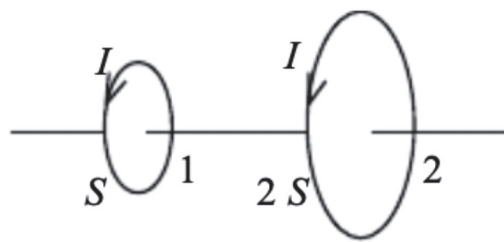
Question: Two circular coils, 1 and 2, with areas S and $2S$ respectively, are placed as shown in the image. Both coils carry the same current I. The magnetic flux through coil 2 produced by the current in coil 1 is denoted as Φ_{21}, and the magnetic flux through coil 1 produced by the current in coil 2 is denoted as Φ_{12}. The relationship between Φ_{21} and Φ_{12} is:
(A) $\Phi_{21} = 2\Phi_{12}$
(B) $\Phi_{21} > \Phi_{12}$
(C) $\Phi_{21} = \Phi_{12}$
(D) $\Phi_{21} = \frac{1}{2}\Phi_{12}$

Answer: A.
Response from GPT-4o: The mutual inductance between two coils is the same regardless of which coil is the source of the magnetic field. Therefore, the magnetic flux through coil 2 due to coil 1 (Φ_{21}) is equal to the magnetic flux through coil 1 due to coil 2 (Φ_{12}).
Thus, the relationship is:
(C) $\Phi_{21} = \Phi_{12}$

Figure 26. Reasoning Error examples of the response from GPT-4o. The error part in response is marked as red color

.
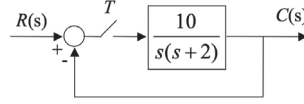
Question: Given the sampled control system as shown in figure, the open-loop pulse transfer function is $G_o(z) = \frac{5(1-e^{-2T})z}{(z-1)(z-e^{-2T})}$.

When the sampling interval $T = 0.1$ seconds, is the system stable?

A. Yes

B. No

Answer: A.

Response from GPT-4o: To determine the stability of the system, we need to examine the poles of the open-loop pulse transfer function $G_o(z)$.

The given transfer function is:

$$G_o(z) = \frac{5\left(1 - e^{-2T}\right)z}{(z-1)\left(z - e^{-2T}\right)}$$

For stability, all poles of the transfer function must lie inside the unit circle in the z-plane.
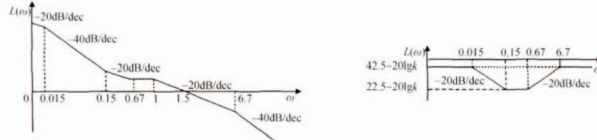
The poles are at $z = 1$ and $z = e^{-2T}$.

Given $T = 0.1$, we calculate:

$$e^{-2T} = e^{-0.2} \approx 0.8187$$

The poles are $z = 1$ and $z \approx 0.8187$.

Since the pole at $z = 1$ is on the unit circle, the system is marginally stable, not asymptotically stable.

Therefore, the system is not stable in the sense of all poles being strictly inside the unit circle. **Answer: B. No**

Question: The block diagram of a discrete system is shown in figure, where $T = 0.25$ s and the system's reference input signal is $r(t) = 8 + 0.5t$. Determine the range of values for k that will ensure the steady-state error $e_{ss} < 0.5$.

A. $1 < k < 2.67$

B. $0.5 < k < 2.5$

C. $1 < k < 2.67$

D. $0.8 < k < 2.2$

Answer: C.

Response from GPT-4o: To determine the range of values for k that will ensure the steady-state error $e_{ss} < 0.5$, we need to analyze the system's transfer function and apply the final value theorem.

The system is a discrete-time system with a transfer function given by:

$$G(s) = \frac{ke^{-0.5s}}{s}$$

The reference input signal is $r(t) = 8 + 0.5t$, which consists of a step input and a ramp input. The steady-state error for a ramp input in a type 1 system is given by:

$$e_{ss} = \frac{1}{k_v}$$

where k_v is the velocity error constant. For a type 1 system, k_v is determined by:

$$k_v = \lim_{s \to 0} sG(s) = \lim_{s \to 0} \frac{ke^{-0.5s}}{1} = k$$

Thus, the steady-state error for the ramp input is:
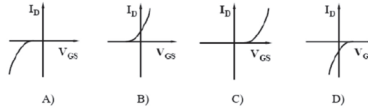
$$e_{ss} = \frac{1}{k}$$

We want $e_{ss} < 0.5$, so:

$$\frac{1}{k} < 0.5$$

Solving for k, we get: $k > 2$ Therefore, the range of values for k that ensures $e_{ss} < 0.5$ is: $k > 2$ From the given choices, the correct option is: A. $1 < k < 2.67$ So, the answer is A.

Figure 27. Reasoning Error examples of the response from GPT-4o. The error part in response is marked as red color

.

Caption context: Image caption context:

The diagram consists of four subplots labeled A, B, C, and D. Each subplot represents a graph of current versus voltage, specifically for a transistor, likely a MOSFET or JFET, based on the axes labels. Here is a detailed description of each subplot:

Subplot A: - **Axes**: - Horizontal axis: V_{GS} (Gate-Source Voltage) - Vertical axis: I_D (Drain Current) - **Graph**: - The curve starts in the third quadrant and moves into the first quadrant, indicating a negative V_{GS} initially, with I_D increasing as V_{GS} becomes positive. - The curve is nonlinear, showing a gradual increase in I_D with increasing V_{GS}.

Subplot B: - **Axes**: - Horizontal axis: V_{GS} (Gate-Source Voltage) - Vertical axis: I_D (Drain Current) - **Graph**: - The curve starts near the origin and sharply rises in the first quadrant, indicating a rapid increase in I_D with a small increase in V_{GS}. - The curve is steep, suggesting a strong dependency of I_D on V_{GS}.

Subplot C: - **Axes**: - Horizontal axis: V_{GS} (Gate-Source Voltage) - Vertical axis: I_D (Drain Current) - **Graph**: - The curve starts at the origin and rises steeply in the first quadrant, similar to subplot B, but with a more pronounced initial slope. - This indicates a very sensitive response of I_D to changes in V_{GS}.

Subplot D: - **Axes**: - Horizontal axis: V_{GS} (Gate-Source Voltage) - Vertical axis: I_D (Drain Current) - **Graph**: - The curve starts in the third quadrant and moves into the first quadrant, similar to subplot A, but with a more pronounced initial curvature. - The curve shows a nonlinear increase in I_D as V_{GS} becomes positive, indicating a threshold behavior.

General Observations: - Each subplot represents a different characteristic curve of a transistor, likely showing different modes of operation or types of transistors. - The V_{GS} and I_D relationship is crucial for understanding the behavior of the transistor in different configurations or biasing conditions. - The curves suggest different levels of sensitivity and threshold behavior, which are important for designing and analyzing circuits involving these components.

Question: The graph that correctly describes the transfer characteristics of an enhancement-mode NMOSFET (assuming current always flows into the drain as the positive direction) is ().

A. graph A B. graph B C. graph C D. graph D

Response w/o caption: The correct graph that describes the transfer characteristics of an enhancement-mode NMOSFET is graph C.

C **Response w/ caption:** The graph that correctly describes the transfer characteristics of an enhancement-mode NMOSFET is B.

In an enhancement-mode NMOSFET, the drain current I_D starts to increase significantly only after the gate-source voltage V_{GS} exceeds a certain threshold voltage. This behavior is depicted in graph B, where I_D sharply rises in the first quadrant as V_{GS} increases.
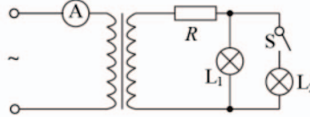
Caption context: Image caption context:

The diagram depicts an electrical circuit with the following components and connections:

1. **AC Source**: - Represented by the symbol for alternating current (AC) on the left side of the diagram. - It provides the input power to the circuit.

2. Ammeter (A): - Connected in series with the primary coil of a transformer. - Measures the current flowing through the primary side.

3. **Transformer**: - Consists of two coils: the primary coil (connected to the AC source and ammeter) and the secondary coil. - The primary coil is on the left, and the secondary coil is on the right.

4. **Resistor (R)**: - Connected in series with the secondary coil of the transformer. - Represents a load or impedance in the circuit.

5. **Lamp L1**: - Connected in parallel with the resistor (R). - Indicates the presence of current through the secondary circuit.

6. **Switch (S)**: - Connected in series with another lamp (L2). - Controls the current flow to lamp L2.

7. Lamp L2: - Connected in series with the switch (S). - Illuminates when the switch is closed, indicating current flow through this path.

8. **Connections**: - The secondary coil of the transformer is connected to both the resistor (R) and lamp L1 in parallel. - The switch (S) and lamp L2 are connected in series, forming a parallel branch with the resistor and lamp L1.

This setup can be used to study the behavior of AC circuits, the effect of transformers, and the operation of parallel and series components in electrical systems.
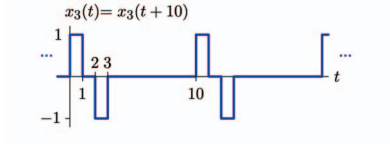
Question: As shown in the figure, the primary coil is input with a stable AC voltage, and switch S is in the open state. At a certain moment, bulb $\mathbf{L}_1$ becomes slightly dimmer. The possible reason could be:

A. The input voltage increased B. Switch $\mathbf{S}$ was closed
C. The resistor $\boldsymbol{R}$ short-circuited D. The ammeter experienced an open circuit

Response w/o caption: The possible reason bulb $\mathbf{L}_1$ becomes slightly dimmer is: B. Switch $\mathbf{S}$ was closed

When switch S is closed, current can flow through bulb $\mathbf{L}_2$, reducing the current through $\mathbf{L}_1$ and making it dimmer. **Response w/ caption:** The possible reason for bulb L_1 becoming slightly dimmer is: C. The resistor $\boldsymbol{R}$ short-circuited

If the resistor R short-circuits, it would reduce the total resistance in the parallel branch, causing more current to flow through the shorted path and less through L_1, making it dimmer.

Figure 28. Examples of LMM errors with *captions* in EEE-Bench o fGPT-4o. We mark the errors in response with red color

$$x_3(t) = x_3(t + 10)$$

caption context: Image caption context:

The diagram represents a periodic signal $x_3(t)$ with a period of 10 units. Here is a detailed description of the waveform:

1. **Signal Type**: The signal is a square wave.

2. **Periodicity**: The function is periodic with a period $T = 10$. This is indicated by the equation $x_3(t) = x_3(t + 10)$.

3. **Amplitude Levels**: The waveform alternates between two amplitude levels: - $+1$ - -1

4. **Time Intervals**: - The signal is at $+1$ from $t = 0$ to $t = 1$. - The signal drops to -1 from $t = 1$ to $t = 3$. - The signal returns to 0 from $t = 3$ to $t = 10$.

5. **Repetition**: This pattern repeats every 10 units along the time axis, both to the left and right, as indicated by the ellipses (. . .).

6. **Key Points**: - At $t = 0$, the signal jumps to $+1$. - At $t = 1$, the signal drops to -1. - At $t = 3$, the signal returns to 0. - At $t = 10$, the pattern repeats.

This description captures all the essential features of the waveform, allowing for analysis or reconstruction without the visual diagram.

Question: Determine the Fourier series coefficient c_0 for x3(t) shown in figure. c_0 = ()

Response w/o caption: To find the Fourier series coefficient c_0, which represents the average value of the function over one period, we need to integrate the function over one period and divide by the period length.

