ChatGPT Generates a Novel Tax Strategy

by Andrew Blair-Stanek, Nils Holzenberger, and Benjamin Van Durme

Andrew Blair-Stanek

Nils Holzenberger

Benjamin Van Durme

Andrew Blair-Stanek is a professor of law at the University of Maryland, Nils Holzenberger is an assistant professor in computer science at the Institut Polytechnique de Paris, and Benjamin Van Durme is an associate professor in computer science at Johns Hopkins University. This article is based on work supported by the National Science Foundation under grant No. 2204926. The authors would like to also recognize MITRE Corp. for their valuable support, assistance, and research support.

In this article, the authors describe the first publicly disclosed novel tax minimization strategy generated by a large language model like ChatGPT.

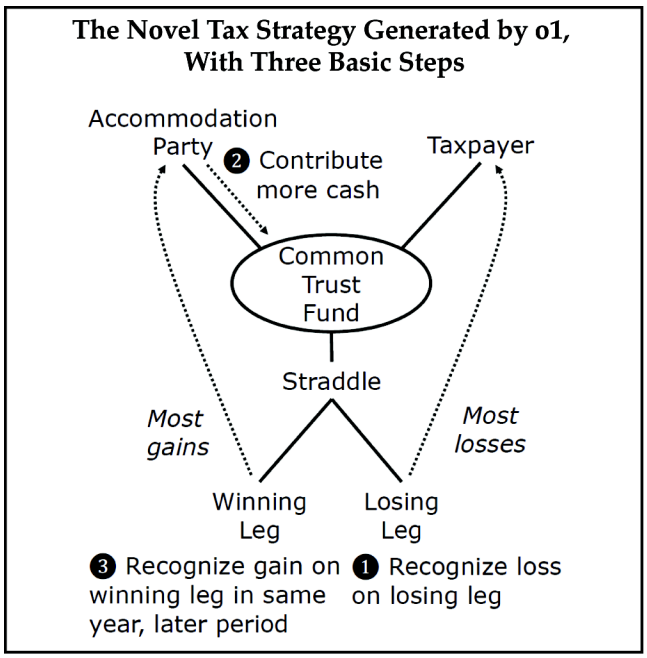
Any opinions, findings, and conclusions or recommendations expressed in this article are the authors' and do not necessarily reflect the views of the National Science Foundation or MITRE.

Three years ago in *Tax Notes*,[1] we predicted that artificial intelligence would be able to generate tax minimization strategies when fed only the text of tax authorities — without any internal IRS data and without any laborious hand coding of tax code sections.[2] That day has arrived.

OpenAI's industrial-grade reasoning large language model o1, which was under the hood of ChatGPT-o1, created what we believe to be the first publicly disclosed tax minimization strategy generated by a large language model. This came from our systematic exploration of how these models understand past tax strategies.

[1] Andrew Blair-Stanek, Nils Holzenberger, and Benjamin Van Durme, "Shelter Check: Proactively Finding Tax Minimization Strategies via AI," *Tax Notes Federal*, Dec. 12, 2022, p. 1515.

[2] Erik Hemberg et al., "Tax Non-Compliance Detection Using Co-Evolution of Tax Evasion Risk and Audit Likelihood," ICAIL (2015) (using hand-coded partnership rules).



The Strategy

The strategy generated by o1 involves straddles and common trust funds. Common trust funds, which are defined in section 584, are a type of passthrough that banks can use to pool assets on behalf of multiple participants.[3] Common trust funds are not themselves subject to tax. Rather, all items of gain and loss pass through to their participants.

Straddles are two opposite financial derivatives contracts, where one contract or "leg" ends up losing money and the other leg ends up making an identical amount of money. Prior to 1981, taxpayers would enter into a straddle, then deduct the loss from the losing contract but indefinitely hold off on triggering income from the winning contract. To shut down this abuse, Congress enacted section 1092, which prevents deducting a loss on a straddle in a tax year before the corresponding gain has been recognized.

