



HAL
open science

Shelter Check: Proactively Finding Tax Minimization Strategies via AI

Andrew Blair-Stanek, Nils Holzenberger, Benjamin van Durme

► **To cite this version:**

Andrew Blair-Stanek, Nils Holzenberger, Benjamin van Durme. Shelter Check: Proactively Finding Tax Minimization Strategies via AI. Tax Notes Federal, 2022, 177. hal-04462115

HAL Id: hal-04462115

<https://hal.science/hal-04462115>

Submitted on 23 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Shelter Check: Proactively Finding Tax Minimization Strategies via AI

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 associate professor in computer science at 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.

In this article, the authors explore how artificial intelligence could be used to automatically find tax minimization strategies in the tax law. Congress or Treasury could then proactively shut down such strategies. But, if large accounting or law firms develop the technology first, the result could be a huge, silent hit to the treasury.

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

Copyright 2022 Andrew Blair-Stanek, Nils Holzenberger, and Benjamin Van Durme.
All rights reserved.

In *Peracchi*,¹ the Ninth Circuit addressed what basis a taxpayer had in a note in which he was the debtor. At the end of the opinion, the court did something unusual but wise, saying, “We take a final look at the result to make sure we have not placed our stamp of approval on some sort of exotic tax shelter.”² Yet the court had neither the tax expertise nor the resources to fully consider

whether its holding would enable future tax shelters. The court’s tax shelter analysis took just two paragraphs, covering only a few possible scenarios.

Artificial intelligence will hopefully one day allow judges deciding tax cases to do a thorough check to ensure that their opinions do not enable new tax minimization strategies. We dub this future software “Shelter Check,” and we are now doing the research to make it a reality. A judge’s law clerk would upload a draft opinion to Shelter Check, which would see whether the opinion

¹*Peracchi v. Commissioner*, 143 F.3d 487 (9th Cir. 1998).

²*Id.* at 496.

not plausible. Other errors are false negatives: For Shelter Check, that would be plausible tax strategies that it never identified. Human tax lawyers are good at determining whether a tax strategy is plausible, but not good at foreseeing how authorities might be misused to create new tax strategies. So Shelter Check's false negatives (that is, missed plausible strategies) are more dangerous, and decreasing them should be the priority.¹⁴

All AI models require data for training and for fine-tuning to minimize errors. There are several sources of attempted tax strategies for training and verifying Shelter Check. All existing listed transactions and transactions of interest cite the tax authorities relied on to avoid taxes. Similarly, all court cases discussing tax avoidance or judicial doctrines like substance over form cite the authorities the taxpayers rely on and describe the facts of the attempted strategy.

New-Authority Shelter Check will be substantially faster than Existing-Authority Shelter Check. Suppose that there are 100,000 existing tax law authorities and, for simplicity, that tax minimization schemes are built using just two authorities. Running one new authority through Shelter Check to see if it interacts with any of these existing authorities to create a tax minimization scheme requires checking 100,000 possible interactions. By contrast, each of the 100,000 existing authorities can interact with each of the 99,999 other authorities in 4,999,950,000 different ways. Although clever heuristics likely can reduce the computational burden without substantially increasing the error rate, Existing-Authority Shelter Check will still likely be thousands of times slower than New-Authority Shelter Check. This is good, because judges' law clerks, congressional staffers, and IRS attorneys will want Shelter Check to run quickly. Running Existing-Authority Shelter Check will likely require extensive computational resources. Unfortunately, black-hat tax advisers looking for tax minimization strategies for their clients can afford those computer resources more easily than academic researchers or the IRS. This mismatch leads us to our one policy suggestion.

¹⁴ In statistical terms, Shelter Check should be high recall at the expense of lower precision.

Make AI-Found Strategies Reportable

Our only immediate policy suggestion is for the IRS or Congress to create a new category of reportable transactions for tax minimization strategies that taxpayers or their tax advisers find using AI. Reportable transactions are those that the IRS, through regulations, "determines as having a potential for tax avoidance or evasion."¹⁵ A taxpayer engaging in anything the IRS has designated a reportable transaction must explicitly report the transaction to the IRS.¹⁶ Some reportable transactions are those in which the IRS has specified the substance of the transaction,¹⁷ such as tax shelters the IRS has come across. But several categories of reportable transactions focus on the process behind the transaction. For example, if a tax adviser tells a taxpayer about a transaction and requests that the taxpayer keep its details confidential, that's a reportable transaction.¹⁸ As another example, if a tax adviser promises to return some of its fees if the IRS successfully challenges a strategy devised by the adviser, that's a reportable transaction.¹⁹

The definition of reportable transactions should be expanded to include strategies that taxpayers or their tax advisers discover using AI tools not available to the general public. We outline Shelter Check with the intent that it be a white-hat tool to prevent tax minimization. But we are still years away from having a working Shelter Check, and we plan to publish our intermediate results. (We hope to attract other white-hat AI researchers to the area, and they too will likely publish their intermediate results.) Yet savvy tax advisers, like the big accounting firms, may have a working Existing-Authority Shelter Check before we and other white-hat researchers do. If the black hats win this race, making the strategies they find into reportable transactions will prevent a massive, silent hit to the treasury.