Given the periodic function $x_3(t)$ with period $T = 10$, we can calculate c_0 as follows:

$$c_0 = \frac{1}{T} \int_0^T x_3(t)\, dt$$

From the graph, the function $x_3(t)$ is:

- 1 from $t = 0$ to $t = 1$ - -1 from $t = 2$ to $t = 3$ - 0 elsewhere in the period

Now, calculate the integral:

$$c_0 = \frac{1}{10} \left(\int_0^1 1\, dt + \int_2^3 (-1)\, dt \right)$$

$$c_0 = \frac{1}{10} \left([t]_0^1 + [-t]_2^3 \right)$$

$$c_0 = \frac{1}{10} \left((1 - 0) + (-(3) - (-(2))) \right)$$

$$c_0 = \frac{1}{10} (1 - 1)$$

$$c_0 = \frac{1}{10} \times 0 = 0$$

Thus, the Fourier series coefficient c_0 is 0.0.

Response w/ spurious caption: To find the Fourier series coefficient c_0 for the signal $x_3(t)$, we need to calculate the average value of the signal over one period. The coefficient c_0 is given by:

$$c_0 = \frac{1}{T} \int_0^T x_3(t)\, dt$$

Given the period $T = 10$, we can break the integral into segments based on the time intervals provided:

1. From $t = 0$ to $t = 1$, $x_3(t) = 1$. 2. From $t = 1$ to $t = 3$, $x_3(t) = -1$. 3. From $t = 3$ to $t = 10$, $x_3(t) = 0$.

Now, calculate the integral:

$$c_0 = \frac{1}{10} \left(\int_0^1 1\, dt + \int_1^3 (-1)\, dt + \int_3^{10} 0\, dt \right)$$

$$= \frac{1}{10} \left([t]_0^1 + [-t]_1^3 + [0]_3^{10} \right)$$

$$= \frac{1}{10} \left((1 - 0) + (-(3 - 1)) + 0 \right)$$

$$= \frac{1}{10} (1 - 2)$$

$$= \frac{1}{10} \times (-1) = -0.1$$

Thus, the Fourier series coefficient c_0 is -0.1.

Figure 29. Examples of LMM errors with *captions* in EEE-Bench o f GPT-4o. We mark the errors in response with red color
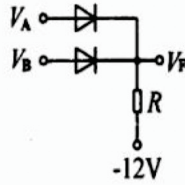
Spurious context: Image caption context:

The diagram depicts a diode AND gate circuit with a positive voltage reference. Here is a detailed description:

1. Components: - Two Diodes: The cathodes of both diodes are connected to separate input voltages. - Resistor (R): Connected in parallel with the anodes of the diodes.

2. Connections: - Input Voltage V_A: Connected to the cathode of the first diode. - Input Voltage V_B: Connected to the cathode of the second diode. - The anodes of both diodes are connected together at a junction point. - Output Voltage V_F: Taken from the junction point where the anodes of the diodes meet. - The other end of the resistor is connected to a +12V supply.

3. Voltage Reference: - The circuit is referenced to a +12V supply, which is connected to the other end of the resistor.

4. Functionality: - This configuration allows the output voltage V_F to follow the lower of the two input voltages V_A or V_B, plus the diode forward voltage drop, as long as it is below +12V.

This setup is typically used in applications where the lowest of two input voltages needs to be selected, with a positive voltage reference for biasing.

Question: In the circuit shown in figure, $V_A = 3V, V_B = 0V$, if the forward voltage drop of the diode is ignored, then V_F is ( ) A. -12 V B. 9 V C. 0 V D. 3V

Response w/o spurious caption: To find V_F, we need to consider the behavior of the diodes. Since the forward voltage drop is ignored, the diodes will conduct if the voltage at their anode is higher than at their cathode.

1. **Diode connected to V_A:** - $V_A = 3V$ - The diode will conduct because $3V > V_F$.

2. **Diode connected to V_B:** - $V_B = 0V$ - The diode will not conduct because $0V < V_F$.

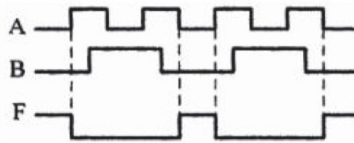Since the diode connected to V_A conducts, V_F will be equal to V_A, which is $3V$.

Therefore, $V_F = 3V$.

The correct option is D.

Response w/ spurious caption: In the given circuit, the output voltage V_F will follow the lower of the two input voltages V_A or V_B since the diodes will conduct for the lower voltage. Here, $V_A = 3V$ and $V_B = 0V$.

Since V_B is lower, V_F will be $0V$.

The correct answer is: C. 0V

Spurious context: Image caption context:

The diagram is a timing diagram showing three digital waveforms labeled A, B, and F. Here is a detailed description:

1. Waveform A: - The waveform is a sine wave with a 75% duty cycle. - It alternates between a high state and a low state. - The high state duration is longer than the low state duration. - The waveform starts with a low state.

2. Waveform B: - This waveform is also a sine wave but with a shorter high state compared to waveform A. - It starts with a high state, then transitions to a low state. - The high state duration is shorter than the low state duration.

3. Waveform F: - This waveform is a combination of waveforms A and B. - It starts with a high state. - The waveform transitions to a low state when either A or B is high. - It returns to a high state when both A and B are low.

4. Timing Relationships: - The transitions of waveforms A and B are misaligned such that waveform F is low only when both A and B are simultaneously high. - The dashed vertical lines indicate the points of transition for synchronization.

This timing diagram is typically used in digital electronics to illustrate the behavior of logic circuits.

Question: In the circuit shown in figure, if the waveform of the input terminal A and output terminal F of the NAND gate is as shown in figure, then this circuit is ( )

A. AND gate B. OR gate C. NOT gate D. NAND gate

Response w/o spurious caption: The waveform at terminal F is the inverse of the waveform at terminal A. This indicates that the circuit is functioning as a NOT gate, which inverts the input signal.
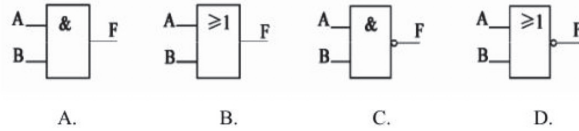
Correct option: C. NOT gate

Response w/ spurious caption: The waveform F is low only when both A and B are high, which is characteristic of a NAND gate. A NAND gate outputs low only when all its inputs are high.

Therefore, the circuit is a:

D. NAND gate

Figure 30. Examples of LMM laziness in EEE-Bench. We mark the *spurious text* in the caption and error in response with red color

A.   B.   C.   D.

Spurious context: Image caption context:

The diagram consists of four logic gate configurations labeled A, B, C, and D. Each configuration has two inputs labeled A and B, and one output labeled F. Here is a detailed description of each configuration:

A. OR Gate - **Symbol**: A rectangle with a ≥'1' inside. - **Inputs**: Two inputs labeled A and B on the left side. - **Output**: One output labeled F on the right side. - **Function**: The output F is true (1) if at least one of the inputs A or B is true (1).

B. AND Gate - **Symbol**: A rectangle with an ampersand (&) inside. - **Inputs**: Two inputs labeled A and B on the left side. - **Output**: One output labeled F on the right side. - **Function**: The output F is true (1) only if both inputs A and B are true (1).

C. NOR Gate Symbol: A rectangle with a " ≥ 1" inside, followed by a small circle (representing negation) at the output. - **Inputs**: Two inputs labeled A and B on the left side. - **Output**: One output labeled F on the right side. - **Function**: The output F is true (1) only if both inputs A and B are false (0).

D. NAND Gate - **Symbol**: A rectangle with an ampersand (&) inside, followed by a small circle (representing negation) at the output. - **Inputs**: Two inputs labeled A and B on the left side. - **Output**: One output labeled F on the right side. - **Function**: The output F is false (0) only if both inputs A and B are true (1); otherwise, it is true (1).

These configurations represent basic logic gates used in digital circuits, each performing a specific logical operation based on the inputs provided.

Question: Among the logic symbols shown in figure, the one that can realize the logic function $F = \overline{AB}$ is ( )

A. circuit A B. circuit B C. circuit C D. circuit D

Response w/o spurious caption: The logic function $F = \overline{AB}$ is a NAND operation. The circuit that represents a NAND gate is circuit C, which shows an AND gate with a NOT symbol (a small circle) at the output.

Correct option: C

Response w/ spurious caption: The logic function $F = \overline{AB}$ is realized by a NAND gate. In the diagram, the NAND gate is represented by configuration D. So, the correct option is D.
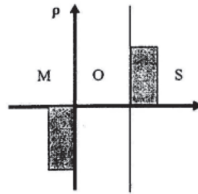
Spurious context: Image caption context:

The diagram appears to be a Bode plot, which is commonly used in control systems to analyze the frequency response of a system. Here is a detailed description:

1. Axes: - The horizontal axis is labeled as the frequency axis. - The vertical axis is labeled as the magnitude in dB.

2. Curves: - There is a single curve that shows the gain as a function of frequency.

3. Points and Regions: - The origin is marked as O. - There are two highlighted points: - One at a low frequency, labeled M. - One at a high frequency, labeled S.

4. Interpretation: - The plot is used to determine the gain and phase margins of a system. - It helps in understanding how the system will respond to different frequencies, particularly in terms of amplification and attenuation.

This description should help in analyzing the frequency response and stability of a control system using the Bode plot method.

Question: For the ideal MOS structure shown in the figure, is the semiconductor in the question N-type or P-type?

A. N-type B. P-type C. can not be determined

Response w/o spurious caption: To determine the type of semiconductor in the MOS structure, we need to look at the charge distribution. In the figure, the charge density (ρ) is positive in the semiconductor (S) region. This indicates that the semiconductor is P-type, as positive charge density in the semiconductor region is due to the presence of holes, which are the majority carriers in P-type semiconductors.

Answer: B. P-type

Response w/ spurious caption: The image provided is not a Bode plot but rather a diagram related to a MOS structure. The vertical axis is labeled ρ, which typically represents charge density.

In the context of MOS structures: - If the charge density is negative (below the horizontal axis) at point M, it indicates the presence of electrons, suggesting an N-type semiconductor. - If the charge density is positive (above the horizontal axis) at point S, it indicates the presence of holes, suggesting a P-type semiconductor.

Given the diagram, the charge density at M is negative, indicating an N-type semiconductor.

Therefore, the correct answer is: A. N-type

Figure 31. Examples of LMM laziness in EEE-Bench. We mark the *spurious text* in the caption and error in response with red color

$$x_1(t) = x_1(t + 10)$$
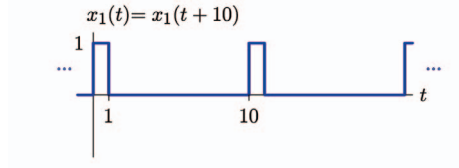
Spurious context: Image caption context:

The diagram represents a periodic rectangular pulse signal $x_1(t)$ with the following characteristics:

1. Signal Description: - The signal is periodic with a period of 20 units of time. - The mathematical expression for the signal is $x_1(t) = x_1(t + 20)$, indicating its periodic nature.

2. Pulse Characteristics: - Each pulse has an amplitude of 2. - The pulse width is 2 units of time, starting at integer multiples of 5 (e.g., 0, 5, 10, etc.).

3. Time Axis: - The horizontal axis represents time t. - The vertical axis represents the amplitude of the signal.

4. Pulse Position: - The first pulse starts at $t = 0$ and ends at t = 2. - The second pulse starts at t = 5 and ends at $t = 7$. - This pattern continues indefinitely in both directions, as indicated by the ellipses (...).

5. Signal Behavior: - The signal is zero for all time intervals except during the pulse width of 2 units at each period.

This description provides a complete understanding of the periodic rectangular pulse signal depicted in the diagram, suitable for analysis in various electrical and electronics engineering contexts.

Question: Determine the Fourier series coefficients a_0 for $x_1(t)$ shown in figure. a_0=().

Response w/o spurious caption: To find the Fourier series coefficient a_0 for the periodic function $x_1(t)$, we need to calculate the average value of the function over one period.

