Time and Contract Interpretation:

Lessons from Machine Learning

Yonathan A. Arbel[*]

Abstract:

Contract interpretation is the task of estimating what distant in time parties meant to say or would have said about a specific contingency. For at least a century, scholars and courts have been debating how to best carry out this task.

Conceiving of the interpretative task as one of prediction, I suggest that there are some valuable lessons to be drawn from a field devoted to building prediction models: machine learning. From this viewpoint, this chapter makes four contributions to the study of contract interpretation. It first defends the view of interpretation-as-prediction against the common linguistic view. The linguistic view perceives interpretation as establishing meaning in the philosophy of language sense. But as applied to *contract* interpretation, such arguments often employ motte-and-bailey argumentation. The second is in explaining a puzzling aspect of the debate about interpretative methods. Both textualists and contextualists insist that their method is more accurate. They can do so because they conflate two senses of the term, precision and accuracy. Third, it brings the hard problem of bias-variance tradeoff to the choice of interpretative methods. Finally, and most speculatively, the chapter distinguishes between interpretation and simulation, and argues that the latter is far more important but far less understood in legal theory. With advances in modeling techniques, the idea of simulation demands serious reconsideration.

* Silver Associate Professor of Law, University of Alabama, School of Law. The author would like to thank Frank Fagan, Murat Mungan, and Saul Levmore.

INTRODUCTION

The source of all interpretative battles in contract litigation is time. Richard Posner once noted that "[t]he main purpose of contracts is to enable performance to unfold over time without either party being at the mercy of the other . . . contracts regulate the future, and interpretive problems are bound to arise simply because the future is unpredictable" (Posner, 2014, at 1582). Judges and juries are late comers to the contract scene. They arrive only after litigation has broken out, and they attempt to deduce how the parties meant to have dealt with the present contingency given the evidentiary breadcrumbs the parties left behind them.[1]

Contracting parties are aware of the corrosive effects of time. They reluctantly recognize that their hopeful moment of execution could eventually erode as new information is revealed, preferences shift, and the environment changes. They lay their best plans for when things go awry. If the price of cotton falls, they might want to give the buyer the right to break free of the contract, and in the event the tractor breaks down, they might want to put the seller responsible to finding a replacement. If the courts are to have any role in this simple narrative, it is one of deferential, literal enforcement of the plans as designed.

Contract interpretation theories attempt to solve two hard problems that the simple narrative glosses over. The first is that the parties' plans are encoded in a diffused manner. The colloquial notion of a 'contract' assumes a unitary, canonical document, but this hardly describes either the law or the practice. Courts enforce contracts that span multiple, disparate documents, and even ones where no document exist at all. More broadly, the expressions that capture the interchange of promises are highly diffused. Parties are socially embedded, consciously or not, in common norms of decency, industry standards, their own course of dealings, advice from third parties, cultural expectations, and so on. These sources, integral to their understanding of their mutual exchange, are never exhaustively reduced into

[1] There are accounts, obviously, where judges do more than reconstruction, and either descriptively or normatively build their own vision of the transaction. But even in such accounts, reconstruction plays a role, and this is our focus here.

text. The interpreter then has to pick and choose which of these diffused sources to use, subject to all the practical constraints of judicial process and economy. The second problem is that the parties have not necessarily planned for each possible contingency. Even if the language would appear absolute on its face, black swan contingencies sunder the assumption that what parties meant to *say* captures what they meant to *do*. This makes the choice of interpretative methods a sensitive one (Barnett, 1992).

These difficulties suggest the desirability and potential of clarifying interpretative frames. Elsewhere, we proposed one such frame: interpretation as prediction (Arbel & Hoffman, 2024). We argued that the prediction frame can be gainfully employed to better understand the judicial task. Under this perspective, what interpretation seeks to do is predict (or, in a sense, retrodict) the parties' intent with respect to a given future contingency. It asks how the parties *would have* allocated rights and liabilities given the current state of the world? If they were to explicitly discuss the current situation as a possibility in real time, what would they say?

My goal in this Chapter is illustrate that the prediction framework naturally invites valuable lessons from the field of machine learning.[2] Admittedly, machine learning would appear at first blush quite distinct from contract interpretation. After all, machine learning is about technology, models, and, well, machines. Interpretation deals with language, doctrine, and humans. But despite appearances, I will try to show that lessons on prediction from machine learning can contribute useful insight and conceptual clarity to the field of contract interpretation.[3]

The prediction frame is best understood in terms of its opposition, what we might call the "linguistic theory" of interpretation. Under this theory, interpretation's ultimate claim is not to predict the parties' intent, but to *understand* or *establish* the meaning of their expressions.[4] The first

[2] Throughout this chapter, I will elide the distinctions (if any) between statistical learning writ large and the subfield of machine learning. I rely on some general ideas developed in Hatie et al., 2017. For a helpful, application-based general introduction tailored for lawyers, see Surden, 2014.

[3] Prediction is a broad category indeed, and today's large language models achieve impressive conversational skills by simply predicting the next token in a sequence autoregressively.

[4] A common expression of this theory is found in the of-repeated passage: "A word is not a crystal, transparent and unchanged, it is the skin of a living thought and may vary greatly in color and content according to the circumstances and the time in which it is used." Towne v. Eisner 245 U.S. 418, 425 (1918).

contribution of this Chapter lies in explaining how the linguistic theory has been misused by contextualists to support their favored interpretative methodology. In contrast, the predictive frame highlights that the methodological choice is grounded in efficacy, not metaphysics.

Another contribution of the prediction frame is its wholehearted acceptance of the idea of *partial* accuracy and its openness to measure it. Given a more nuanced account of accuracy can contribute to theory, and the discussion here shows how contemporary contract interpretation debates could benefit from adopting two different meanings of accuracy: precision and accuracy.

The development of interpretation theory would also benefit from a clear understanding of the bias-variance tradeoff, common from the context of machine learning. The discussion offers an introduction and some lessons.

I conclude by proposing using the idea of *simulation* to clarify the two modalities of interpretation—predicting meanings and predicting purposes. These are conflated in practice, but involve very different models, data, and normative implications. My ultimate goal is to illustrate that legal interpretation, while unique in some ways, also shares a common ground with other disciplines that deal with prediction and modeling. There is therefore ample opportunity for lawyers to learn and teach.

I. TWO AND HALF THEORIES OF INTERPRETATION

Almost all modern commentators and courts agree that the objective of contract interoperation is to "ascertain the parties' intention at the time [they] made their contract." (Burton, 2008, at 1)[5] It is then said that the "The best evidence of what parties to a written agreement intend is what they say in their writing."[6] This puts courts in the position of estimating how past parties, which are separated by time, information, and even preferences from

[5] See also Greenfield v. Philles Records, Inc., 98 N.Y.2d 562, 569 (2002) ("Under statutory rules of contract interpretation, the mutual intention of the parties at the time the contract is formed governs interpretation.")

This view does not negate that judges do more than interpret contracts and sometimes reach outcomes that are directly opposed to the contract interpretation proper, such as when they refuse to enforce a clear term violating minimum wage laws. Such activities often fall under the term construction. For a thorough exposition, see Klass (2018).