AI is increasingly woven into many tools used by lawyers, including the search functions of Westlaw or Lexis and spreadsheet programs like

¹⁵ Section 6707A.

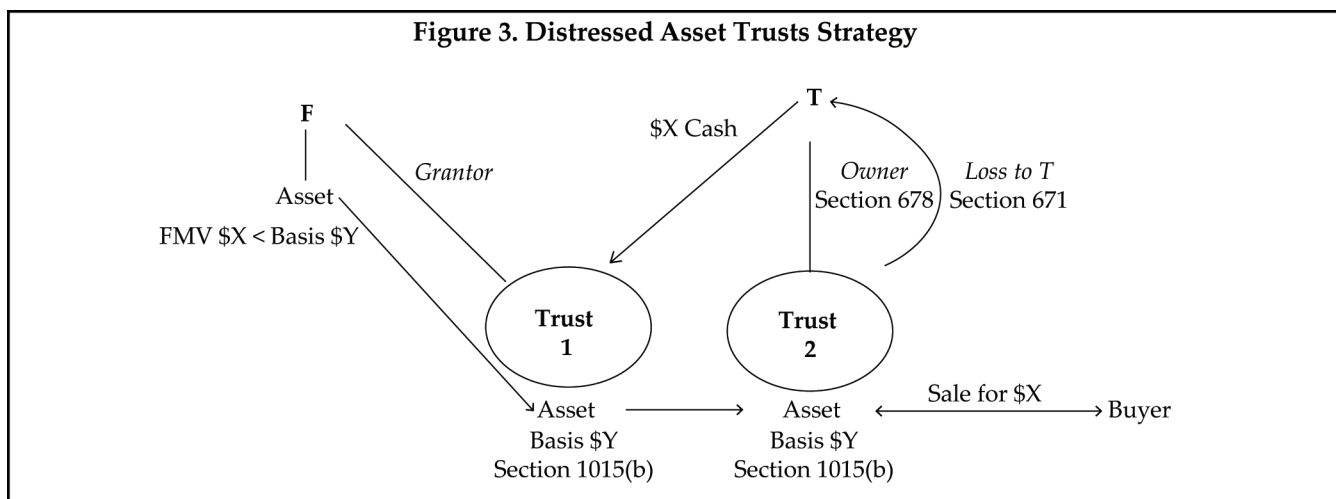
¹⁶ *Id.*; reg. section 1.6011-4(b).

¹⁷ Reg. section 1.6011-4(b)(2) and (6).

¹⁸ Reg. section 1.6011-4(b)(3).

¹⁹ Reg. section 1.6011-4(b)(4).

Figure 3. Distressed Asset Trusts Strategy



strategy came to light.²⁵) Alternatively, the IRS could have issued the ruling, but with an explicit limitation that its holding would not apply if any consideration other than X stock had been contributed to Y. That also would have foiled PwC's later strategy.

This example also illustrates how tax law is particularly amenable to using AI to find abusive strategies. Tax law is unique in that it boils down to a concrete number — money owed to the government.²⁶

Example 2: *Summa Holdings* arrangement.

This transaction is another example of what we call the tax minimization strategy of legal cleverness. It involved the combination of the Roth IRA provisions with the domestic international sales corporation provisions. Congress granted DISCs an explicit tax exemption for commissions on exports to cut taxes on exports and reduce the trade deficit. The taxpayers in *Summa Holdings*²⁷ and other taxpayers who used this strategy owned businesses that exported goods. They also owned Roth IRAs, tax-favored retirement accounts explicitly allowed by Congress that allow tax-free growth and withdrawals. They had their Roth IRAs own DISCs, which collected commissions on their businesses' exports. These commissions were deductible by the

taxpayers' businesses and excluded from the DISCs' income.²⁸

The IRS challenged these tax savings, attempting to recharacterize the commissions as nondeductible deemed dividends to the businesses' owners, followed by excess Roth IRA contributions. But the Sixth Circuit allowed the tax savings, noting that Congress had explicitly enacted both provisions the taxpayers were using. In other words, the taxpayers' strategy was valid tax planning, not tax avoidance.

Shelter Check would seek to catch all tax minimization strategies — not only tax avoidance like PwC's combination of Rev. Rul. 74-503 and section 956, but also tax planning like a Roth IRA holding a DISC. Both tax avoidance and tax planning siphon money from the treasury, and both could be caught before being used — with the proper software.