The function $x_1(t)$ is a periodic square wave with a period $T = 10$. It has a value of 1 from $t = 0$ to $t = 1$ and 0 from $t = 1$ to $t = 10$. The formula for a_0 is:

$$a_0 = \frac{1}{T} \int_0^T x_1(t)\, dt$$

Substituting the given values:

$$a_0 = \frac{1}{10} \left(\int_0^1 1\, dt + \int_1^{10} 0\, dt \right)$$

$$a_0 = \frac{1}{10} \left([t]_0^1 + 0 \right)$$

$$a_0 = \frac{1}{10}(1 - 0) = 0.1$$

Thus, the Fourier series coefficient a_0 is 0.1.

Response w/o spurious caption: To determine the Fourier series coefficient a_0 for the periodic rectangular pulse signal $x_1(t)$, we need to calculate the average value of the signal over one period.

The formula for a_0 is:

$$a_0 = \frac{1}{T} \int_0^T x_1(t)\, dt$$

where T is the period of the signal. For this signal, $T = 20$.

The signal $x_1(t)$ has an amplitude of 2 for a duration of 2 units of time in each period and is zero otherwise. Therefore, the integral over one period is:

$$\int_0^{20} x_1(t)\, dt = \int_0^2 2\, dt + \int_5^7 2\, dt + \int_{10}^{12} 2\, dt + \int_{15}^{17} 2\, dt$$

Calculating each integral:

$$\int_0^2 2\, dt = 2 \times 2 = 4, \quad \int_5^7 2\, dt = 2 \times 2 = 4$$

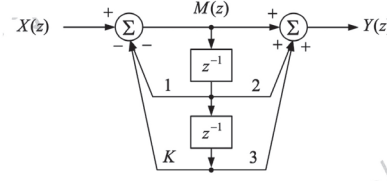$$\int_{10}^{12} 2\, dt = 2 \times 2 = 4, \quad \int_{15}^{17} 2\, dt = 2 \times 2 = 4$$

Adding these, the total integral over one period is: 4 + 4 + 4 + 4 = 16.

Now, calculate a_0:

$$a_0 = \frac{1}{20} \times 16 = 0.8$$

Thus, the Fourier series coefficient a_0 is 0.8.

Figure 32. Examples of LMM laziness in EEE-Bench. We mark the *spurious text* in the caption and error in response with red color

Spurious context: Image caption context:

The diagram is a block diagram representing a continuous-time signal processing system using the Laplace transform. Here is a detailed description:

1. Inputs and Outputs: - The input to the system is denoted as $X(s)$. - The output of the system is denoted as $Y(s)$.

2. Summing Junctions: - There are three summing junctions in the diagram. - The first summing junction has three inputs: $X(s)$ with a negative sign and a feedback path with a positive sign. - The second summing junction has two inputs, all with negative signs.

3. Paths and Connections: - The output of the first summing junction is connected to a path labeled $N(s)$. - This path leads to the second summing junction.

4. Delay Elements: - There is one delay element in the diagram, represented by a block labeled s^{-1}. - The delay element is connected in parallel with the output of the first summing junction.

5. Feedback and Feedforward Paths: - The output of the delay element is fed back to the first summing junction with a gain of K. - The output of the delay element is connected to the second summing junction with a gain of 3. - The output of the second summing junction is also connected to the second summing junction with a gain of 2.

6. Gains: - The feedback path from the output of the delay element to the first summing junction has a gain of K. - The feedforward path from the output of the delay element to the second summing junction has a gain of 3. - The feedforward path from the output of the second summing junction to the second summing junction has a gain of 2.

This block diagram represents a continuous-time system with feedback and feedforward paths, incorporating a delay element and gains, which can be analyzed using Laplace transform techniques to determine the system's behavior and response.

Question: If the system is stable when $a < K < b$, then $a + b = ($ $)$.

Response w/o spurious caption: To determine the stability of the system, we need to find the characteristic equation of the closed-loop transfer function. The system has two delay elements, each represented by z^{-1}.

The transfer function $M(z)$ can be derived from the block diagram:

1. The forward path gain is 1. 2. The feedback path includes two delay elements and a gain K.

The characteristic equation for the closed-loop system is: $1 + Kz^{-2} = 0$ Solving for z, we get:

$$Kz^{-2} = -1, z^2 = -K$$

For stability, the poles of the system must lie inside the unit circle in the z-plane. This means the magnitude of z must be less than 1: $|z| < 1$. Substituting $z^2 = -K$, we have:

$$|-K| < 1, K < 1$$

Additionally, since $z^2 = -K$, K must be positive for the system to have real poles: $K > 0$ Thus, the system is stable when $0 < K < 1$. Therefore, $a = 0$ and $b = 1$, so:

$$a + b = 0 + 1 = 1.0$$

The final value is 1.0.

Response w/ spurious caption: To determine the stability of the system, we need to analyze the characteristic equation derived from the block diagram. The system is stable if all poles of the transfer function have negative real parts.

Given the block diagram, the characteristic equation can be derived from the feedback loop. The transfer function $H(z)$ can be expressed as:

$$H(z) = \frac{Y(z)}{X(z)} = \frac{M(z)}{1 + M(z) \cdot G(z)}$$

where $M(z)$ is the forward path transfer function and $G(z)$ is the feedback path transfer function.

From the diagram: - The forward path $M(z)$ includes the two delay elements, so $M(z) = \frac{1}{z^2}$. - The feedback path $G(z)$ includes the gain K and the delay elements, so $G(z) = K \cdot \frac{1}{z^2}$. The characteristic equation is:

$$1 + M(z) \cdot G(z) = 1 + \frac{1}{z^2} \cdot K \cdot \frac{1}{z^2} = 1 + \frac{K}{z^4} = 0$$

This simplifies to: $z^4 + K = 0$. For stability, the roots of this equation must lie inside the unit circle in the z-plane. The roots are:

$$z = \sqrt[4]{-K}$$

For the system to be stable, the magnitude of z must be less than 1: $|\sqrt[4]{-K}| < 1$ This implies: $|K| < 1$ Thus, the system is stable when $-1 < K < 1$. Therefore, $a = -1$ and $b = 1$, so $a + b = 0$. The final value is 0.0.

Figure 33. Examples of LMM laziness in EEE-Bench. We mark the *spurious text* in the caption and error in response with red color
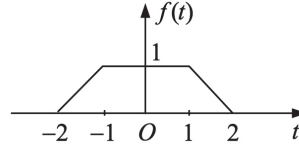
Spurious context: Image caption context:

The diagram represents a sinusoidal function $f(t)$ plotted against time t. Here is a detailed description:

1. Axes: - The horizontal axis is labeled t, representing time. - The vertical axis is labeled $f(t)$, representing the function value.

2. Function Description: - The function $f(t)$ is defined over the interval from $t = -2$ to $t = 2$. - The function is zero for $t < -2$ and $t > 2$.

3. Piecewise Segments: - **From $t = -2$ to $t = -1$**: - The function linearly decreases from 1 to 0. - **From $t = -1$ to $t = 1$**: - The function remains constant at 0. - **From $t = 1$ to $t = 2$**: - The function linearly increases from 0 to 1.

4. Key Points: - At $t = -2$, $f(t) = 1$. - At $t = -1$, $f(t) = 0$. - At $t = 0$, $f(t) = 0$. - At $t = 1$, $f(t) = 0$. - At $t = 2$, $f(t) = 1$.

This function is a sinusoidal waveform, often used in signal processing and control systems to represent a pulse with linear rise and fall times.

Question: The signal $f(t)$ is as shown in the figure, and the frequency spectrum function $F(j\omega)$ is equal to:

A. $\text{Sa}^2\left(\frac{\omega}{2}\right)[1 + 2\cos\omega]$ B. $\text{Sa}^2\left(\frac{\omega}{2}\right)[1 + 2\sin\omega]$ C. $\text{Sa}^2\left(\frac{\omega}{2}\right)[1 + \cos\omega]$ D. $\text{Sa}^2\left(\frac{\omega}{2}\right)[1 + 2\cos\frac{\omega}{2}]$

Response w/o spurious caption: To find the frequency spectrum $F(j\omega)$ of the given signal $f(t)$, we need to analyze the shape of the signal. The signal is a trapezoidal pulse, which can be decomposed into a rectangular pulse and two triangular pulses.

1. **Rectangular Pulse**: The central part from -1 to 1 is a rectangular pulse of width 2 and height 1. The Fourier transform of a rectangular pulse of width T and height A is $A \cdot T \cdot \text{Sa}(\omega T/2)$.

2. **Triangular Pulses**: The two triangular parts on either side can be considered as the difference between two shifted rectangular pulses. The Fourier transform of a triangular pulse is the square of the sinc function.

The overall Fourier transform will be the sum of these components. The sinc function is defined as $\text{Sa}(x) = \frac{\sin(\pi x)}{\pi x}$.

Given the symmetry and the shape of the signal, the Fourier transform will involve the square of the sinc function and a cosine term due to the symmetry about the origin.

The correct choice is: A. $\text{Sa}^2\left(\frac{\omega}{2}\right)[1 + 2\cos\omega]$

Response w/ spurious caption: The function $f(t)$ is a triangular pulse, which can be represented as the convolution of two rectangular pulses. The Fourier transform of a triangular pulse is the square of the sinc function, often denoted as $\text{Sa}^2\left(\frac{\omega}{2}\right)$.

Given the symmetry and the linear rise and fall, the frequency spectrum function $F(j\omega)$ is:

$$F(j\omega) = \text{Sa}^2\left(\frac{\omega}{2}\right)\left[1 + 2\cos\frac{\omega}{2}\right]$$

Thus, the correct choice is: **D.** $\text{Sa}^2\left(\frac{\omega}{2}\right)\left[1 + 2\cos\frac{\omega}{2}\right]$
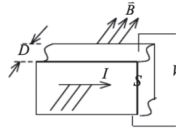
Spurious context: Image caption context:

The diagram appears to represent a section of a simple electrical circuit involving basic resistive components. Here is a detailed description of the components and parameters:

1. Rectangular Block: - The main component is a rectangular block, which could represent a resistor or a non-conductive barrier. 2. Magnetic Flux Density (**B**): - Arrows pointing into the bottom surface of the block indicate the direction of the magnetic flux density (**B**), suggesting that the block is outside of any magnetic field influence. 3. Electric Current (I): - An arrow labeled I is shown outside the block, indicating the direction of electric current flow around the block, not through it. 4. Electric Displacement Field (**D**): - Arrows labeled **D** are shown exiting the right side of the block, indicating the direction of the electric displacement field. This suggests the absence of an electric field within the block. 5. Voltage (V): - A curly bracket on the left side of the block is labeled V, indicating the potential difference or voltage along the side of the block. 6. Surface (S): - The bottom surface of the block is labeled S, which could represent the area unrelated to the magnetic flux density.

This diagram likely represents a scenario involving basic electrical principles, such as a simple resistor circuit, where electric fields and currents are minimally interacting with magnetic fields. The parameters **B**, **D**, I, V, and S are used to analyze the basic electrical behavior of the system.

Question: A conductor carrying a current I has a thickness D and a cross-sectional area S. It is placed in a uniform magnetic field with a magnetic flux density B, where the magnetic field direction is perpendicular to the side surface of the conductor, as shown in the figure. The potential difference between the top and bottom surfaces of the conductor is measured to be V. The Hall coefficient of this conductor is:

Choices: (A) $\frac{VDS}{IB}$ (B) $\frac{IBV}{DS}$ (C) $\frac{VS}{IBD}$ (D) $\frac{IVS}{BD}$ (E) $\frac{VD}{IB}$

Response w/o spurious caption: The Hall coefficient R_H is given by the formula: $R_H = \frac{VD}{IB}$ Thus, the correct option is: (E) $\frac{VD}{IB}$

Response w/ spurious caption: The Hall coefficient (R_H) is given by the formula: $R_H = \frac{V}{IBD}$ This formula relates the Hall voltage (V), current (I), magnetic flux density (B), and thickness (D) of the conductor.