[6] Slamow v. Del Col, 79 N.Y.2d 1016, 1018 (1992)

their contemporaneous disputing selves, would answer contractual questions. According to scholars like Scott and Schwartz this exercise has a theoretically correct answer (Schwartz & Scott, 2003).[7]

The fly in the ointment is that hardly anyone agrees on how to operationalize this objective. In modern American jurisprudence, two and a half systems of interpretation vie for dominance. The first is textualism, a system of interpretation that focuses on the text of a final document, the canonical "contract." Textualism confines attention to linguistic information that appears within a small set of intermingling sources: the text within the four corners of the document, the judge's own semantic knowledge, semantic knowledge contained in dictionaries, and meaning heuristics contained within Latinized canons of interpretation. Somewhat confusingly, textualists do care and integrate context. They are deeply committed to the way certain words are used within the contract itself. But they limit this type of contextual inquiry to the limited range of sources just noted. Only when these tools prove inapt, because they result in an inescapable ambiguity or because the document text is alleged to be voidable, that the court is authorized to look to other sources of information.

Contextualism is a system that considers a much broader evidentiary base. The contractual text itself is one part of the base, usually carrying heavier inferential weight than other sources (textualists' portrayal to the contrary notwithstanding). But text is certainly not the only authoritative source. Contextualists hold a capacious understanding of what counts as potentially probative information, leading them to adopt a promiscuous approach to evidence. One notable contextualist, Melvin Eisenberg, defined relevant contextual evidence as including "the context in which the contract was made" as well as "special- community usages, trade usages, course of dealing, and course of performance". (Eisenberg, 2018) Others may be even more permissive.[8]

[7] Schwartz & Scott (2003, at 568) ("There is a consensus among courts and commentators that the appropriate goal of contract interpretation is to have the enforcing court find the 'correct answer.'").

[8] Nanakuli Paving & Rock Co. v. Shell Oil Co., 664 F.2d 772, 803, n.43 (9th Cir. 1981); Columbia Nitrogen Corp. v. Royster Co., 451 F.2d 3 (4th Cir. 1971); Chase Manhattan Bank v. First Marion Bank, 437 F.2d 1040 (5th Cir. 1971).

The "half" is purposive and literal interpretations (for a broader review, see Mitchell 2019), which are assigned here a lighter weight because they are rarely discussed as separate schools in American jurisprudence and are often blended in with the other approaches.[9] The idea of purposive interpretation is to establish the parties' goals from a variety of sources, which is then used to determine the meaning of specific provisions. Purposivism is best contrasted with literalism where the task is to find the meaning of the language used by the parties. To put this contrast in practical light: suppose a contract says that the seller shall deliver the car within 14 days, and that it becomes necessary to answer from what time the clock runs. A literalist approach would seek to find evidence of meaning in the text itself alongside other provisions of the contract, whereas a purposivist approach would ask why the parties chose to set late delivery in the first place.

I will not offer much more on these half choices other than to say that they are compatible with both textualism and contextualism. Indeed, contextualists will often motivate the introduction of contextual evidence to understand not just the parties' meaning, but also their purposes. Conversely, textualists will sometimes divine parties' intent from the document itself, which they can do when the document announces its purposes. In practice, however, there seems to be a clear correlation between textualists and literalists and between contextualists and purposivists.

*

The debates between the approaches are longstanding. Most (American) legal commentators sway contextualist, (Martinez and Tobia, 2023) while most courts lean textualist (Stempel and Knutsen, 2021). The differences between these approaches are sometimes framed in metaphysical terms, other times in pragmatic ones (Silverstein, 2021). There is also arguably a political bent at play, with textualism often dyed with more conservative colors, and contextualism painted progressive. But it is also clear that the argument are burdened with commentary that lacks clarity and is itself open to interpretation. My first goal, then, is to use the prediction frame to

[9] In some parts, the Restatement focuses on a more literalist reading, in others, a more purposive. See e.g., Restatement (Second) of Contracts § 202 (1981) cmt (b) ("interpretation is directed to the meaning of [an integrated] writing in the light of the circumstances."); and cmt (c) ("if such a purpose is disclosed further interpretation is guided by it.")

highlight what I believe is a common misunderstanding about the role of linguistic theory in contract interpretation.

II. CONTRACT INTERPRETATION: FINDING MEANING OR PREDICTING INTENT?

Of the various ways of thinking about the goals of interpretation, the most familiar one is that the goal is to establish the *meaning* of the words *actually* expressed by the parties. I will call this the linguistic theory of interpretation, because it emphasizes linguistic ideas of meaning and invokes linguistic tools and theories of language. Another theory is that interpretation is essentially an exercise in prediction; it seeks to predict what the parties *wanted* to mean. Here, the tools are more utilitarian and ecumenical. The words the parties used are a valid source of inference, but so are the judges' priors on the industry norms, or their familiarity with the contracting parties. The best means of predicting parties' intent are the ones that predict them best.

While this tautology may sound non-committal at first, the prediction frame espouses serious commitments, although their source is empirical rather than theoretical. The goal is to use a bottom-up approach to consider interpretative processes that yield outcomes with high fidelity to parties' intent. The question it emphasizes is which model would yield the optimal predictive accuracy, accounting for adjudication and error costs. This approach is constitutionally open to contestation and adaptation. Unlike the linguistic approach, it is also very sensitive to institutions. For weak institutions, reliance on text may be recommended, for stronger ones, broader sources would be advisable (Posner, 2004). More than anything, the prediction frame recognizes and embraces the probabilistic, uncertain nature of the interpretative endeavor.

Many arguments about interpretation adopt a linguistic theory frame, and use it to support some variant of contextualism. They hold that to *truly* determine meaning we must be attentive to contextual evidence. The prediction frame challenges this view, and would see it as a motte-and-bailey argument rather than a true argument for contextualism.[10]

[10] Burton, ____, § 4.6.1. at 144 ("The chief criticism of the plain meaning and four corners rules has been that there are no plain meanings that an interpreter can find on a contract document's face."); Gilson et al., *supra* note 6, at 36 ("Contextualist jurisdictions … reject the

A motte and bailey fallacy is the argumentative strategy of relying on strong, uncontroversial claims (motte) to make expansive, unsupported assertions (bailey) while conflating the two (Shackel, 2005).

The motte of legal interpretation is well expressed here by Stanley Fish:

> *"A sentence is never not in a context. We are never not in a situation. . . . A set of interpretive assumptions is always in force. A sentence that seems to need no interpretation is already the product of one." Fish (1978)*

We learn from this passage that the meaning of words is always affected by their context, which is a central tenet of linguistic pragmatics (Geoffry & Thomas, 1989). This idea is easily demonstrated when we consider how the referent of *car* can depend on which one a speaker is pointing at. This observation is both useful and uncontroversial. It makes for a good motte.

However, the motte does not make for a sound legal theory of contextualism. Rather, contextualism occupies the bailey of the argument, it is the argument that *because* of indeterminacy of language, judges *should* admit into evidence testimony about past conversations, trade usage, or past dealings.

The arc of the argument starts by first portraying an imagined version of textualism:

> *[Textualists believe that t]he [judge can] . . . retir[e] into that lawyer's Paradise where all words have a fixed, precisely ascertained meaning; where men may express their purposes, not only with accuracy, but with fullness; and where, if the writer has been careful, a lawyer, having a document referred to him, may sit in his chair, inspect the text, and answer all questions without raising his eyes "(Thayer, 1898, at 428-29)*

Although this platonic view of language is hardly descriptive of the view of real textualists, launching this attack sets the stage up for the bailey. Consider Corbin's move in the following passage:

> *[I]t can hardly be insisted on too often or too vigorously that language at its best is always a defective*

notion that words in a contract can have a plain or unambiguous--context free--meaning at all.").