In o1's strategy, the taxpayer (*T*) and accommodation party (*A*) would become participants in a common trust fund (CTF), with the taxpayer initially having the larger stake. For example, *T* might borrow $90 cash and contribute it to the CTF, while *A* contributes $10 cash. The CTF chooses to have monthly valuation dates for allocating gain and loss to participants. The CTF then enters into a straddle, such as on foreign currencies. When the foreign currency changes in value, one leg has unrecognized gain and the other has equal or nearly equal unrecognized loss. For simplicity, assume the unrecognized gain is $100 and the unrecognized loss is also $100.

Step 1 occurs right before a valuation date that does not coincide with the end of the calendar year. The CTF disposes of the losing leg, recognizing $100 in loss, which passes through pro rata, with $90 in loss going to *T* and $10 to *A*. These losses reduce their bases in their CTF participation interests (akin to outside basis) to $0.

Step 2 involves the accommodation party *A* making a substantial cash contribution to the CTF. Assume *A* contributes $800, which means it now has a 90 percent interest, having contributed a total of $810 cash, whereas *T* contributed only $90. *A*'s contribution raises its basis in its CTF participation interest to $800.

Step 3 occurs promptly after the valuation date, with the CTF disposing of the winning leg, recognizing $100 in gain, which passes pro rata $90 to *A* and $10 to *T*, raising their bases to $890 and $10, respectively.

Section 1092, which normally stops straddle abuse with the loss being recognized before the gain, does not come into play. Section 1092 prevents taking a loss on a straddle if the corresponding gain has not been recognized by the end of the tax year. But in o1's strategy, the gain and loss are recognized in the same tax year, albeit in different valuation periods.

The accommodation party *A* contributed $810 in cash and is left with a CTF interest worth $810, making the transaction economically neutral. For tax purposes, *A* has a $10 loss and a $90 gain, for an $80 net gain. This gain has no impact if *A* is a tax-indifferent party like a charity or foreign party. But if *A* is taxable, it has an $80 built-in loss on its CTF interest, which has a basis of $890 but a value of $810. *A* could withdraw most of its

[3] Calvin Johnson, "Common Trust Funds: The Living Fossil of Passthroughs," *Tax Notes*, Apr. 5, 2010, p. 103; Lee A. Sheppard, "A Shelter That Only Banks Can Sell," *Tax Notes*, Mar. 26, 2001, p. 1755.

interest in the CTF, which is treated as a sale or exchange and would recognize most of the loss.[4]

The taxpayer *T* borrowed $90 cash and is left holding a CTF interest worth $90, meaning there is no economic loss. For tax purposes, *T* has a $90 loss and $10 gain, for an $80 net loss. The only tax downside for *T* is having an $80 built-in gain on the CTF interest, which has a basis of $10 but a value of $90. This built-in gain can be deferred indefinitely as long as the CTF remains in existence and *T* makes no withdrawals. If *T* dies, the built-in gain disappears.

In sum, *T* has generated an $80 tax loss, with no economic loss except, perhaps, for fees. If this works for a loss of $80, it works for a loss of $80 million.

How We Found It

For a human tax professional, devising an appropriate tax strategy for a client requires knowing something about their factual background, their goals, and the tax provisions that already apply to their situation. Similarly, it makes no sense to simply ask an AI model, "Make me a tax strategy."

We did what computer scientists often do to tackle problems: We built a dataset. Specifically, we created a dataset of 36 known past tax strategies, meticulously setting out the required background facts, the goals the strategies served, the legal authorities required, the steps of the strategy, and a legal analysis of why the strategy purportedly works. One of the 36 was Notice 2003-54, 2003-2 C.B. 1223, which involved straddles and common trust funds to manufacture losses.

We used this 36-strategy dataset for numerous computer science experiments,[5] most of which we do not recount here. We found, for instance, that AI models can implement different levels of scrutiny; they find more legal analysis steps pass muster under lower levels of scrutiny. We also found that even the best AI models have a "yes-man" tendency to say a tax strategy works, even if we give it a deliberately broken strategy.