Therefore, the correct choice is: (C) $\frac{VS}{IBD}$

Figure 34. Examples of LMM laziness in EEE-Bench. We mark the *spurious text* in the caption and error in response with red color
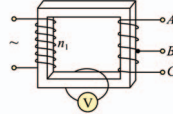
Spurious context: Image caption context:

The diagram depicts a transformer with a voltmeter connected to it. Here is a detailed description:

1. Transformer Core: - The transformer has a circular core, typically made of solid steel or another non-magnetic material.

2. Primary Winding: - The primary winding is on the right side of the core. - It is connected to a DC voltage source, indicated by a plus and minus symbol. - The number of turns in the primary winding is denoted as n_2.

3. Secondary Winding: - The secondary winding is on the left side of the core. - It has two terminals labeled A and B. - Terminal C is a tap on the winding, indicating that the secondary winding is tapped to provide different current levels.

4. Connections: - The primary winding is connected to a DC source. - The secondary winding has two output terminals (A, B) for different current outputs.

5. Voltmeter: - A voltmeter is connected across the primary winding. - The voltmeter is shown in a rectangular symbol with the letter "A" inside.

This setup is typical for a step-up transformer with a tapped primary winding, allowing for multiple current outputs. The voltmeter measures the current across the selected terminals of the primary winding.

Question: The transformer shown in the figure has an input voltage of 220 V and can output voltages of 12 V, 18 V, and 30 V. The primary coil with n_1 turns has a voltage varying with time as $\mu = U_\mathrm{m} \cos(100\pi t)$. A single-turn coil wound around the iron core is connected to an AC voltmeter, which shows a reading of 0.1 V. When a resistor R with a resistance of 12 Ω is connected across points B and C, the power is 12 W. Which of the following statements is correct?

A. n_1 is 1100 turns, and U_m is 220 V

B. The number of turns in the coil between B and C is 120, and the current through R is 1.4 A

C. If R is connected across A and B, the voltage across R is 18 V, with a frequency of 100 Hz

D. If R is connected across A and C, the current through R is 2.5 A, with a period of 0.02 s

Response w/o spurious caption: To solve this problem, we need to analyze the transformer and the given conditions.

1. Single-turn coil reading: The voltmeter connected to a single-turn coil reads 0.1 V. This means the induced EMF in one turn is 0.1 V.

2. **Power across R between B and C**: The power is given as 12 W, and the resistance R is 12 Ω. Using the power formula $P = I^2 R$, we can find the current I: $12 = I^2 \times 12 \implies I = 1\,\mathrm{A}$. The voltage across R is $V = I \times R = 1 \times 12 = 12$ V.

3. **Number of turns between B and C**: Since the voltage across B and C is 12 V and the voltage per turn is 0.1 V, the number of turns is: $\frac{12}{0.1} = 120$ turns. 4. **Frequency and period**: The input voltage varies as $\mu = U_\mathrm{m} \cos(100\pi t)$, which indicates a frequency of $f = \frac{100\pi}{2\pi} = 50\,\mathrm{Hz}$. The period T is $\frac{1}{f} = 0.02\,\mathrm{s}$.

Now, let's evaluate the options:

- **A**: n_1 is 1100 turns, and U_m is 220 V. This option does not provide enough information to verify.

- **B**: The number of turns in the coil between B and C is 120, and the current through R is 1.4 A. The number of turns is correct, but the current is 1 A, not 1.4 A.

- **C**: If R is connected across A and B, the voltage across R is 18 V, with a frequency of 100 Hz. The frequency is 50 Hz, not 100 Hz.

- **D**: If R is connected across A and C, the current through R is 2.5 A, with a period of 0.02 s. The period is correct, but we need to verify the current.

Since none of the options are entirely correct, let's verify the current for option D:

- Voltage across A and C is 30 V (given). - Current $I = \frac{V}{R} = \frac{30}{12} = 2.5\,\mathrm{A}$.

Thus, option D is correct regarding the current and period. **Correct Answer: D** **Response w/ spurious caption:** To solve this problem, we need to analyze the given information and apply transformer principles.

1. **Transformer Basics**: - The voltage ratio in a transformer is given by $\frac{V_1}{V_2} = \frac{n_1}{n_2}$. - Power in a resistor is given by $P = I^2 R$.

2. **Given Information**: - Input voltage $V_1 = 220\,\mathrm{V}$. - Output voltages: 12 V, 18 V, 30 V. - Voltmeter reading: 0.1 V. - Resistor $R = 12\,\Omega$. - Power across B and C is 12 W.

3. **Analysis**: - **Power Calculation**:

$$P = I^2 R \implies 12 = I^2 \times 12 \implies I = 1\,\mathrm{A}$$

- **Voltage across B and C**:

$$V = I \times R = 1 \times 12 = 12\,\mathrm{V}$$

- **Turns Ratio**: - For 12 V output, the turns ratio $\frac{n_1}{n_2} = \frac{220}{12}$.

4. **Checking Options**: - **A**: n_1 is 1100 turns, and U_m is 220 V. - This option does not provide enough information to verify without additional calculations. - **B**: The number of turns in the coil between B and C is 120, and the current through R is 1.4 A. - Current through R is 1 A, not 1.4 A. - **C**: If R is connected across A and B, the voltage across R is 18 V, with a frequency of 100 Hz. - The frequency is correct (100 Hz), but the voltage needs verification. - **D**: If R is connected across A and C, the current through R is 2.5 A, with a period of 0.02 s. - The period of 0.02 s corresponds to 50 Hz, not 100 Hz. **Correct Option**: C

Figure 35. Examples of LMM laziness in EEE-Bench. We mark the *spurious text* in the caption and error in response with red color
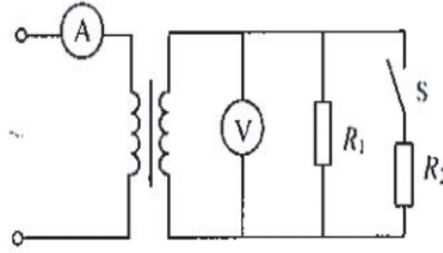
Spurious context: Image caption context:

The diagram represents an electrical circuit with the following components and connections:

1. Ammeter (A): - Positioned at the end of the circuit. - Measures the voltage across the circuit.

2. Transformer: - Consists of two coils (primary and secondary) without a core symbol between them. - The primary coil is connected in parallel with the ammeter. - The secondary coil is connected to the rest of the circuit.

3. Voltmeter (V): - Connected in series with the secondary coil of the transformer. - Measures the current through the secondary coil.

4. Resistor R_1: - Connected in series with the voltmeter. - Part of the load on the primary side of the transformer.

5. Switch (S): - Connected in parallel with resistor R_2. - Controls the disconnection of R_2 from the circuit.

6. Resistor R_2: - Connected in parallel with the switch. - When the switch is closed, R_2 is removed from the circuit.

Connections: - The primary side of the transformer is connected in parallel with the ammeter. - The secondary side of the transformer is connected in series to the voltmeter and resistor R_1. - The switch S and resistor R_2 are connected in parallel, forming a series branch with R_1 when the switch is closed.

This setup can be used to measure the voltage and current in a transformer circuit, and to analyze the effects of removing or adding a series load (resistor R_2) on the circuit.

Question: In the circuit shown in figure, an ideal transformer has its primary coil connected to a high voltage and its secondary coil connected to a load, with the line resistance being negligible. When S is initially closed and $R_1 = R_2$, and then S is opened, the AC voltage reading U, AC current reading I, and power P_1 on the load R_1 will

A. U increases, I increases, P_1 increases

B. U remains unchanged, I decreases, P_1 increases

C. U decreases, I decreases, P_1 decreases

D. U remains unchanged, I decreases, P_1 remains unchanged

Response w/o spurious caption: When the switch S is initially closed, the total resistance in the secondary circuit is $R_1 + R_2$. When S is opened, the resistance becomes just R_1.

Since $R_1 = R_2$, opening the switch doubles the resistance in the secondary circuit.

1. **Voltage U:** The voltage across the load R_1 remains unchanged because the transformer maintains the same voltage ratio regardless of the load.

2. **Current I:** The current decreases because the total resistance in the circuit increases when S is opened.

3. **Power P_1:** The power $P_1 = U^2/R_1$ remains unchanged because the voltage across R_1 remains the same and R_1 itself does not change.

Therefore, the correct answer is:

D. U remains unchanged, I decreases, P_1 remains unchanged

Response w/ spurious caption: In the given circuit, when the switch S is initially closed, resistors R_1 and R_2 are in parallel, effectively reducing the total resistance on the secondary side. When S is opened, R_2 is removed, leaving only R_1 in the circuit, which increases the total resistance.

1. **Voltage U:** The voltage across the secondary coil remains unchanged because the transformer maintains a constant voltage ratio between the primary and secondary coils.

2. **Current I:** With the increase in resistance (since R_2 is removed), the current through the circuit decreases.

3. **Power P_1:** Power is given by $P = I^2 R$. Since the current decreases and the resistance remains constant at R_1, the power P_1 decreases.

Therefore, the correct choice is:

C. U decreases, I decreases, P_1 decreases

Figure 36. Examples of LMM laziness in EEE-Bench. We mark the *spurious text* in the caption and error in response with red color
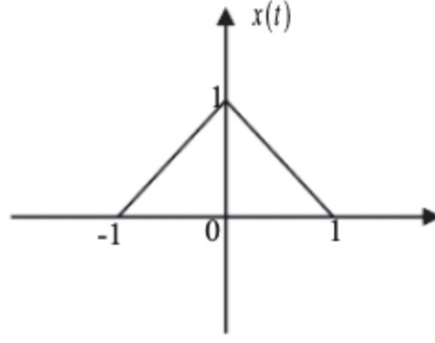
Spurious context: Image caption context:

The diagram depicts a sinusoidal waveform centered on the time axis, which is commonly used in signal processing and control systems. Here is a detailed description:

1. Axes: - The horizontal axis represents time, labeled as t. - The vertical axis represents the amplitude of the signal, labeled as $x(t)$.

2. Waveform Shape: - The waveform is a symmetric sine wave centered at the origin.

3. Key Points: - The sine wave peaks at $t = 0$ with an amplitude of 1. - The base of the sine wave extends from $t = -1$ to $t = 1$.

4. Slopes: - From $t = -1$ to $t = 0$, the waveform smoothly increases from 0 to 1. - From $t = 0$ to $t = 1$, the waveform smoothly decreases from 1 to 0.

5. Mathematical Representation: - For $-1 \leq t < 0$, $x(t) = \sin(\pi t + \pi)$. - For $0 \leq t \leq 1$, $x(t) = \sin(\pi t)$.

This sinusoidal waveform is often used in applications involving signal modulation, filtering, and analysis in electrical and electronics engineering.

Question: Given that the Fourier transform of $x(t)$ is $X(j\omega)$, and the waveform of $x(t)$ is as shown in the figure, determine $X(0) = ($).

Response w/o spurious caption: To find $X(0)$, which is the Fourier transform of $x(t)$ evaluated at $\omega = 0$, we need to calculate the area under the curve of $x(t)$. This is because $X(0)$ represents the integral of $x(t)$ over all time:

$$X(0) = \int_{-\infty}^{\infty} x(t)\, dt$$

The waveform shown is a triangle with a base from -1 to 1 and a height of 1. The area A of a triangle is given by:

$$A = \frac{1}{2} \times \text{base} \times \text{height}$$

For this triangle:
- Base = $1 - (-1) = 2$ - Height = 1
Thus, the area is:

$$A = \frac{1}{2} \times 2 \times 1 = 1$$

Therefore, $X(0) = 1.0$.