> *and uncertain instrument, that words to not define themselves, that terms and sentences in a contract, a deed or a will do not apply themselves to external objects and performances, that the meaning of such terms and sentences consists of the ideas that they induce in the mind of some individual person who uses or hears or reads them, and that seldom in a litigated case do the words of a contract convey one identical meaning to the two contracting parties or to third persons.* ***Therefore****, it is invariably necessary, before a court can give any meaning to the words of a contract and can select one meaning rather than other possible ones as the basis for the determination of rights and other legal effects, that extrinsic evidence shall be heard to make the court aware of the "surrounding circumstances," including the other persons, objects, and events to which the words can be applied and which caused the words to be used. (emphasis added) (Corbin, 1960, at §535)*

The word 'therefore' here highlights the shift to the bailey, making it seem like a logical entailment. The problem for him is that if we are to accept that language is "endlessly self-referential" in the Sausserian sense (de Saussere, 1959),[11] then there is no stopping moment. Judges, and any other interpreter for that matter, would be eternally out of water if they were to accept Fish's idea that sentences are never "not in context" as a jurisprudential theory of meaning. If meaning can only be determined from context then the relevant set of evidence must include the contract, every circumstance that led its formation, and then every circumstance that led to those circumstances, and on, and on, and on.

Of course, no contextualist has ever proposed such an infinitely malleable inquiry or even found one to be possible or necessary. Everyone understands—including dye in the wool contextualists—that ultimately, there is such a thing as 'good enough' in legal interpretation. A party may proffer an interpretation that is so remote, implausible, and abstruse that, even if theoretically relevant, can be safely ignored.

But this means that interpretation is ultimately about finding the best probabilistic approximation of meaning, subject to practical constraints, not about determining the veritable linguistic meaning of the parties'

[11]

expressions. Thus, even the contextualists yield to the empirical and pragmatic, at the expense of the metaphysical.

With this in mind, let us restate the case for contextualism and textualism not as one that is rooted in the philosophy of language, but as one based on the psychology of the parties. The claim that context is important is essentially a claim that context adds accuracy. In the words of Gilson, Sabel, and Scott, the claim "rests on the powerful intuition that fair and efficient contracting takes place in a social context" (Gilson et al., 2014). When parties say "computer" they mean the one on display, and when they say "payment" they mean that form of payment that they discussed over dinner yesterday. They do not mean the platonic ideal of computation machines or even some prototypical exemplar of computers. This sensitivity to implicit incorporation of context is notable in Llewellyn's approach to the UCC (Scott, 2002). The point is that one does not need to resort to vaguely stated arguments about semiotics to justify contextualism. The argument is, quite simply, that without context the predictive accuracy of judgments is low.

And likewise, the case for textualism is best rooted in predictive accuracy. It is the argument that when courts interpret the contract following its plain meaning, notwithstanding the vagueness of this term, they can produce interpretations that are sufficiently accurate, again, in light of various practical constraints. That is, the predictions of this cheap method are sufficiently accurate to justify ignoring the broader context.

The prediction frame reorients our thinking towards the 'good enough.' It further gives us a common benchmark to evaluate interpretative methods. Thinking about predictive accuracy, weighted for costs, fairness, and similar considerations, offers a conceptual metric for the evaluation of different interpretative models. It clears part of the conceptual thicket that grew in this debate. And with predictive accuracy being at the epicenter, some other disciplinary arbitrage might be productive. In the remainder of this Chapter I will present more ideas from machine learning (and statistics more generally) regarding prediction that sharpen our use of language, suggest the limits of a different approach, and offers a way through to a more integrative approach to evidence.

III. Precision and Accuracy

What do we mean when we talk about accuracy? Ironically, the term itself is ambiguous, and some of the disagreement on which system of interpretation is more accurate trades on this ambiguity. This allows both sides to claim that their method is more accurate, while pointing at the other's faults (Schwartz & Scott, 2010).

Because a model is only as good as its predictions, machine learning is deeply invested in developing various measures of accuracy. We will borrow its sensitivity to conceptual clarity regarding accuracy, by distinguishing between precision and accuracy *stricto sensu.*[12] In order to stay grounded, let us consider how the various meanings play out in the context of an old contract chestnut, the case of *Varney v. Ditmars* 217 N.Y. 223 (N.Y. 1916).

In *Varney*, an architect working for a firm was considering an outside job offer. His boss, the defendant, swayed him to stay by offering him both a salary increase and, if the year ends well, an end-of-year bonus comprising a "fair share of my profits." The worker stayed with the employer and worked diligently throughout the year. But in November a dispute arose between the parties, and he was fired. The boss refused to pay any of the year's profits and the architect sued.

The court denied the claim, saying that the contract term "fair share" was too vague and indefinite to support any specific share of profits. Thus, under the doctrine of certainty, the damage award could not be supported. The minority, by Cardozo, held that the problem was not the vagueness of the contractual text, but rather the lack of contextual evidence. Had the architect brought evidence of the industry custom about profit sharing, Cardozo reasoned, the court could have integrated it into the contract and used it a metric for evaluating the fair share of profits. But absent such a showing, the court had nothing to integrate.

At first glance, the majority opinion appears self-contradictory. It is evident that while the parties might not have held in common a unique understanding of the term "fair share," they did mean to share *some* profits. Perhaps it was one sharing scheme, perhaps another, but one thing is certain: they meant to share *some* profits. In effect, the employer gets to retain the full extent of the business' profits. This means that in the infinite space of

[12] Precision and accuracy belong to the larger field of statistical analysis, and machine learning has not invented these concepts. But they are a deep part of the field's "ethos." For an overview, see Shalev-Shwartz & Ben-David, 2014.

sharing ratios, the court managed to choose the single one that is certainly wrong. In the name of accuracy, then, the court settled on the least accurate interpretation.

Perhaps the majority's decision can be salvaged. A more coherent interpretation of the majority stance would be that the court recognized the weakness of its own priors. Feeling little confidence in its own sense of the term 'fair share,' it thought to sacrifice accuracy in the current case for greater accuracy in future cases. If parties would learn that courts do not enforce such vague terms, parties would offer greater guidance in the future, resulting in more accurate judgments. Under this view, the court was essentially setting a "penalty default rule," one that intentionally goes against the parties' wishes, in the hope of shaping their future behavior.