Generating strategies was only a small part of this computer science research. For generating strategies, we would pass only the background facts, the goals, and the relevant tax authorities to an AI model with the instructions, "You will be coming up with a tax strategy that meets specified goals, given background facts and particular tax-law authorities that the strategy should employ to reach the goals." We have published the full text of the prompt to o1 and its response,[6] as well as our python code and a chunk of our dataset.[7] We ran the strategy-generation test for all 36 in our dataset, against both o1 and Anthropic's AI model Claude-3.5, and we had a human lawyer consider each strategy to determine whether it works legally, meets the stated goals, and accommodates basic economic reality. Overall, the models performed poorly, with Claude generating only seven valid strategies and o1 generating 11. All but one of these valid strategies were merely regurgitations of the known strategy we had been working from. But the remaining strategy was the one depicted in the figure.

Differences From Notice 2003-54

How is o1's strategy different from the one in Notice 2003-54, which provided the background, goals, and authorities that prompted o1 to come up with the strategy? There are two main differences: They involve opposite steps, and they prey on different weaknesses of section 1092.

Notice 2003-54's strategy can be summarized as "*gain, join, loss*." In the notice, the IRS describes an abuse where two or more tax-indifferent parties other than the taxpayer form a common trust fund, which enters into the straddle. The *gain* from the winning leg is recognized and passed through to the tax-indifferent parties. Only later does the taxpayer *join* the common trust fund, which recognizes the losing leg's *loss* and passes much of it through to the taxpayer.

By contrast, o1's strategy can be summarized as "*loss, dilute, gain*." The taxpayer *T* and the accommodation party *A* start out being in the common trust fund, which recognizes the *loss*

[4] Section 584(e).

[5] Blair-Stanek, Holzenberger, and Van Durme, "Can LLMs Identify Tax Abuse?" arXiv (Aug. 10, 2025).

[6] Blair-Stanek, "Novel Tax Strategy From o1-Preview," Github (last updated Oct. 2025).

[7] Blair-Stanek, "Code and Tax-Strategy Data From Paper 'Can LLMs Identify Tax Abuse?'" Github (last updated Oct. 2025).

first. Then, a contribution by *A dilutes T*'s proportionate interest in the fund, which then recognizes the *gain*.

The weakness in section 1092 that Notice 2003-54's strategy exploits has to do with ordering: Section 1092 bars only losses taken in a year before the corresponding gains are recognized. Notice 2003-54's strategy is entirely viable if spread across multiple years. For example, the common trust fund with the tax-indifferent parties could recognize the gain in one year, with the taxpayer not brought in until a subsequent year.

By contrast, o1's strategy exploits section 1092's reliance on the tax year as the relevant period, whereas common trust funds use monthly or quarterly valuation periods. Unlike the notice's strategy, o1's strategy cannot be spread across multiple years.

The Strategy's Origin

How was o1 able to create this tax strategy? Models that came before o1, like GPT-4, immediately generated next words based on the prompt provided by the user. By contrast, o1 was a "reasoning" model that used an internal series of scratch pads to generate a chain of reasoning about a problem.[8] Different approaches and hypotheses could be explored on this internal scratch pad, with the model evaluating these and returning only the best. We think it is likely that this reasoning capacity was central to o1 creating this strategy.

We cannot rule out, however, the possibility that this strategy has been used before and was later seen during o1's training. We performed a thorough legal search for cases, released IRS documents, and secondary sources indicating that "loss, dilute, gain" has been used or described before.[9] We found none. That is why we believe o1's strategy to be novel. But we may have missed some publicly available source. Or some adviser may have described it to clients or in internal memoranda, without it appearing in the sources we searched. Large language models are typically trained on as much data as their creators can get their hands on, even if it has ethically questionable origins. If an adviser had described it, OpenAI may have gotten access to the adviser's description, even though that may implicate client confidentiality. OpenAI, like other creators of these models, does not disclose its training data.