Response w/ spurious caption: To find $X(0)$, which is the Fourier transform of $x(t)$ evaluated at zero frequency, we need to calculate the area under the curve of $x(t)$ over its period. This is equivalent to finding the integral of $x(t)$ over its domain.

Given the piecewise function: - For $-1 \leq t < 0$, $x(t) = \sin(\pi t + \pi) = -\sin(\pi t)$. - For $0 \leq t \leq 1$, $x(t) = \sin(\pi t)$.

The integral of $x(t)$ from -1 to 1 is:

$$X(0) = \int_{-1}^{0} -\sin(\pi t)\, dt + \int_{0}^{1} \sin(\pi t)\, dt$$

Calculating each integral:

1. $\int_{-1}^{0} -\sin(\pi t)\, dt = \left[\frac{\cos(\pi t)}{\pi}\right]_{-1}^{0} = \frac{\cos(0) - \cos(-\pi)}{\pi} = \frac{1 - (-1)}{\pi} = \frac{2}{\pi}$

2. $\int_{0}^{1} \sin(\pi t)\, dt = \left[-\frac{\cos(\pi t)}{\pi}\right]_{0}^{1} = \frac{-\cos(\pi) + \cos(0)}{\pi} = \frac{-(-1) + 1}{\pi} = \frac{2}{\pi}$

Adding these results:

$$X(0) = \frac{2}{\pi} + \frac{2}{\pi} = \frac{4}{\pi}$$

Thus, the value of $X(0)$ is approximately 1.3.

Figure 37. Examples of LMM laziness in EEE-Bench. We mark the *spurious text* in the caption and error in response with red color

oval beads	$2/kilogram
rectangular beads	$3/kilogram
star-shaped beads	$2/kilogram
spherical beads	$2/kilogram
heart-shaped beads	$3/kilogram
square beads	$2/kilogram
flower-shaped beads	$2/kilogram

Spurious context: Image caption context:

The diagram is a table listing different types of beads along with their prices per kilogram. The table has two columns: the first column lists the bead shapes, and the second column lists the corresponding prices. The table is organized as follows:

1. Oval beads - priced at $3 per kilogram. 2. Rectangular beads - priced at $2 per kilogram. 3. Star-shaped beads - priced at $3 per kilogram. 4. Spherical beads - priced at $3 per kilogram. 5. Heart-shaped beads - priced at $2 per kilogram. 6. Square beads - priced at $3 per kilogram. 7. Flower-shaped beads - priced at $3 per kilogram.

The table is presented with a light blue background, and each row alternates in a slightly darker shade of blue for clarity. This table can be used for calculations involving the cost of purchasing different types of beads by weight.

Question: Chase wants to buy 4 kilograms of oval beads and 5 kilograms of star-shaped beads. How much will he spend? (Unit: $)

Response w/o spurious caption: Chase wants to buy 4 kilograms of oval beads and 5 kilograms of star-shaped beads.

- Cost of oval beads: 4 kg × $2/kg = $8 - Cost of star-shaped beads: 5 kg × $2/kg = $10

Total cost: $8 + $10 = $18

18

Response w/ spurious caption: To calculate the total cost:

1. **Oval beads**: 4 kg at $3 per kg = 4 * 3 = $12 2. **Star-shaped beads**: 5 kg at $3 per kg = 5 * 3 = $15

Total cost = $12 + $15 = $27

27

Houses per block

Stem	Leaf
0	5 7
1	0 4 5
2	0 0 1 2 3 3 4 6 7 8 9
3	0 1 2 4 6 8

Spurious context: Image caption context:

The diagram is a stem-and-leaf plot titled "Houses per block." It is used to display quantitative data in a compact form, showing the distribution of values.

- Title: Houses per block
- Structure: The plot is divided into two columns: "Stem" and "Leaf."
- Stem Column: - The stem represents the hundreds digit of the data values. - Values in the stem column are: 0, 1, 2, 3.
- Leaf Column: - The leaf represents the tens digit of the data values. - For each stem value, the corresponding leaf values are listed in descending order.
- Data: - Stem 0: Leaves are 7, 5. This represents the numbers 70 and 50. - Stem 1: Leaves are 5, 4, 0. This represents the numbers 150, 140, and 100. - Stem 2: Leaves are 9, 8, 7, 6, 4, 3, 3, 2, 1, 0, 0. This represents the numbers 290, 280, 270, 260, 240, 230, 230, 220, 210, 200, and 200. - Stem 3: Leaves are 8, 6, 4, 2, 1, 0. This represents the numbers 380, 360, 340, 320, 310, and 300.

This stem-and-leaf plot provides a clear visual representation of the distribution of the number of houses per block, allowing for easy identification of patterns, such as clustering and spread of the data.

Question: A real estate agent drove around the neighborhood and counted the number of houses on each block. How many blocks have exactly 36 houses? (Unit: blocks)

Response w/o spurious caption: To find the number of blocks with exactly 36 houses, look at the stem "3" and the leaf "6" in the stem-and-leaf plot.

There is 1 block with exactly 36 houses.

Final value: 1

Response w/ spurious caption: The stem-and-leaf plot does not show any blocks with exactly 36 houses. The closest values in the plot are 30 and 40, but not 36.

Therefore, the number of blocks with exactly 36 houses is 0.

Figure 38. Examples of LMM laziness in MathVista. We mark the *spurious text* in the caption and error in response with red color
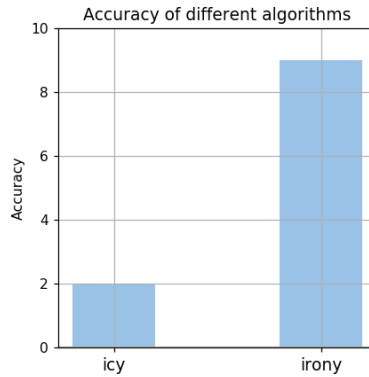
Spurious context: Image caption context:

The diagram is a bar chart titled "Accuracy of different algorithms." It displays the accuracy of two algorithms, labeled "icy" and "irony," on the x-axis. The y-axis is labeled "Accuracy" and ranges from 0 to 100, with grid lines at intervals of 20 units.

- The bar for "icy" is light blue and reaches an accuracy level of 20. - The bar for "irony" is also light blue and reaches an accuracy level of 80.

The chart is designed to compare the performance of the two algorithms based on their accuracy.

Question: How many algorithms have accuracies higher than 9?

Response w/o spurious caption: The number of algorithms with accuracies higher than 9 is 0.

Response w/ spurious caption: 2
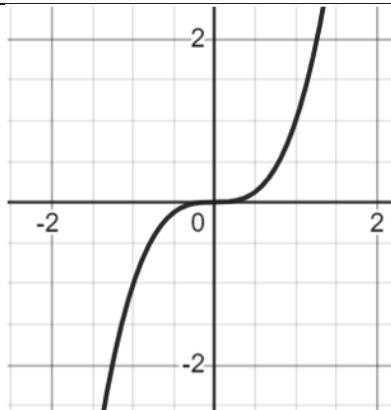
Spurious context: Image caption context:

The diagram is a graph of a mathematical function plotted on a Cartesian coordinate system. Here is a detailed description:

1. Axes: - The graph has a horizontal x-axis and a vertical y-axis, both intersecting at the origin (0,0). - The axes are labeled with integer values at regular intervals: -3, 0, and 3 on both axes.

2. Grid: - The background consists of a grid with squares, each representing two units on both the x and y axes.

3. Function: - The graph depicts an S-shaped curve, characteristic of a hyperbolic sine function or a logistic function. - The curve passes through the origin (0,0), indicating it is symmetric about this point.

4. Behavior: - As x approaches negative infinity, the curve approaches y = -3, suggesting a horizontal asymptote. - As x approaches positive infinity, the curve approaches y = 3, indicating another horizontal asymptote. - The curve is steepest around the origin, where it transitions from negative to positive values.

5. Quadrants: - The curve starts in the second quadrant, passes through the origin, and continues into the fourth quadrant.

This graph is likely representing a function with a sigmoid shape, commonly used in statistics, neural networks, and logistic growth models.

Question: What is the degree of this function?

Response w/o spurious caption: The graph shown is that of a cubic function, which typically has the form $f(x) = ax^3 + bx^2 + cx + d$. The characteristic S-shape indicates that the degree of the function is 3.

3

Response w/ spurious caption: The graph represents a sigmoid function, which is not a polynomial function. Therefore, it does not have a degree in the traditional sense used for polynomials. The degree of this function is 0.

Figure 39. Examples of LMM laziness in MathVista. We mark the *spurious text* in the caption and error in response with red color

A	B	C	F
0	0	0	1
0	0	1	1
0	1	0	1
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	0

Question: In the circuit shown in figure, the corresponding logical expression is ( ).
A. $F = A + B + C$
B. $F = A \cdot B \cdot C$
C. $F = \overline{A + B + C}$
D. $F = \overline{A \cdot B \cdot C}$
Answer: D

GPT4-o: The truth table shows that the output F is 1 for all combinations of A, B, C except when all are 1. This corresponds to the logical expression for a NOR gate, which is the negation of the AND operation. The correct logical expression is:

$$F = \overline{A \cdot B \cdot C}$$

So, the correct option is: D. $F = \overline{A \cdot B \cdot C}$

Gemini-Pro 1.5: The truth table shows that F is 0 only when A, B, and C are all 1. This corresponds to a NAND gate. Therefore, the correct logical expression is:
D. $F = \overline{A \cdot B \cdot C}$

Qwen-VL-Max: The logical expression for the circuit shown in the image is $F = A + B + C$, which represents the sum of three inputs A, B, and C. Therefore, the correct choice is:
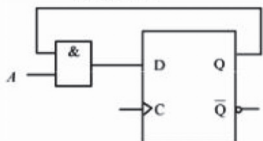A. $F = A + B + C$

InternVL2-26B: The given circuit is a combination of AND gates. The output F is 1 only when all the inputs A, B, and C are 1. Therefore, the logical expression for F is:

$$F = A \cdot B \cdot C$$

So, the correct option is B.

MiniCPM-V-2-6: Answer: D

LLaVA-v1.6-7b: The correct option letter is C. The logical expression for the circuit shown in the figure is the complement of the sum of A, B, and C.

Question: The logical circuit shown in figure, when A = "0", after the arrival of C pulse, the D flip-flop ( )
A. Has counter function B. Set to "0"
C. Set to "1" D. Uncertain Answer is B
Answer: B

GPT4-o: When A = "0", the output of the AND gate will be "0" regardless of the other input. Therefore, the D input of the flip-flop will be "0". After the arrival of the C pulse (clock pulse), the D flip-flop will be set to "0".
Answer: B

Gemini-Pro 1.5: When A = "0", the output of the AND gate will be "0" regardless of the other input. Therefore, the D input of the flip-flop will be "0". After the arrival of the C pulse (clock pulse), the D flip-flop will be set to "0".
Answer: B

Qwen-VL-Max: When A = "0", the AND gate output will be "0" regardless of the value of B. When the C pulse arrives, it will pass through the OR gate and trigger the D flip-flop. Since the input to the D flip-flop is "0" due to the AND gate output, the D flip-flop will set to "0". Therefore, the correct answer is:
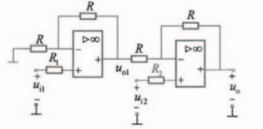B. Set to "0"

InternVL2-26B: B. Set to "0"

MiniCPM-V-2-6: Answer: C

LLaVA-v1.6-7b: The correct answer is B. The D flip-flop is set to "0" when A = "0" and after the arrival of C pulse.

Figure 40. Examples of responses among 6 baselines in EEE-Bench. The error part os each response is marked as red.