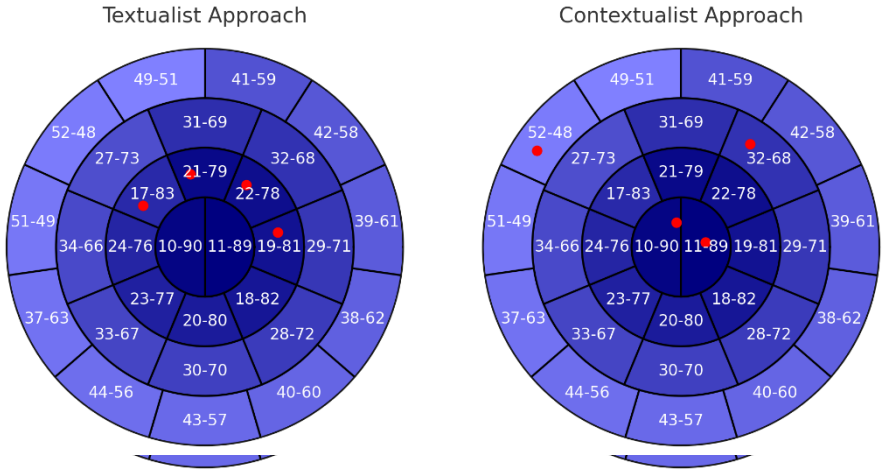
But such a penalty default rule is not particularly appealing in this context. If employers are best positioned to draft employment agreements—as they often are—they only stand to gain from this unexpected penalty default rule which relives them from any duty to share profits when the terms are vague. As such, it is not clear that this rule would indeed invite greater clarity in the future.

The court could have reached a more sensible outcome by simply enforcing its vague intuitions. A textualist court adopting a rough understanding of the term "fair share" might rely on some general view that a (say) 20-80 split is appropriate. Admiteddly, the court would likely be mistaken by this semi-arbitrary choice. Let us stipulate that the custom in that specific industry is actually 10-90 for worker incentives programs. That is we stipulate that by following its imprecise priors, a textualist court would deviate from what the parties likely have meant by "fair share."

Now consider a contextualist court. If the court were presented with contextual evidence, it is more likely that it would discover the industry custom. At the same time, the court will also have exposed itself to motivated testimony and unreliable evidence, so it is also possible (even if less likely) that it would find a much different split appropriate. Given the breadth of potential contextual evidence, many possibilities regarding the court's eventual findings present themselves.

The following figure depicts a potential distribution of answers we might expect the court to make under each approach:

Figure 1 Contextualism vs. Textualism as Precision vs. Accuracy



On the left panel, we see the results of a hypothetical textualist approach. When the court renders its decision, it reaches a resolution that has low variance, somewhere around 20-80. This resolution would be *inaccurate* because we know that the industry standard is 10-90. But it would be quite *precise*. Parties could look at precedents and see how courts interpret "fair share" in a more a-contextual manner, and so the parties could predict the court's next decision with *precision*.

On the right panel, we see the results of a contextualist investigation. Two dots are very close to the center, showcasing a court that reaches the true outcome by consulting contextual evidence. However, this higher degree of accuracy is counterbalanced by a greater scatter or spread. The two other results are very from both the correct result and for the other ones. This means that the precision, and therefore the predictability, of the court's decision is much lower.

Lawyers have a sense of this distinction, which is what they mean by the "certainty" advantage of textualism over contextualism (Linzer, 2002). However, they often seem to treat it as a quality rather than a quantity. As the discussion above illustrates, it is possible to achieve precision (or certainty) arbitrarily and mechanically: you can always choose to do the same thing (e.g., rule for the plaintiff), and this method too will maximize precision. The court's decision in *Varney* does something similar to this: it chooses an arbitrary but certain decision rule and enforces it. The

precision—accuracy distinction is useful in showing that we can aim for rules that have more precision *and* more accuracy (that is, they are on the Pareto frontier).

Consider Peter Linzer's lament that "It is comforting to live in a world of plain meaning. But that world just isn't real. Rather than indulging in a fantasy of certainty, we should opt for a world of reality, however untidy it may be." (Linzer, 2002, 839). Here, we are presented with a choice: between certain but fantastically inaccurate textualism and accurate even if untidy contextualism. Yet, our choice space is much broader! Textualism a-la Varney has high certainty but low accuracy; textualism that integrates some world knowledge has high certainty *and* higher accuracy.

The trade-off between precision and accuracy is obviously a normative question. Its resolution depends on how much we care about judicial interpretations that are clustered together, and therefore easier to predict, relative to results that are closer to the true intention of the parties but can be harder to predict. For theories of justice that emphasize treating similar cases similarly, precision would be more important than accuracy. For theories that emphasize individual merit, accuracy is more important. For efficiency reasons, parties might prefer precision, because it allows them to plan, and settle cases, more easily. This counsels in favor of textualism, but only if the court is reasonably accurate. If the court's interpretation under textualism veers too far from what they actually meant, it will become necessary to opt for a contextualist interpretation, or otherwise, the parties will be limited in their ability to contract for specific outcomes.

The precision-accuracy distinction illuminates other important points about current debates. Take Schwartz and Scott's seminal work on contract interpretation. They argued that when it comes to sophisticated parties, it is less important that courts be correct in any specific case than that they will be correct "on average." (Schwartz & Scott , 2003) This is because sophisticated parties are risk-neutral and would therefore not care about the variance of the decision rule, so long as it is unbiased. This argument is important to them because they concede that textualism is less accurate than contextualism, but argue that it does not matter since it does not affect the mean resolution. [13] However, as we have discussed here, the choice of

[13] Schwartz & Scott, 2003, at 933 ("we concede that a court is more likely to make an accurate interpretation if it sees more evidence")

textualism versus contextualism also affects the locus of the mean of judicial interpretations. Contextualism is not just higher variance, but it is also more accurate than textualism.[14] This makes their case for textualism weaker. However, the analysis also highlights an advantage of textualism, which is its higher relative predictability, making settlement more likely and reducing the probability of disputes.

To summarize, we introduced here the tradeoff between accuracy and precision, a persistent issue in building machine learning models. We saw how both concepts capture different senses of the term accuracy that is used in interpretation debates, and how distinguishing between them is useful in understanding the debate and the choices it makes. We now move to consider another related aspect of predictive accuracy: the bias-variance tradeoff.

IV. BIAS V. VARIANCE

It is commonly argued in the interpretative debates that more contextual information will improve judicial accuracy. The central value proposition of contextualists is exactly that: that with more information will come greater fidelity to the parties' true intents. They have been quite successful in making this claim. Even the Scott and Schwartz' otherwise intrepid attack on contextualism concedes as much, agreeing that contextualism is more accurate and then pivoting to attack its cost-effectiveness. (Schwartz & Scott, 2006).

Scott and Schwartz may be conceding too much. The marginal value of more extrinsic evidence is not only declining, it might actually be negative. This is the case whenever we have reason to worry that parol evidence is likely, on average, to be forged, inaccurate, or biasing in its effects (Bridgeman, 2006).[15] A big part of the law of evidence is premised on exactly

[14] To clarify, the court's prior will be informed by the norm in all industries. This prior does not have to be biased to lead to systemic bias in a decision that relates to any specific industry, like architecture. I believe that Scott & Schwartz' response would be that the parties would draft enough of the contract to ensure that the court has sufficient context to avoid bias.

[15] Residential Mktg. Grp., Inc. v. Granite Inv. Grp., 933 F.2d 546, 548 (7th Cir. 1991) ("Desire for certainty and predictability, perhaps combined with some distrust of juries, has resulted in a presumption that the judge will try to puzzle out the meaning of the contract without recourse to inevitably self-serving, often protracted, and typically inconclusive oral testimony").

this idea: that some types of evidence are likely to be so problematic as to have a net negative informational value.[16]

There is an even broader sense in which this concession goes too far, and it concerns the classic bias-variance tradeoff (Geman et al, 1992). Interpretative processes are like models. They are fed inputs (evidence) and then are expected to produce a prediction (say, a verdict). In building models, we should pay close attention to their complexity. A model's complexity is important because it affects its accuracy, but also because it affects its explainability (a simpler model would tend to be easier to understand) and its predictability (the behavior of a simpler model tend to be easier to predict).