Suppose congressional staffers had had access to New-Authority Shelter Check when Congress was preparing to enact the Roth IRA provisions in 1997. Shelter Check would have identified the possible combination of those provisions with the DISC provisions that Congress had added over two decades earlier. Congress could have expressly barred Roth IRAs from directly or indirectly owning DISCs. That would have prevented the *Summa Holdings* strategy entirely.

Example 3: Distressed asset trusts (Notice 2008-34, 2008-1 C.B. 645). This transaction, as

²⁵ Rev. Rul. 2006-2, 2006-1 C.B. 261.

²⁶ Cf. Lawsky, *supra* note 3, at 122 ("Other forms — immigration forms, for example — primarily collect and organize information. Tax forms, however, do things with numbers.")

²⁷ *Summa Holdings Inc. v. Commissioner*, 848 F.3d 779 (6th Cir. 2017).

²⁸ Section 995(b)(1)(F)(i) and (g).

shown in Figure 3, is an example of the tax minimization strategy of arbitrage between two taxpayers. Let's say you have a U.S. taxpayer T with income to shelter. Through advisers, T is put in touch with a foreign taxpayer F that is not subject to U.S. tax and owns an asset with substantial built-in loss. Let's say that F's basis in the asset is \$Y, but it is now worth much less, \$X.

F would contribute the asset to a grantor trust, called Trust 1. Section 1015(b) states that the asset still has basis \$Y. T pays \$X in cash to become the beneficiary of Trust 1. Then the asset is transferred to a second trust, Trust 2, designed so that T has the rights described in section 678(a)(1), which means T is treated as the owner of Trust 2. Again, section 1015(b) governs basis, saying that Trust 2 holds the asset with basis \$Y. Then Trust 2 sells the asset to some unrelated third party for the \$X it is worth, resulting in the recognition of the full built-in loss. Under sections 678(a)(1) and 671, this loss goes onto T's tax return, sheltering T's income. Other than taxes, this transaction is a wash for T, who paid \$X cash to F but got back \$X cash on the sale to a third party.

If Congress had New-Authority Shelter Check when it enacted section 678, it would have been alerted to this strategy. Congress could have amended section 1015(b) to set the asset's basis to \$X in these circumstances.

This strategy also demonstrates how Shelter Check need not have a full understanding of all legal concepts. Section 678 applies if a beneficiary has "a power exercisable solely by himself to vest the corpus or the income therefrom in himself," which is likely a hard legal concept for AI to understand or represent. The Shelter Check modeling could simply assume that T might have that power over Trust 2 in finding the Notice 2008-34 strategy. But confirming that a trust can have that power in the real world is precisely why it would be necessary to have a human tax lawyer review potential tax minimization strategies identified by Shelter Check.

The Technological Challenges

The AI capabilities required for Shelter Check do not yet exist, or, if some accounting firm or law firm has developed them, their existence remains secret. We are working to develop the required capabilities and welcome the efforts of other

white-hat researchers. The required AI capabilities fall into two basic categories — natural language understanding and strategy modeling.

We are focused on a form of natural language understanding called semantic parsing. Semantic parsing is concerned with extracting the logical, structured representation of the meaning of language.²⁹ For example, with a revenue ruling, a semantic parse would extract the relevant parties, their relations, their transactions, and the holdings. It might also extract the reasoning the IRS used to reach its holdings. There are also other approaches we are exploring that involve less explicit structure but rely on huge computational models with hundreds of billions of artificial neurons.³⁰

There are two basic types of legal authority — statutory, which sets out rules in the abstract, and case-based, which gives facts and explains how the law applies to them. In tax law, the IRC and most of the Treasury regulations are the statutory type authorities. Tax law's case-based authorities include not only decisions by courts on tax law cases but also the Treasury regulations' examples, revenue rulings, private letter rulings, technical advice memoranda, field service advice, general counsel memoranda, and generic legal advice memoranda. The two types of legal authority pose different challenges for semantic parsing, and we are working on technologies for the semantic parsing of both.³¹

Any area of AI benefits from having lots of data to train predictive models, and semantic parsing is no exception. The raw text of many tax law authorities is now available, and that adds up to a substantial corpus of text. Even more promising is transfer learning, in which an AI

²⁹ Li Dong and Mirella Lapata, "Language to Logical Form With Neural Attention," Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, August 7-12, 2016, Berlin, Germany, Volume 1: Long Papers.

³⁰ Tom B. Brown et al., "Language Models Are Few-Shot Learners," 33 *Advances in Neural Information Processing Systems* 1877 (2020); Cade Metz, "Meet GPT-3. It Has Learned to Code (and Blog and Argue)," *The New York Times*, Nov. 24, 2020, at D6.