Question: In the circuit shown in figure, it is known that $u_{i1} = 2V$, $u_{i2} = -2V$. The output voltage u_o is ( )V.

Answer: -8.

GPT4-o: To solve this problem, we need to analyze the circuit, which consists of two operational amplifiers (op-amps) in an inverting configuration.

First Op-Amp: - The first op-amp is in an inverting configuration. - The input voltage $u_{i1} = 2V$. -The resistors R_1 and R form the feedback network.

The voltage at the output of the first op-amp, u_{o1}, is given by the inverting amplifier formula: $u_{o1} = -\left(\frac{R}{R_1}\right)u_{i1}$ Assuming $R = R_1$, we have: $u_{o1} = -2V$

Second Op-Amp: - The second op-amp is also in an inverting configuration. - The input voltage $u_{i2} = -2V$. - The resistors R_2 and R form the feedback network.

The voltage at the output of the second op-amp, u_o, is given by: $u_o = -\left(\frac{R}{R_2}\right)(u_{o1} + u_{i2})$ Assuming $R = R_2$, we have:

$$u_o = -(u_{o1} + u_{i2}) = -(-2V - 2V) = 4V$$

Therefore, the output voltage u_o is 4.0 volts.

Gemini-Pro 1.5: The circuit shown is a two-stage inverting amplifier. The gain of each stage is determined by the ratio of the feedback resistor to the input resistor. Since the open-loop gain of the op-amps is infinite, the voltage at the inverting input of each op-amp is equal to the voltage at the non-inverting input, which is 0V.

For the first stage, the output voltage u_{o1} can be calculated as:
$u_{o1} = -\frac{R}{R}u_{i1} = -u_{i1} = -2V$
For the second stage, the output voltage u_o can be calculated as:
$u_o = -\frac{R}{R}u_{o1} = -u_{o1} = 2V$
Therefore, the output voltage u_o is **2.0**.

Qwen-VL-Max: The circuit shown in the image is a simple voltage divider circuit. To find the output voltage u_o, we need to calculate the voltage across the resistor R_2 using Ohm's law.

Ohm's law states that $V = IR$, where V is the voltage, I is the current, and R is the resistance.

In this case, the voltage across R_2 can be calculated as follows: $V_{R2} = I * R_2$ First, let's find the total resistance of the circuit (R_T): $R_T = R + R_1 + R_2$ Then, we can find the current through the circuit (I) by dividing the input voltage (u_{i1}) by the total resistance: $I = \frac{u_{i1}}{R_T}$ Finally, we can calculate the voltage across R_2 (V_{R2}):
$V_{R2} = I * R_2$
Since $u_i1 = 2V$ and $u_i2 = -2V$, the net input voltage is zero. Therefore, the output voltage u_o will also be zero.

So, the output voltage u_o is 0.0 V.

InternVL2-26B: To determine the output voltage u_o in the given circuit, we need to analyze the circuit and apply the appropriate rules for resistors and operational amplifiers.

1. **Identify the Operational Amplifiers (Op-Amps):** - The circuit has two operational amplifiers, each with an input resistance of R and a feedback resistance of R_1 and R_2 respectively.

2. **Determine the Input Voltage for Each Op-Amp:** - For the left Op-Amp: - $u_{i1} = 2V$ - $u_{i2} = -2V$ - For the right Op-Amp: - $u_{i1} = u_o$ - $u_{i2} = u_o$

3. **Apply the Op-Amp Voltage Rule:** - The voltage difference between the inverting and non-inverting inputs of each Op-Amp must be zero for proper operation. - For the left Op-Amp:

$$u_{i1} - u_{i2} = 2V - (-2V) = 4V$$

This indicates that the left Op-Amp is not functioning correctly, as the voltage difference is not zero.

4. **Analyze the Circuit Configuration:** - The circuit configuration suggests that the resistors R and R_1 are connected in series with the left Op-Amp, and R and R_2 are connected in series with the right Op-Amp. - The output voltage u_o is connected to both Op-Amps, indicating that the circuit is likely designed to sum the input voltages.

5. **Summing the Input Voltages:** - Since the circuit is designed to sum the input voltages, we can write the equation for u_o: $u_o = u_{i1} + u_{i2}$ - Substituting the given values: $u_o = 2V + (-2V) = 0V$
Therefore, the output voltage u_o is $0V$.

Figure 41. Examples of responses among 4 baselines in EEE-Bench, where all baselines fail to reason correctly for this problem.

References

[1] Marah Abdin, Sam Ade Jacobs, Ammar Ahmad Awan, Jyoti Aneja, Ahmed Awadallah, Hany Awadalla, Nguyen Bach, Amit Bahree, Arash Bakhtiari, Harkirat Behl, et al. Phi-3 technical report: A highly capable language model locally on your phone. *arXiv preprint arXiv:2404.14219*, 2024. 4

[2] Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023. 1, 4

[3] Stanislaw Antol, Aishwarya Agrawal, Jiasen Lu, Margaret Mitchell, Dhruv Batra, C Lawrence Zitnick, and Devi Parikh. Vqa: Visual question answering. In *Proceedings of the IEEE international conference on computer vision*, pages 2425–2433, 2015. 8, 14

[4] Jinze Bai, Shuai Bai, Shusheng Yang, Shijie Wang, Sinan Tan, Peng Wang, Junyang Lin, Chang Zhou, and Jingren Zhou. Qwen-vl: A versatile vision-language model for understanding, localization, text reading, and beyond. 2023. 4

[5] Bimal K Bose. Energy, environment, and advances in power electronics. In *ISIE'2000. Proceedings of the 2000 IEEE International Symposium on Industrial Electronics (Cat. No. 00TH8543)*, pages TU1–T14. IEEE, 2000. 1, 8, 14

[6] Bimal K Bose. Power electronics and motor drives: advances and trends. 2020. 1, 8, 14

[7] Jennifer AN Brophy and Christopher A Voigt. Principles of genetic circuit design. *Nature methods*, 11(5):508–520, 2014. 1

[8] Tom B Brown. Language models are few-shot learners. *arXiv preprint arXiv:2005.14165*, 2020. 1, 8, 14

[9] Mathilde Caron, Hugo Touvron, Ishan Misra, Hervé Jégou, Julien Mairal, Piotr Bojanowski, and Armand Joulin. Emerging properties in self-supervised vision transformers. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 9650–9660, 2021. 8, 14

[10] Kaiyan Chang, Ying Wang, Haimeng Ren, Mengdi Wang, Shengwen Liang, Yinhe Han, Huawei Li, and Xiaowei Li. Chipgpt: How far are we from natural language hardware design. *arXiv preprint arXiv:2305.14019*, 2023. 1

[11] Kaiyan Chang, Kun Wang, Nan Yang, Ying Wang, Dantong Jin, Wenlong Zhu, Zhirong Chen, Cangyuan Li, Hao Yan, Yunhao Zhou, et al. Data is all you need: Finetuning llms for chip design via an automated design-data augmentation framework. *arXiv preprint arXiv:2403.11202*, 2024. 8, 14

[12] Yupeng Chang, Xu Wang, Jindong Wang, Yuan Wu, Linyi Yang, Kaijie Zhu, Hao Chen, Xiaoyuan Yi, Cunxiang Wang, Yidong Wang, et al. A survey on evaluation of large language models. *ACM Transactions on Intelligent Systems and Technology*, 15(3): 1–45, 2024. 14

[13] Jun Chen, Deyao Zhu, Xiaoqian Shen, Xiang Li, Zechun Liu, Pengchuan Zhang, Raghuraman Krishnamoorthi, Vikas Chandra, Yunyang Xiong, and Mohamed Elhoseiny. Minigpt-v2: large language model as a unified interface for vision-language multi-task learning. *arXiv preprint arXiv:2310.09478*, 2023. 8, 14

[14] Wenhu Chen, Ming Yin, Max Ku, Pan Lu, Yixin Wan, Xueguang Ma, Jianyu Xu, Xinyi Wang, and Tony Xia. Theoremqa: A theorem-driven question answering dataset. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 7889–7901, 2023. 8, 14

[15] Xinlei Chen, Saining Xie, and Kaiming He. An empirical study of training self-supervised vision transformers. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 9640–9649, 2021. 8, 14

[16] Zhe Chen, Weiyun Wang, Hao Tian, Shenglong Ye, Zhangwei Gao, Erfei Cui, Wenwen Tong, Kongzhi Hu, Jiapeng Luo, Zheng Ma, et al. How far are we to gpt-4v? closing the gap to commercial multimodal models with open-source suites. *arXiv preprint arXiv:2404.16821*, 2024. 2, 4

[17] Zhe Chen, Jiannan Wu, Wenhai Wang, Weijie Su, Guo Chen, Sen Xing, Muyan Zhong, Qinglong Zhang, Xizhou Zhu, Lewei Lu, et al. Internvl: Scaling up vision foundation models and aligning for generic visual-linguistic tasks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 24185–24198, 2024. 4, 21

[18] Wei-Lin Chiang, Zhuohan Li, Zi Lin, Ying Sheng, Zhanghao Wu, Hao Zhang, Lianmin Zheng, Siyuan Zhuang, Yonghao Zhuang, Joseph E Gonzalez, et al. Vicuna: An open-source chatbot impressing gpt-4 with 90%* chatgpt quality. *See https://vicuna. lmsys. org (accessed 14 April 2023)*, 2(3):6, 2023. 1

[19] Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, Maarten Bosma, Gaurav Mishra, Adam Roberts, Paul Barham, Hyung Won Chung, Charles Sutton, Sebastian Gehrmann, Parker Schuh, Kensen Shi, Sasha Tsvyashchenko, Joshua Maynez, Abhishek Rao, Parker Barnes, Yi Tay, Noam Shazeer, Vinodkumar Prabhakaran, Emily Reif, Nan Du, Ben Hutchinson, Reiner Pope, James Bradbury, Jacob Austin, Michael Isard, Guy Gur-Ari, Pengcheng Yin, Toju Duke, Anselm Levskaya, Sanjay Ghemawat, Sunipa Dev, Henryk Michalewski, Xavier Garcia, Vedant Misra, Kevin Robinson, Liam Fedus, Denny Zhou, Daphne Ippolito, David Luan, Hyeontaek Lim, Barret Zoph, Alexander Spiridonov, Ryan Sepassi, David Dohan, Shivani Agrawal, Mark Omernick, Andrew M. Dai, Thanumalayan Sankaranarayana Pillai, Marie Pellat, Aitor Lewkowycz, Erica Moreira, Rewon Child, Oleksandr Polozov, Katherine Lee, Zongwei Zhou, Xuezhi Wang, Brennan Saeta, Mark Diaz, Orhan Firat, Michele Catasta, Jason Wei, Kathy Meier-Hellstern, Douglas Eck, Jeff Dean, Slav Petrov, and Noah Fiedel. Palm: Scaling language modeling with pathways, 2022. 7, 19

[20] Can Cui, Yunsheng Ma, Xu Cao, Wenqian Ye, Yang Zhou, Kaizhao Liang, Jintai Chen, Juanwu Lu, Zichong Yang, Kuei-Da Liao, et al. A survey on multimodal large language models for autonomous driving. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 958–979, 2024. 14

[21] Wenliang Dai, Junnan Li, Dongxu Li, Anthony Meng Huat Tiong, Junqi Zhao, Weisheng Wang, Boyang Li, Pascale Fung, and Steven Hoi. Instructblip: Towards general-purpose vision-language models with instruction tuning, 2023. 1

[22] Anna C. Doris, Daniele Grandi, Ryan Tomich, Md Ferdous Alam, Mohammadmehdi Ataei, Hyunmin Cheong, and Faez Ahmed. Designqa: A multimodal benchmark for evaluating large language models' understanding of engineering documentation, 2024. 2, 14