Broadly speaking, textualism offers a (relatively) low complexity model because it focuses on a limited set of "features" or "dimensions."[17] The court makes its analysis based only on the text of contract and a limited set of tools (like dictionaries and canons). By comparison, the contextualist model is more complex. The number of factors that inform the court includes everything that a textualist would consider, and then also a large number of indicators of circumstances. This will include an old email, testimony of a water fountain conversation, and photographic evidence of a smiling handshake. There is no set list of features, nor established procedures for their weighting, and the process becomes as complex as the parties make it.

This leads us to a central issue in designing models: the bias-variance tradeoff. To explain this idea, first technically and then intuitively, consider the way it is often formulated in the context of prediction models. Every model's predictions involve errors, and those are commonly decomposed into the following measure of mean squared errors:[18]

$$MSE = Bias^2 + variance + irreducible\ errors$$

[16] Rules of Evid., Rule 403, ("The court may exclude relevant evidence if its probative value is substantially outweighed by a danger of one or more of the following: unfair prejudice, confusing the issues, misleading the jury, undue delay, or needlessly presenting cumulative evidence.").

[17] It is worth remembering that textualism can mimic the complexity of contextualism. The parties can instruct the court to consider parol evidence as part of their agreement. The difference is that a contextualist judge would assume, perhaps irrebuttably so, that the parties intended for context to be included.

[18] While the MSE equation provides a helpful framework, it assumes a linear additive relationship, which may not capture the full complexity of legal decision-making.

The bias and variance are defined as:

$$Bias^2 = \left(E_D(\hat{f}(x; D) - f(x)\right)^2$$

$$Variance = E_D[\left(\hat{f}(x; D) - E_D\left[\hat{f}(x; D)\right]\right)^2]$$

Where D is the dataset used to estimate the model, x is a vector of features, $\hat{f}$ is the specific model and f(x) is the true underlying function. The bias-variance tradeoff emerges because a process that is meant to reduce variance would often involve smoothing, which in turn will increase bias. (Geman et al., 1992, at 10).

Put in intuitive terms, bias captures the idea that simple models or decision rules will possess a known level of error. If you always carry an umbrella during the winter, you will often be lugging one along on clear days. The decision rule – if winter then umbrella – has a clear bias (in its statistical sense), and if we look back, we will find that, while it is better than the decision rule of never carrying an umbrella, it has an expected level of error. Despite these errors, the decision rule has the double virtue in that it is (a) not irrational and (b) simple to understand.

It will be tempting to adopt a more nuanced approach to the umbrella decision. Looking back, one could note many other features that were correlated with the rain, such as a weather forecast or a simple glance outside the window at the gathering clouds. Even more complex features exist, and one might notice a striking regularity: that last year it was always rainy on Tuesday but rarely on Wednesday, or that for some reason, every time the weatherperson wore blue, it would rain the next day. Applying a decision rule that also incorporates these historical features could yield a model that shows remarkable performance in predicting every occurrence of rain in our dataset. It will be much more accurate than the simplistic decision rule. Yet, it will also be highly complex and recklessly brittle.

The simple decision rule was biased because it "underfits" the data. It fails to consider some more nuanced features. It inevitably results in sunny days with an umbrella, contributing to the model's "bias." Conversely, the complex umbrella rule is an illustration of a model with high variance due to "overfitting" of the data. This model will perform poorly on new data, like a new weatherperson that always wears blue, because it is too sensitive to the context of each past example.

The bias-variance dilemma holds that there is some degree of complexity that is optimal. Simpler models tend to have a high bias, they fail to capture relevant factors so make predictable errors (like carrying an umbrella on sunny winter days). More complex models can have high variance, they overfit historical data and perform poorly on new data (like umbrella rules based on the weatherperson's shirt color). As model complexity increases, bias typically decreases but variance usually increases. A good model is one that balances well between these two sources of inaccuracy.

Applying these distinctions to contract interpretation, recall that court decisions do more than resolve the current dispute, they also set precedent for future cases. This is true even for interpretation, as courts take heed of how past courts have interpreted similar concepts and build on their interpretations.[19] To illustrate, in one of the most important interpretation cases of the last decades, the court tried to define the meaning of "flood" within an exclusion term in an insurance policy, following hurricane Katrina.[20] One factor in the court's decision was how past courts defined other sources of flood, specifically, whether flood could be caused by breakage of a water main.[21] In addition to the court's anticipation of future cases, when parties draft contracts they also do so in anticipation of how the court would interpret their words. And so, we have a system where all the actors are trying to anticipate each other's actions and decisions, in a dynamic feedback loop. The question we focus on now is the *dynamic* or *precedential* utility of each system of interpretation.

A rule of interpretation that is radically textual will produce precedential holdings that tend to be biased. A court that does not consider how words are informed by *anything* other than the words alone—excluding other parts of the contract—will be predictably wrong in deciphering parties' meaning and intent. Every time the model mis-predicts the parties' meaning using its procedure it will contribute to the measure of its bias.

Contextualists correctly argue that a more complex model, one that incorporates more evidence, will reduce this bias. They express this sentiment

[19] For example, according to Larry DiMatteo, courts have developed a common sticky interpretation of the "duty to defend" (DiMatteo, 2004)

[20] *In re Katrina Canal Breaches Litig.*, 495 F.3d 191, 199 (5th Cir. 2007)

[21] Id, at 216 ("Unlike a canal, a water main is not a body of water or watercourse.").Sher v. Lafayette Ins. Co., 2007-2441 (La. 4/8/08), 988 So. 2d 186, 195, on reh'g in part (July 7, 2008) ("inundation of property due to broken water mains . . . would not be excluded as a 'flood'").

in terms of capturing meaning, rather than accuracy or bias, but this is a fair interpretation of their goals. However, what they either not recognize or admit is that their own contextualist model would also tend to introduce variance.[22]

The variance comes from the fact that decision rules that are too immersed in the context of past cases would fail to generalize to new cases. When a court references a past decision of a contextualist court that interprets a term like "flood," it will rely on decisions that are overly sensitive to the context of previous cases and will be misleading in the current context. This would lead the court to adopt a meaning that can stray significantly from the meaning intended by the parties.

The contextualist might seek to salvage the rule by arguing that the court should come *tabula rasa* to each case, ignoring all past interpretative decisions. But this is not a realistic expectation of humans or institutions and it involves a great degree of duplication of effort. But even if it were possible, it would mean that parties cannot effectively choose words today based on a reliable estimate of how they would be interpreted tomorrow. An all-things-considered model of interpretation only function predictably with a Solomonic judge and no restriction on available evidence. But in the real world—well, who knows. The variance is very high.