³¹ E.g., Nils Holzenberger et al., "A Dataset for Statutory Reasoning in Tax Law Entailment and Question Answering," Proceedings of the 2020 Natural Language Processing Workshop (statutes); Andrew Blair-Stanek and Benjamin Van Durme, "Improved Induction of Narrative Chains via Cross-Document Relations," 11th Joint Conference on Lexical and Computational Semantics (2022) (case-based).

model is trained on a large set of related data before being trained on the domain of interest (here, tax law). There is now a huge trove of that related data for our purposes. The Harvard Law School Library has made the text of virtually all published U.S. case law available to researchers,³² and the U.S. Code and Code of Federal Regulations are available from government sources in structured, computer-readable format.

This huge quantity of available text upends one of the arguments made in favor of the conventional bottom-up approach of using AI on the IRS's huge trove of return data, as opposed to the top-down approach we advocate. The bottom-up approach can use more data, since the IRS receives 261 million tax returns and 4.6 billion information returns each year.³³ But with several billion words of legal text now available, plus transfer learning technology to use nontax law text to train a tax-law-focused model, the top-down approach looks increasingly feasible.

Shelter Check's second technological challenge is strategy modeling. Once we have parsed several tax authorities, how can we combine them to create a tax minimization strategy? For example, in PwC's plan to combine Rev. Rul. 74-503 and section 956, simply understanding the meaning of the text of both authorities is not enough to know how to create the strategy. Rather, that also requires several modestly creative steps: creating a U.S. subsidiary to stand in for Y in the ruling, plus later transferring the cash from the U.S. subsidiary to the parent company. AI can exhibit this limited sort of creativity.

One possible approach to tax strategy modeling is reinforcement learning, which is a branch of AI concerned with agents that interact with an environment to fulfill goals. To fulfill their goals, agents take actions, which elicit reactions from the environment and may yield rewards. For tax strategy modeling, the agents are the taxpayers, whose goal is to maximize their net worth. The actions they can take are any legal action (for example, forming a subsidiary or transferring cash), and the environment consists

of all tax law authorities. Reinforcement learning has managed amazing feats,³⁴ including the development of winning strategies for complex multi-player video games — making it a promising approach to shelter modeling.

A second possible AI approach to tax strategy modeling is language modeling. Language models can generate novel, coherent, logically plausible text, ranging from poetry to news stories. A language model trained on legal authorities and then tax law authorities could be used to systematically generate possible tax minimization strategies, which would then be evaluated using the semantic parses of all existing tax law authorities.

A third possible approach is modeling the associations between different areas of tax law. Computer scientists Jamshid Sourati and James Evans built a data set of links between scientific publications, authors, and scientific entities, such as types of physical materials.³⁵ They then used AI methods applied to that data set to mimic associations that researchers make. Astonishingly, this produced relevant scientific hypotheses. Shelter modeling faces a challenge similar to coming up with scientific hypotheses: associating tax law authorities from different areas of tax law that can interact in unexpected ways.

Conclusion

We have proposed a new approach to using AI in tax law. Rather than the consensus bottom-up approach of feeding the torrent of data the IRS receives into AI models, we propose the top-down approach of understanding the text of all tax law authorities and modeling how these authorities may be manipulated in new and unusual ways to minimize taxes. Our approach has several advantages, including being proactive, allowing outside researchers to help, and using the actual text of tax law authorities without requiring human lawyers to manually encode them. But our approach has a dark side —

³² See Harvard Law School, Caselaw Access Project (2022).

³³ IRS Data Book 2021, at 4, Table 2 (2022); *id.* at 54, Table 22.

³⁴ Reinforcement learning has been applied to tax economics with promising results. Stephan Zheng et al., "The AI Economist: Taxation Policy Design Via Two-Level Deep Multiagent Reinforcement Learning," *Sci. Adv.* 8 (2022).

³⁵ Jamshid Sourati and James Evans, "Accelerating Science With Human Versus Alien Artificial Intelligences," arXiv (Apr. 12, 2021).

it also might be used by tax advisers to find new strategies. To counter that, we propose that the IRS immediately make tax strategies found using AI into reportable transactions.

We should note that the novel top-down approach we propose is not mutually exclusive with the bottom-up approach. Combining the two may turn out to be the most powerful approach to attacking tax minimization. For example, top-down semantic parses of tax law authorities and models of possible strategies may be entered into bottom-up models that review reams of tax return data. Conversely, the plentiful data available to the IRS might be used as input to help train top-down models. ■

taxnotes®

Federal | State | International



**Want to be renowned
in the tax community?**
Write for Tax Notes.

If you do, you will:

- ▶ Receive exposure in the world's leading tax publication
- ▶ Join a network of the best and brightest minds in tax

taxnotes.com/acquisitions

Your byline here.