[23] Peng Gao, Jiaming Han, Renrui Zhang, Ziyi Lin, Shijie Geng, Aojun Zhou, Wei Zhang, Pan Lu, Conghui He, Xiangyu Yue, et al. Llama-adapter v2: Parameter-efficient visual instruction model. *arXiv preprint arXiv:2304.15010*, 2023. 8, 14

[24] Timin Gao, Peixian Chen, Mengdan Zhang, Chaoyou Fu, Yunhang Shen, Yan Zhang, Shengchuan Zhang, Xiawu Zheng, Xing Sun, Liujuan Cao, and Rongrong Ji. Cantor: Inspiring multimodal chain-of-thought of mllm, 2024. 7, 19

[25] Yunfan Gao, Yun Xiong, Xinyu Gao, Kangxiang Jia, Jinliu Pan, Yuxi Bi, Yi Dai, Jiawei Sun, Meng Wang, and Haofen Wang. Retrieval-augmented generation for large language models: A survey, 2024. 7, 19

[26] Team GLM, Aohan Zeng, Bin Xu, Bowen Wang, Chenhui Zhang, Da Yin, Diego Rojas, Guanyu Feng, Hanlin Zhao, Hanyu Lai, et al. Chatglm: A family of large language models from glm-130b to glm-4 all tools. *arXiv preprint arXiv:2406.12793*, 2024. 4

[27] Yash Goyal, Tejas Khot, Douglas Summers-Stay, Dhruv Batra, and Devi Parikh. Making the v in vqa matter: Elevating the role of image understanding in visual question answering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 6904–6913, 2017. 8, 14

[28] Danna Gurari, Qing Li, Abigale J Stangl, Anhong Guo, Chi Lin, Kristen Grauman, Jiebo Luo, and Jeffrey P Bigham. Vizwiz grand challenge: Answering visual questions from blind people. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3608–3617, 2018. 8, 14

[29] Dan Hendrycks, Collin Burns, Steven Basart, Andy Zou, Mantas Mazeika, Dawn Song, and Jacob Steinhardt. Measuring massive multitask language understanding. *arXiv preprint arXiv:2009.03300*, 2020. 8, 14

[30] Dan Hendrycks, Collin Burns, Saurav Kadavath, Akul Arora, Steven Basart, Eric Tang, Dawn Song, and Jacob Steinhardt. Measuring mathematical problem solving with the math dataset. *arXiv preprint arXiv:2103.03874*, 2021. 8, 14

[31] Anwen Hu, Haiyang Xu, Jiabo Ye, Ming Yan, Liang Zhang, Bo Zhang, Chen Li, Ji Zhang, Qin Jin, Fei Huang, et al. mplug-docowl 1.5: Unified structure learning for ocr-free document understanding. *arXiv preprint arXiv:2403.12895*, 2024. 14

[32] Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*, 2021. 14

[33] Drew A Hudson and Christopher D Manning. Gqa: A new dataset for real-world visual reasoning and compositional question answering. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 6700–6709, 2019. 8, 14

[34] J David Irwin and R Mark Nelms. *Basic engineering circuit analysis*. John Wiley & Sons, 2020. 1, 8, 14

[35] Albert Q Jiang, Alexandre Sablayrolles, Antoine Roux, Arthur Mensch, Blanche Savary, Chris Bamford, Devendra Singh Chaplot, Diego de las Casas, Emma Bou Hanna, Florian Bressand, et al. Mixtral of experts. *arXiv preprint arXiv:2401.04088*, 2024. 1, 8, 14

[36] Carlos E Jimenez, John Yang, Alexander Wettig, Shunyu Yao, Kexin Pei, Ofir Press, and Karthik Narasimhan. Swe-bench: Can language models resolve real-world github issues? *arXiv preprint arXiv:2310.06770*, 2023. 8, 14

[37] Takeshi Kojima, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. Large language models are zero-shot reasoners, 2023. 7, 19

[38] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. Retrieval-augmented generation for knowledge-intensive nlp tasks, 2021. 19

[39] Bohao Li, Yuying Ge, Yixiao Ge, Guangzhi Wang, Rui Wang, Ruimao Zhang, and Ying Shan. Seed-bench: Benchmarking multimodal large language models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 13299–13308, 2024. 8, 14

[40] Junnan Li, Dongxu Li, Caiming Xiong, and Steven Hoi. Blip: Bootstrapping language-image pre-training for unified vision-language understanding and generation. In *International conference on machine learning*, pages 12888–12900. PMLR, 2022. 8, 14

[41] Junnan Li, Dongxu Li, Silvio Savarese, and Steven Hoi. Blip-2: Bootstrapping language-image pre-training with frozen image encoders and large language models. In *International conference on machine learning*, pages 19730–19742. PMLR, 2023. 8, 14

[42] Kaixin Li, Yuchen Tian, Qisheng Hu, Ziyang Luo, and Jing Ma. Mmcode: Evaluating multi-modal code large language models with visually rich programming problems. *arXiv preprint arXiv:2404.09486*, 2024. 8, 14

[43] Ming Li, Jike Zhong, Chenxin Li, Liuzhuozheng Li, Nie Lin, and Masashi Sugiyama. Vision-language model fine-tuning via simple parameter-efficient modification. *arXiv preprint arXiv:2409.16718*, 2024. 14

[44] Zhenwen Liang, Kehan Guo, Gang Liu, Taicheng Guo, Yujun Zhou, Tianyu Yang, Jiajun Jiao, Renjie Pi, Jipeng Zhang, and Xiangliang Zhang. Scemqa: A scientific college entrance level multimodal question answering benchmark. *arXiv preprint arXiv:2402.05138*, 2024. 8, 14

[45] Dongyang Liu, Renrui Zhang, Longtian Qiu, Siyuan Huang, Weifeng Lin, Shitian Zhao, Shijie Geng, Ziyi Lin, Peng Jin, Kaipeng Zhang, et al. Sphinx-x: Scaling data and parameters for a family of multi-modal large language models. In *Forty-first International Conference on Machine Learning*, 2024. 2, 4

[46] Fuxiao Liu, Kevin Lin, Linjie Li, Jianfeng Wang, Yaser Yacoob, and Lijuan Wang. Mitigating hallucination in large multi-modal models via robust instruction tuning. In *The Twelfth International Conference on Learning Representations*, 2023. 3, 7

[47] Haotian Liu, Chunyuan Li, Yuheng Li, and Yong Jae Lee. Improved baselines with visual instruction tuning, 2023. 4, 8, 14

[48] Haotian Liu, Chunyuan Li, Qingyang Wu, and Yong Jae Lee. Visual instruction tuning. *Advances in neural information processing systems*, 36, 2024. 1, 2, 4, 8, 14

[49] Jiazhen Liu, Yuhan Fu, Ruobing Xie, Runquan Xie, Xingwu Sun, Fengzong Lian, Zhanhui Kang, and Xirong Li. Phd: A prompted visual hallucination evaluation dataset. *arXiv preprint arXiv:2403.11116*, 2024. 3, 7

[50] Mingjie Liu, Teodor-Dumitru Ene, Robert Kirby, Chris Cheng, Nathaniel Pinckney, Rongjian Liang, Jonah Alben, Himyanshu Anand, Sanmitra Banerjee, Ismet Bayraktaroglu, et al. Chipnemo: Domain-adapted llms for chip design. *arXiv preprint arXiv:2311.00176*, 2023. 1

[51] Yuan Liu, Haodong Duan, Yuanhan Zhang, Bo Li, Songyang Zhang, Wangbo Zhao, Yike Yuan, Jiaqi Wang, Conghui He, Ziwei Liu, et al. Mmbench: Is your multi-modal model an all-around player? *arXiv preprint arXiv:2307.06281*, 2023. 8, 14

[52] Haoyu Lu, Wen Liu, Bo Zhang, Bingxuan Wang, Kai Dong, Bo Liu, Jingxiang Sun, Tongzheng Ren, Zhuoshu Li, Yaofeng Sun, et al. Deepseek-vl: towards real-world vision-language understanding. *arXiv preprint arXiv:2403.05525*, 2024. 4

[53] Pan Lu, Ran Gong, Shibiao Jiang, Liang Qiu, Siyuan Huang, Xiaodan Liang, and Song-Chun Zhu. Inter-gps: Interpretable geometry problem solving with formal language and symbolic reasoning. *arXiv preprint arXiv:2105.04165*, 2021. 8, 14

[54] Pan Lu, Swaroop Mishra, Tanglin Xia, Liang Qiu, Kai-Wei Chang, Song-Chun Zhu, Oyvind Tafjord, Peter Clark, and Ashwin Kalyan. Learn to explain: Multimodal reasoning via thought chains for science question answering. *Advances in Neural Information Processing Systems*, 35:2507–2521, 2022. 1

[55] Pan Lu, Swaroop Mishra, Tony Xia, Liang Qiu, Kai-Wei Chang, Song-Chun Zhu, Oyvind Tafjord, Peter Clark, and Ashwin Kalyan. Learn to explain: Multimodal reasoning via thought chains for science question answering, 2022. 2

[56] Pan Lu, Hritik Bansal, Tony Xia, Jiacheng Liu, Chunyuan Li, Hannaneh Hajishirzi, Hao Cheng, Kai-Wei Chang, Michel Galley, and Jianfeng Gao. Mathvista: Evaluating mathematical reasoning of foundation models in visual contexts. *arXiv preprint arXiv:2310.02255*, 2023. 1, 2, 4, 7, 8, 14, 17, 20

[57] Subir Majumder, Lin Dong, Fatemeh Doudi, Yuting Cai, Chao Tian, Dileep Kalathil, Kevin Ding, Anupam A Thatte, Na Li, and Le Xie. Exploring the capabilities and limitations of large language models in the electric energy sector. *Joule*, 8(6):1544–1549, 2024. 1, 8, 14

[58] Kenneth Marino, Mohammad Rastegari, Ali Farhadi, and Roozbeh Mottaghi. Ok-vqa: A visual question answering benchmark requiring external knowledge. In *Proceedings of the IEEE/cvf conference on computer vision and pattern recognition*, pages 3195–3204, 2019. 8, 14

[59] Minesh Mathew, Dimosthenis Karatzas, and CV Jawahar. Docvqa: A dataset for vqa on document images. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pages 2200–2209, 2021. 8, 14

[60] Fanqing Meng, Wenqi Shao, Quanfeng Lu, Peng Gao, Kaipeng Zhang, Yu Qiao, and Ping Luo. Chartassisstant: A universal chart multimodal language model via chart-to-table pre-training and multitask instruction tuning. *arXiv preprint arXiv:2401.02384*, 2024. 14

[61] OpenAI. Gpt-4v (vision) system card, 2023. Citekey: gptvision. 4

[62] OpenAI. Hello gpt-4o. https://www.openai.com/, 2024. 2, 4, 8, 14

[63] OpenAI. GPT-4o Mini: Advancing cost-efficient intelligence, 2024. 4, 7, 20

[64] Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel Haziza, Francisco Massa, Alaaeldin El-Nouby, et al. Dinov2: Learning robust visual features without supervision. *arXiv preprint arXiv:2304.07193*, 2023. 8, 14

[65] Liangming Pan, Michael Saxon, Wenda Xu, Deepak Nathani, Xinyi Wang, and William Yang Wang. Automatically correcting large language models: Surveying the landscape of diverse self-correction strategies, 2023. 7, 19

[66] Yusu Qian, Haotian Zhang, Yinfei Yang, and Zhe Gan. How easy is it to fool your multimodal llms? an empirical analysis on deceptive prompts. *arXiv preprint arXiv:2402.13220*, 2024. 3, 7