The key lesson is that there is a sweet spot between the minimization of bias and variance. To paraphrase Einstein, we want interpretative modalities that are as simple as possible, but not any simpler. The corresponding idea in contract interpretation is that if we want an interpretation rule that both predicts parties' intent and guides it, we should aim for intermediate approaches. The corner solutions of extreme contextualism or radical textualism suffer from an inherent flaw: they save bias at the expense of variance, or sacrifice bias to save reduce variance. But in fact, optimal prediction models should have some mix of the two. To be sure, bias and variance are important factors but not exclusive ones. We also care, as the preceding discussion explains, about questions of certainty, litigation or drafting cost, fairness, and the direction of errors, among others. But the discussion here should clarify that, even if hold all of these considerations

[22] The relationship between added variables and increased variance is not necessarily linear, and so it is not necessarily the case that every additional variable would introduce variance.

stable, we would still not necessarily want a fully contextualist or textualist analysis.

V. INTERPRETATION VERSUS SIMULATION

Up to this point, we conceptualized interpretation as the task of predicting parties' intent based on their evidentiary breadcrumb trail. The leading camps differ on which, or how many, breadcrumbs to use. However, they both agree that the goal of this evidence is to inform the task of contract interpretation. Here, I want to offer, in an exploratory manner, a different paradigm altogether: from interpretation to simulation.

Again, to keep us grounded, let us start the analysis with a case. This time, Krell v Henry [1903] 2 KB 740. The case takes us back to England in 1902, and specifically, to Pall Mall street, where the coronation procession of King Edward VII was planned to take place. To watch it, Henry leased for two days Krell's apartment which had a strategically positioned balcony overseeing the busy street. Henry paid in advance £25 on account of the agreed upon sum of £75. When the procession was canceled due to the King's appendicitis, Henry refused to pay arguing that he had no use of the apartment. Krell sued Henry for the remainder, and Henry countersued for the deposit.

The question to the court, as it put it, is whether the contract contained an implied condition that the contract is conditional on the procession taking place. The court answered this in the positive, on the grounds that parol evidence supported the view that both parties had an understanding that the lease concerned the procession. They have implicitly agreed, the court thought, that shall the procession fail to take place, Henry would be under no obligation to pay Krell for the apartment.

This decision has been criticized on many grounds. The contract's language is absolute, and some commentators took issue with the court's imaginative introduction of a clause to a contract that the parties could—but chose not to—introduce themselves (McElroy and Williams, 1941). An alternative, they argued, would be to ask "what did the parties trade?" then the answer might be "a view to watch the procession." And if this is the interpretation given to the contract, then it is no long clear that the obligation to pay arises when the seller cannot provide the good promised.

On careful reflection, neither approach appears satisfactory. It is eminently possible that the parties actually agreed to make an absolute

payment, the way we often do when we make a nonrefundable reservation. Krell, after all, would have had to hold the apartment clean and available, and suffers a loss of opportunities by entering into the contract. Henry, on the other hand, may have arranged for a family gathering in the apartment, and could still use it. The court's interpretation, that the parties included an implied condition, would give both parties the option to cancel the agreement at will once the procession was cancelled, overriding the parties actual purposes.

The court must have sensed the difficulty of its position because it sought to, even if unpersuasively, distinguish this case from the case of cab fares. The scenario considered is of a person ordering a cab to a watch a sports game that is later cancelled. The court believes that even if both parties were well aware of the purpose of the ride, such a contract would not include an implied condition, and the obligation to pay fare would not be discharged.[23] The reason, mysteriously stated, is that the special reason for the trip, which may well be reflected in the cab fare, was not a "foundation of the contract."

In reality, the solution to such cases lies in the embedded option view of remedies articulated by Vic Goldberg (Goldberg, 1988). Parties enter into contracts with cross cutting expectations and constraints. There are considerations of opportunity costs, risk aversion, information availability, agreements with third parties, and so on. Taken together, parties often agree on some division of losses in the event of termination. Indeed, the use of advance deposit, reservation fees, and sub-compensatory liquidated damages often reflects exactly such decisions regarding the allocations of losses. Taken to the case at hand, it is quite reasonable that the parties would find some allocation of losses that is not as drastic as 100% loss to one of them. Indeed, this possibility is so likely that Goldberg contends that it is exactly what the court (covertly) did here. In this case, the court offhandedly dismissed the claim for £25, thus effectively allocating the loss 2/3-1/3.

We can distinguish here between what the parties meant to say and what they really meant to do. Judged only on meaning, the textualist may take the absolute obligation approach; and the contextualist might read an implied

[23] It is difficult to know with any certainty, but the court may have erroneously confused a completed ride to a booking ahead in time. The modern standard in carriage contracts allows some degree of flexibility for both parties—Uber permits passengers to cancel rides scheduled for the future and airlines often permit rescheduling of a flight for a premium. But I take here the court at its word regarding English customs.

condition into the contract. But both views would focus on the *meaning* of the agreement. If we knew the parties well enough, we might be able to offer better interpretations of their agreements based on our understanding of their *purposes*. What did they actually sought to achieve, *given* their personal circumstances, preferences, and constraints? With enough information, we could well deduce that a given allocation is more or less likely in the Krell case.

This is an illustration of a subtle difference between interpretation that is set on understanding the meaning of the agreement and one that is focused on determining its purpose. Yet, courts and commentators often blend the two modes, presumably because they often overlap (Klass, 2018). It is not very surprising that Williston defines interpretation as the task of finding word-meaning.[24] But it is more surprising that Corbin, who holds a more purposivist view, also defines interpretation in terms of word-meaning: "[b]y 'interpretation of language' we determine what ideas that language induces in other persons" (Corbin, 1950, at § 534). Both do not feel much tension there between these two definitions, because it is natural for people to use words that mean what they want to express. Indeed, Williston's strong emphasis on plain language is rooted in the idea that parties' actual purpose is to follow the plain language meaning of their agreement, so the meaning and purpose overlap (Williston, 2008). But common as it is, the Krell case reminds us that the two are critically distinct.

To further hone this distinction and its implications, consider a hypothetical developed by Professor Bagchi in the context of the evergreen debates whether contracts are moral promises or bare legal agreements. Bagchi describes a case of a person, let us call him Bob, who is scheduled to return home from a flight and requires a ride home from the airport. She then posits two scenarios. In the first, Alice, his close friend, promises that she will pick him up. In the second, Charlie a professional driver promises to pick Bob up for $50, the common market rate.

Bagchi's primary point in this hypothetical is that the parties' obligations differ between these scenarios. Alice, but not Charlie, is permitted to

[24] WILLISTON, 2008, at 1159-60 ("Interpretation is the art of finding out the true sense of any form of words: that is, the sense which their author intended to convey, and of enabling others to derive from them the very same idea which the author intended to convey.")

unilaterally cancel when she hears from a third party that Bob has met his romantic love interest on the plane and that they offered Bob a ride back:

> *"Should [Alice] be absolutely confident that [Bob] prefers the new arrangement, it would not be a breach of [Alice's] promise for her to leave a message for [Bob] simply informing [Bob] of the new plan. [Bob's]'s actual consent is not important where there is no uncertainty about [Bob's]'s understanding of her interests."* (Bagchi, 2011, at 720)

Bagchi goes on to make a distinction between intimates and strangers and the set of obligations each relationship entail. But this is only related to the analysis here because of the following logic:

> *"Because intimates know more about each other, they can more reliably assess and act on a richer account of each other's evolving interests; to the extent this holds true, they can adopt and continually update an ex-post view."* (Bagchi, 2011, at 722)

In the terms developed here, the distinction between strangers and intimates is related to predictive accuracy. Alice is allowed to account for Bob's purposes because she can accurately predict them, whereas Charlie can be trusted to make inferences that go beyond the wording of the agreement.