Policy Implications

If we human authors had discovered the strategy, it would not be particularly noteworthy. The IRS already has plenty of tools to challenge its use, including judicial doctrines like economic substance and statutory grounds such as it not being a "transaction entered into for profit" under section 165(c)(2). And, although "loss, dilute, gain" is different than Notice 2003-54's "gain, join, loss," the standard for a transaction to be "substantially similar" and thus a reportable transaction is "broadly construed in favor of disclosure."[10] Both strategies involve common trust funds, straddles, and artificial losses. Although aggressive advisers could point to all the differences from Notice 2003-54, o1's strategy is likely a reportable transaction.

But we human authors did not discover the strategy. The writing is on the wall: AI can discover novel tax strategies. AI's threat to tax revenues will only grow, for three reasons. First, models continue to become more capable. OpenAI's current best model, GPT-5, is already much better than o1.

Second, advances in the field of information retrieval[11] will make it possible to retrieve combinations of tax-law authorities that form ever more novel strategies. Our research was based on the authorities, like the common trust fund provisions and section 1092, already combined in past tax strategies. Advances in information retrieval will allow AI to find new combinations more efficiently than human tax lawyers or accountants.

[8] OpenAI, "o1 System Card" (last updated Dec. 5, 2024); Jun Wang, "A Tutorial on LLM Reasoning: Relevant Methods Behind ChatGPT o1," arXiv (Feb. 15, 2025).

[9] *E.g.*, Sheppard, *supra* note 3 (describing the gain coming before the loss); Johnson, *supra* note 3 (same); Yaron Z. Reich, "The Case for a Super-Matching Rule," 65 *Tax L. Rev.* 241, 311 (2012) (same); *Malone v. Clark Nuber*, 2008 WL 2545069, at *2 (W.D. Wash. June 23, 2008) (same).

[10] Reg. section 1.6011-4(c)(4).

[11] *E.g.*, Orion Weller et al., "Rank1: Test-Time Compute for Reranking in Information Retrieval," arXiv (last revised Aug. 8, 2025); Rulin Shao et al., "ReasonIR: Training Retrievers for Reasoning Tasks," arXiv (Apr. 29, 2025).

Third, accounting and law firms will be able to use AI to create bespoke tax strategies for their clients, based on their clients' unique background facts.[12] By comparison, we fed o1 only the most basic background facts: "There are a Bank and two tax-indifferent parties A and B that are not related to T and are willing to facilitate this transaction."[13] Bespoke tax strategies created by AI will be more sophisticated and harder for the IRS to detect.

In our *Tax Notes* article three years ago,[14] we proposed that Congress or the IRS make AI-generated strategies into reportable transactions, just like strategies with contingent adviser fees.[15] Form 8886, "Reportable Transaction Disclosure Statement," even has extra space, right below box 2e, for a sixth checkbox labeled "AI-generated strategy." Drafting the scope of what AI-generated strategies are covered would require care. For example, a strategy thought up by tax lawyers who collaborated in a document in Microsoft Word, which now incorporates extensive AI, should not count. But the alternative may be a silent decline in tax revenues as AI increasingly allows taxpayers to optimize their tax affairs with bespoke strategies.

There are other lessons. The novel strategy suggests revisiting Calvin Johnson's 2010 proposal in these pages[16] that Congress repeal section 584 and require common trust funds to be taxed as either partnerships or mutual funds. And, of course, it is important to continue computer science research into AI that can aid the IRS.

Conclusion

Our systematic exploration of the capabilities of large language models produced what we believe to be the first novel tax strategy generated by these models. This strategy is fundamentally different from Notice 2003-54 in several ways. Although this strategy seems to pose little danger to the Treasury — it is probably a loser in court and likely a reportable transaction — it should be a wake-up call about the risks AI poses to tax revenues. ■

[12] Benjamin Alarie, "AI and the Future of Tax Avoidance," *Tax Notes Federal*, Dec. 4, 2023, p. 1809.

[13] Blair-Stanek, *supra* note 6.

[14] Blair-Stanek, Holzenberger, and Van Durme, *supra* note 1, at 1518.

[15] Reg. section 1.6011-4(b)(4).

[16] Johnson, *supra* note 3.