[67] Runqi Qiao, Qiuna Tan, Guanting Dong, Minhui Wu, Chong Sun, Xiaoshuai Song, Zhuoma GongQue, Shanglin Lei, Zhe Wei, Miaoxuan Zhang, et al. We-math: Does your large multimodal model achieve human-like mathematical reasoning? *arXiv preprint arXiv:2407.01284*, 2024. 8, 14

[68] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In *International conference on machine learning*, pages 8748–8763. PMLR, 2021. 8, 14

[69] Machel Reid, Nikolay Savinov, Denis Teplyashin, Dmitry Lepikhin, Timothy Lillicrap, Jean-baptiste Alayrac, Radu Soricut, Angeliki Lazaridou, Orhan Firat, Julian Schrittwieser, et al. Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context. *arXiv preprint arXiv:2403.05530*, 2024. 4

[70] Hadi Saadat et al. *Power system analysis*. McGraw-hill, 1999. 1

[71] Tanik Saikh, Tirthankar Ghosal, Amish Mittal, Asif Ekbal, and Pushpak Bhattacharyya. Scienceqa: A novel resource for question answering on scholarly articles. *International Journal on Digital Libraries*, 23(3):289–301, 2022. 8, 14

[72] Freda Shi, Xinyun Chen, Kanishka Misra, Nathan Scales, David Dohan, Ed Chi, Nathanael Schärli, and Denny Zhou. Large language models can be easily distracted by irrelevant context, 2023. 8, 19

[73] Wenhao Shi, Zhiqiang Hu, Yi Bin, Junhua Liu, Yang Yang, See-Kiong Ng, Lidong Bing, and Roy Ka-Wei Lee. Math-llava: Bootstrapping mathematical reasoning for multimodal large language models. *arXiv preprint arXiv:2406.17294*, 2024. 14

[74] Amanpreet Singh, Vivek Natarajan, Meet Shah, Yu Jiang, Xinlei Chen, Dhruv Batra, Devi Parikh, and Marcus Rohrbach. Towards vqa models that can read. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 8317–8326, 2019. 8, 14

[75] Ross Taylor, Marcin Kardas, Guillem Cucurull, Thomas Scialom, Anthony Hartshorn, Elvis Saravia, Andrew Poulton, Viktor Kerkez, and Robert Stojnic. Galactica: A large language model for science. *arXiv preprint arXiv:2211.09085*, 2022. 1

[76] Gemini Team, Rohan Anil, Sebastian Borgeaud, Yonghui Wu, Jean-Baptiste Alayrac, Jiahui Yu, Radu Soricut, Johan Schalkwyk, Andrew M Dai, Anja Hauth, et al. Gemini: a family of highly capable multimodal models. *arXiv preprint arXiv:2312.11805*, 2023. 1, 2, 4, 7, 8, 14

[77] Shengbang Tong, Ellis Brown, Penghao Wu, Sanghyun Woo, Manoj Middepogu, Sai Charitha Akula, Jihan Yang, Shusheng Yang, Adithya Iyer, Xichen Pan, Austin Wang, Rob Fergus, Yann LeCun, and Saining Xie. Cambrian-1: A fully open, vision-centric exploration of multimodal llms, 2024. 8, 14, 21

[78] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*, 2023. 1, 8, 14

[79] Haixin Wang, Xinlong Yang, Jianlong Chang, Dian Jin, Jinan Sun, Shikun Zhang, Xiao Luo, and Qi Tian. Parameter-efficient tuning of large-scale multimodal foundation model. *Advances in Neural Information Processing Systems*, 36:15752–15774, 2023. 14

[80] Ke Wang, Junting Pan, Weikang Shi, Zimu Lu, Mingjie Zhan, and Hongsheng Li. Measuring multimodal mathematical reasoning with math-vision dataset. *arXiv preprint arXiv:2402.14804*, 2024. 8, 14

[81] Ke Wang, Junting Pan, Weikang Shi, Zimu Lu, Mingjie Zhan, and Hongsheng Li. Measuring multimodal mathematical reasoning with math-vision dataset, 2024. 1, 2

[82] Xiaoxuan Wang, Ziniu Hu, Pan Lu, Yanqiao Zhu, Jieyu Zhang, Satyen Subramaniam, Arjun R Loomba, Shichang Zhang, Yizhou Sun, and Wei Wang. Scibench: Evaluating college-level scientific problem-solving abilities of large language models. *arXiv preprint arXiv:2307.10635*, 2023. 8, 14

[83] Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc Le, Ed Chi, Sharan Narang, Aakanksha Chowdhery, and Denny Zhou. Self-consistency improves chain of thought reasoning in language models, 2023. 7, 19

[84] Zirui Wang, Mengzhou Xia, Luxi He, Howard Chen, Yitao Liu, Richard Zhu, Kaiqu Liang, Xindi Wu, Haotian Liu, Sadhika Malladi, et al. Charxiv: Charting gaps in realistic chart understanding in multimodal llms. *arXiv preprint arXiv:2406.18521*, 2024. 8, 14

[85] Jason Wei, Maarten Bosma, Vincent Y Zhao, Kelvin Guu, Adams Wei Yu, Brian Lester, Nan Du, Andrew M Dai, and Quoc V Le. Finetuned language models are zero-shot learners. *arXiv preprint arXiv:2109.01652*, 2021. 8, 14

[86] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc Le, and Denny Zhou. Chain-of-thought prompting elicits reasoning in large language models, 2023. 7, 19

[87] Haoyuan Wu, Zhuolun He, Xinyun Zhang, Xufeng Yao, Su Zheng, Haisheng Zheng, and Bei Yu. Chateda: A large language model powered autonomous agent for eda. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2024. 1

[88] Yuan Yao, Tianyu Yu, Ao Zhang, Chongyi Wang, Junbo Cui, Hongji Zhu, Tianchi Cai, Haoyu Li, Weilin Zhao, Zhihui He, et al. Minicpm-v: A gpt-4v level mllm on your phone. *arXiv preprint arXiv:2408.01800*, 2024. 4

[89] Qinghao Ye, Haiyang Xu, Guohai Xu, Jiabo Ye, Ming Yan, Yiyang Zhou, Junyang Wang, Anwen Hu, Pengcheng Shi, Yaya Shi, et al. mplug-owl: Modularization empowers large language models with multimodality. *arXiv preprint arXiv:2304.14178*, 2023. 8, 14

[90] Longhui Yu, Weisen Jiang, Han Shi, Jincheng Yu, Zhengying Liu, Yu Zhang, James T Kwok, Zhenguo Li, Adrian Weller, and Weiyang Liu. Metamath: Bootstrap your own mathematical questions for large language models. *arXiv preprint arXiv:2309.12284*, 2023. 1

[91] Xiang Yue, Xingwei Qu, Ge Zhang, Yao Fu, Wenhao Huang, Huan Sun, Yu Su, and Wenhu Chen. Mammoth: Building math generalist models through hybrid instruction tuning. *arXiv preprint arXiv:2309.05653*, 2023. 1

[92] Xiang Yue, Yuansheng Ni, Kai Zhang, Tianyu Zheng, Ruoqi Liu, Ge Zhang, Samuel Stevens, Dongfu Jiang, Weiming Ren, Yuxuan Sun, et al. Mmmu: A massive multi-discipline multimodal understanding and reasoning benchmark for expert agi. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9556–9567, 2024. 1, 2, 4, 8, 14

[93] Xiang Yue, Tianyu Zheng, Yuansheng Ni, Yubo Wang, Kai Zhang, Shengbang Tong, Yuxuan Sun, Botao Yu, Ge Zhang, Huan Sun, Yu Su, Wenhu Chen, and Graham Neubig. Mmmu-pro: A more robust multi-discipline multimodal understanding benchmark, 2024. 4, 21

[94] Sukmin Yun, Haokun Lin, Rusiru Thushara, Mohammad Qazim Bhat, Yongxin Wang, Zutao Jiang, Mingkai Deng, Jinhong Wang, Tianhua Tao, Junbo Li, Haonan Li, Preslav Nakov, Timothy Baldwin, Zhengzhong Liu, Eric P. Xing, Xiaodan Liang, and Zhiqiang Shen. Web2code: A large-scale webpage-to-code dataset and evaluation framework for multimodal llms, 2024. 2

[95] Rowan Zellers, Ari Holtzman, Yonatan Bisk, Ali Farhadi, and Yejin Choi. Hellaswag: Can a machine really finish your sentence? *arXiv preprint arXiv:1905.07830*, 2019. 8, 14

[96] Pan Zhang, Xiaoyi Dong, Yuhang Zang, Yuhang Cao, Rui Qian, Lin Chen, Qipeng Guo, Haodong Duan, Bin Wang, Linke Ouyang, et al. Internlm-xcomposer-2.5: A versatile large vision language model supporting long-contextual input and output. *arXiv preprint arXiv:2407.03320*, 2024. 4

[97] Renrui Zhang, Jiaming Han, Chris Liu, Peng Gao, Aojun Zhou, Xiangfei Hu, Shilin Yan, Pan Lu, Hongsheng Li, and Yu Qiao. Llama-adapter: Efficient fine-tuning of language models with zero-init attention. *arXiv preprint arXiv:2303.16199*, 2023. 1, 8, 14

[98] Renrui Zhang, Dongzhi Jiang, Yichi Zhang, Haokun Lin, Ziyu Guo, Pengshuo Qiu, Aojun Zhou, Pan Lu, Kai-Wei Chang, Peng Gao, and Hongsheng Li. Mathverse: Does your multi-modal llm truly see the diagrams in visual math problems?, 2024. 1, 2

[99] Renrui Zhang, Dongzhi Jiang, Yichi Zhang, Haokun Lin, Ziyu Guo, Pengshuo Qiu, Aojun Zhou, Pan Lu, Kai-Wei Chang, Peng Gao, et al. Mathverse: Does your multi-modal llm truly see the diagrams in visual math problems? *arXiv preprint arXiv:2403.14624*, 2024. 1, 8, 14

[100] Renrui Zhang, Xinyu Wei, Dongzhi Jiang, Yichi Zhang, Ziyu Guo, Chengzhuo Tong, Jiaming Liu, Aojun Zhou, Bin Wei, Shanghang Zhang, et al. Mavis: Mathematical visual instruction tuning. *arXiv preprint arXiv:2407.08739*, 2024. 14

[101] Ge Zheng, Bin Yang, Jiajin Tang, Hong-Yu Zhou, and Sibei Yang. Ddcot: Duty-distinct chain-of-thought prompting for multimodal reasoning in language models, 2023. 7, 19

[102] Lianmin Zheng, Wei-Lin Chiang, Ying Sheng, Siyuan Zhuang, Zhanghao Wu, Yonghao Zhuang, Zi Lin, Zhuohan Li, Dacheng Li, Eric Xing, et al. Judging llm-as-a-judge with mt-bench and chatbot arena. *Advances in Neural Information Processing Systems*, 36: 46595–46623, 2023. 8, 14

[103] Denny Zhou, Nathanael Schärli, Le Hou, Jason Wei, Nathan Scales, Xuezhi Wang, Dale Schuurmans, Claire Cui, Olivier Bousquet, Quoc Le, and Ed Chi. Least-to-most prompting enables complex reasoning in large language models, 2023. 7, 19

[104] Kaiyang Zhou, Jingkang Yang, Chen Change Loy, and Ziwei Liu. Conditional prompt learning for vision-language models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 16816–16825, 2022. 14

[105] Kaiyang Zhou, Jingkang Yang, Chen Change Loy, and Ziwei Liu. Learning to prompt for vision-language models. *International Journal of Computer Vision*, 130(9):2337–2348, 2022. 14

[106] Wenwen Zhuang, Xin Huang, Xiantao Zhang, and Jin Zeng. Math-puma: Progressive upward multimodal alignment to enhance mathematical reasoning. *arXiv preprint arXiv:2408.08640*, 2024. 14