What animates Alice's analysis and lends reliability to her inferences, and what distinguishes her from Charlie, is her ability to *simulate* Bob based on her understanding of his preferences and purposes. That is, she has a model of her friend that is sufficiently nuanced that she can probe it and assess how would Bob react to a given scenario. She will be guided in her judgments by past interactions – "Oh, I remember how Bob was angry when his plans with Dorothy changed at the last minute" – and she will try to generalize from them to the present case. And just as her past interactions with her friend will serve as a repository of data, her care for him suggest that she will conduct the simulation with fidelity.

In contrast, Charlie would have little to draw on, and self interest might cloud his simulation of Bob. Charlie might be able to make some guesses based on his experience with passengers and with Bob's voice over the phone, but these will be quite general.

If Alice can preform this task well—and in this hypothetical, this would be the case—she is subject to a different set of normative obligations and expectations than Charlie.

In short, meaning-based interpretation is hermeneutic activity, whereas purpose-based interpretation is a simulative activity. And while the practice of the former is well analyzed in the scholarship, the latter is poorly understood and left mostly to intuitive guidance. Notably, Bagchi's example is rooted in intimates. But there are instances, such as the Krell case, where even strangers would be able to simulate the other with high accuracy. It would have taken little effort to deduce that, should the even be terminated or postponed, Henry would have little use for the apartment.

The concept of simulation sounds alien and unduly technological at first, and I can imagine some readers put off by these aspects. But those who persevere will find it to be a deeply humanistic exercise. After all, every instance of empathy is an exercise in simulating a mind different than our own. We anticipate how people would react to our gestures, invitations, and speech all the time. We try to predict which comment might offend and which one would give joy. We tailor our apologies to the other person's perceived needs. When we sell or persuade, we try to emphasize how our ideas or products fit within the plans and goals of our interlocutor. In all these instances, we create a model of the person in front of us and use that to simulate their reactions.

These simulative skills are quite organic. They tap into an innate human capability, the theory of mind, which allows us to see the world through another's eyes (Wellman, 2011). It is a skill that children normally develop at a young age. Because it is a skill, it means that people can differ quite significantly in their ability (id, at 259). But generally speaking, the predictive accuracy rises when we know the other person well. We find therefore special meaning in a carefully selected gift, because it delivers more than the object, also a credible signal that the other person knows us well enough to anticipate our preferences and goals.

This brings us back to Bagchi's hypothetical. It shows dual normative systems: one where simulative activity is unreliable (strangers) and one when it is reliable (intimates). When simulations are unreliable, the parties would want to rely on criteria that are less person-dependent such as the writing and its meaning. Because Alice's model is reliable, she can safely ignore the text

of her correspondence with Bob. Yet, Charlie is bound by the text as it gives greater predictive accuracy in his case.

The overarching idea is that the difference between meaning-searching and purpose-simulation goes deeper than a different point of emphasis. These are two different 'models' with different objective functions and with corresponding differences in their evidentiary 'datasets.' The orientation of meaning-searching is finding goals in how express themselves, the orientation of simulation is to find people's goal in what they wanted to express.

My normative claim, which I will make but not fully defend here, is that search for purpose is more important for accurate contract interpretation than the search for meaning.[25] Words are downstream of purpose, and if there is a mismatch, the general normative commitments of contract law (as opposed to say insurance contracts) mean strong preference for purpose over meaning. The Krell case is again illustrative of this point, as it shows how courts can sometimes ignore clear meaning altogether.[26] Purpose simulation also offer a coherent approach to gap filling. Trying to use literal approaches to even identify when a gap exists or when the contract is intentionally silent is notoriously difficult, as illustrated by one court finding gaps based on whether there is a "yawning void" in the contract.[27] Against this perplexing standard, purposive interpretation can offer concrete guidance.

The main challenges to purposive-simulation is that it relies on too much data about the parties, requires too much of the adjudicator, restricts third-parties' ability to rely, and ultimately results in inaccurate predictions. These complaints are reasonable: it is hard for a sitting judge to develop a nuanced understanding of the parties, their context, and their industry. The parties will offer conflicting versions and motivated evidence, and this can confound the analyst further. In the terms developed here, it means that 'simulation' rather than 'interpretation' is a predictive model that presents

[25] This is aligned with Karl Llewlyn's attempt to introduce purposive interpretation into the UCC itself. (Maggs, 2000). I note that there are other reasons, such as third party reliance, that would support more textual methods of interpretation.

[26] There are other ways to understand the Krell case. One such alternative is to understand contract excuses as tools of supplementation rather than interpretation. But even this is not availing, because excuses are default rules, and if the proper interpretation of the contract is that the parties did not want to excuse certain behavior then, as §154 Restatement (Second) of Contracts show, the excuse cannot be supplemented.

[27] Bidlack v. Wheelabrator Corp., 993 F.2d 603, 608 (7th Cir. 1993)

both problems of precision (the interpreter's sense of purpose may deviate largely from one to another) and accuracy (can predictions of purpose based on evidence better predict the parties' intent than contractual text).

Yet, these objections do not foreclose the viability or desirability of simulative interpretation. As the Krell case demonstrates, simulating parties answers to contractual questions can yield results that are predictable, information-efficient, sensible, and accurate. Conversely, it is the accurate parsing of the text that can lead to violations of parties' purposes. Bagchi's point regarding intimates is useful in showing intuitively how, despite such objections, simulation is a useful common activity. And cases like Nanakuli further show how even in commercial contracts, courts feel more confident moving from language to purpose.[28] Such cases show that it is realistic, despite all the various constraints, to incorporate purpose and parties' preferences to contract interpretation. Perhaps most importantly, these objections respond to simulation as practiced today, but what if we could improve it?

As indicated earlier, Posner once suggested that we should tie interpretative methods to judges' commercial expertise (Posner, 2014). More commercially experienced judges should rely more on commercial purposes. This points to the relevance of the adjudicator's training and would be one way to think about improving predictive accuracy for simulative exercises.

Another possibility is to think differently about the relevant evidence. To effectively model language, the relevant evidence will be linguistic. But inferring purpose involves a much broader inquiry, one that includes thick information about the parties, their commercial settings, and importantly, details about choices they have made throughout the negotiation process. Sophisticated parties recognize this. They try to improve simulative accuracy when they provide recitals and declarations of purpose (Scott & Triantis, 2006).[29] They work to improve linguistic accuracy through the inclusion of elaborate definition sections in their agreements.

The parties, however, do not always include all the relevant data and this leaves the adjudicator with the need to integrate large amounts of data. This puts an upper limit on what we can expect judges or juries to realistically do.

[28] Nanakuli Paving & Rock Co. v. Shell Oil Co. 664 F.2d 772 (1981)

[29] I disagree that text alone is not indicative of purpose, as in Greene, 1939 ("I]f it is actual intent we are after,... [w]hy then... does the court exclude all evidence but that of the writing itself, which is necessarily inadequate evidence of motive and intention?")

Arbitrators embedded in an industry may be able to execute it with some accuracy, but such an inquiry is likely too burdensome outside of specific domains.

But consider that we witness today the embryonic stages of the integration of language models into the adjudicative process. It is hard to know what role such models will play in future cases, but it is becoming quite clear that, at the very least, they will be used extensively to summarize and organize information. Whether they can be trained or fine-tuned to learn the features of specific industries and even specific parties remains to be seen. Still, even in their more modest role as helpful summarizers we can already anticipate a change in the balance of constraints on data processing. This means that, inasmuch as the limits on purposive interpretation are adjudicator subjectivity and data processing limits, these assumptions are ripe for disruption. In this case, we could expect future interpretative methods to move away from meaning and towards purposes.

To summarize the ideas so far, purposive interpretation and meaning-oriented interpretations are different interpretative modalities. As they seek to predict different measures, they involve different models and most importantly, different inputs. The process of discerning purposes is through mentally simulating the parties and their hypothetical resolution of specific transactional questions. There are some good objections that modeling meaning is more stable—more precise and more accurate—than modeling purposes. Yet, these objections rely on traditional adjudicative technology. As artificial intelligence grows in its ability to process and model data, these limitations will grow weaker and weaker. Future trends will arc towards purposive interpretation which, in my view, is a superior interpretative goal.

An appropriate conclusion to this section is the observation that contract specificity is a function of interpretation technology (and vice versa). If we can improve purposive interpretation techniques, this will make it easier and cheaper for people to enter into transactions. Of course, it is essentially free to enter into a contract today: but it is very expensive to enter into a good one. For a good contract, the parties must deliberate among many states of the world and use careful language and expensive lawyering. Advances in purposive interpretation would lower barriers. These lower barriers could allow people to organize their lives in novel, complex ways. Perhaps this is all a reinvention of relational contracting, but if so, it is one that can be done at scale.

CONCLUSION

Contract interpretation is the task of predicting parties' intentions across time. As a predictive task, it shares much in common with other disciplines that try to build predictive models. While I emphasized the similarity between the legal task and work in machine learning in this Chapter, it should be taken for granted that the different enterprise involve different normative considerations, sociological norms, and legal constraints. But even after accounting for these differences, it is still an open puzzle why legal prediction has been so disinterested in evaluation of its techniques. Legal interpretation largely proceeds in an unempirical fashion, and only recently scholars have started turning attention into the matter (Listokin, 2010). I hope that the theoretical emphasis here on the predictive components of the legal task and the sharpening of the conceptual tools would invite more work into validating our interpretative techniques. How accurate is textualism? How precise is contextualism? What is the level of bias under both? By having the language to ask, we might start answering these questions.

Bibliography

1. Arbel, Yonathan and David Hoffman. “Generative Interpretation.” New York University Law Review (Forthcoming, 2024).
2. Bagchi, Aditi. "Separating Contract and Promise." Florida State University Law Review 38 (2011): 709-758.
3. Barnett, Randy E. "The Sound of Silence: Default Rules and Contractual Consent." Virginia Law Review 78, no. 4 (1992): 821-911.
4. Bridgeman, Curtis. "Default Rules, Penalty Default Rules, and New Formalism." Florida State University Law Review 33, no. 3 (2006): 683-720.
5. Corbin, Arthur Linton. *Corbin on Contracts*. 1st ed. St. Paul, MN: West Publishing Co., 1960.
6. DiMatteo, Larry A. "Reason and Context: A Dual Track Theory of Interpretation." Penn State Law Review 109, no. 2 (2004): 397-486.
7. Eisenberg, Melvin A. *Foundational Principles of Contract Law*. New York: Oxford University Press, 2018.
8. Fish, Stanley. "Normal Circumstances, Literal Language, Direct Speech Acts, the Ordinary, the Everyday, the Obvious, What Goes without Saying, and Other Special Cases." Critical Inquiry 4, no. 4 (1978): 625-644.
9. Geman, Donald, Edir Bienenstock, and René Doursat. "Neural Networks and the Bias/Variance Dilemma." Neural Computation 4, no. 1 (1992): 1-58.
10. Gilson, Ronald J., Charles F. Sabel, and Robert E. Scott. "Text and Context: Contract Interpretation as Contract Design." Cornell Law Review 100, no. 1 (2014): 23-98.
11. Goldberg, Victor P. "Impossibility and Related Excuses." Journal of Institutional and Theoretical Economics (JITE) / Zeitschrift Für Die Gesamte Staatswissenschaft 144, no. 1 (1988): 100-116.
12. Greene, Marjorie. "Theories of Interpretation in the Law of Contracts." The University of Chicago Law Review 6, no. 3 (1939): 374-394.
13. Klass, Gregory. "Interpretation and Construction in Contract Law." Georgetown Law Faculty Publications and Other Works, 2018. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2913228 .

14. Leech, Geoffrey, and Jenny Thomas. "Language, Meaning and Context: Pragmatics." In *An Encyclopedia of Language*, edited by N.E. Collinge, 173-206. London: Routledge, 1989.
15. Linzer, Peter. "The Comfort of Certainty: Plain Meaning and the Parol Evidence Rule." Fordham Law Review 71, no. 3 (2002): 799-836.
16. Listokin, Yair. "The Meaning of Contractual Silence: A Field Experiment." Journal of Legal Analysis 2, no. 2 (2010): 397-416.
17. Maggs, Gregory E. "Karl Llewellyn's Fading Imprint on the Jurisprudence of the Uniform Commercial Code." University of Colorado Law Review 71, no. 2 (2000): 541-588.
18. Martinez, Eric, and Kevin Tobia. "What Do Law Professors Believe About Law and the Legal Academy?" Georgetown Law Journal 112 (forthcoming 2023): 1120189.
19. McElroy, R. G., and Glanville Williams. "The Coronation Cases—I." Modern Law Review 4, no. 4 (April 1941): 241-60. https://doi.org/10.1111/j.1468-2230.1940.tb00777.x.
20. Mitchell, Catherine. *Interpretation of Contracts*. 2nd ed. London: Routledge-Cavendish, 2019.
21. Posner, Richard A. "The Law and Economics of Contract Interpretation." Law & Economics Working Paper No. 229, John M. Olin Program in Law and Economics, University of Chicago Law School, 2004.
22. Saussure, Ferdinand de. *Course in General Linguistics*. Edited by Charles Bally and Albert Sechehaye. Translated by Wade Baskin. New York: Philosophical Library, 1959.
23. Schwartz, Alan, and Robert E. Scott. "Contract Interpretation Redux." The Yale Law Journal 119, no. 5 (2010): 926-964.
24. Schwartz, Alan, and Robert E. Scott. "The Limits of Contract Law." Yale Law Journal 113 (2003): 541, 573.
25. Scott, Robert E. "The Rise and Fall of Article 2." Louisiana Law Review 62, no. 4 (2002): 1009-1064.
26. Scott, Robert E., and George G. Triantis. "Anticipating Litigation in Contract Design." The Yale Law Journal 115, no. 4 (2006): 814-879.
27. Shackel, Nicholas. "The Vacuity of Postmodernist Methodology." Metaphilosophy 36, no. 3 (2005): 295-320.
28. Shalev-Shwartz, Shai, and Shai Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. New York: Cambridge University Press, 2014.

29. Silverstein, Joshua M. "The Contract Interpretation Policy Debate: A Primer." Stanford Journal of Law, Business & Finance 26, no. 2 (2021): 222-294.
30. Stempel, Jeffrey W., and Erik S. Knutsen. "Rejecting Word Worship: An Integrative Approach to Judicial Construction of Insurance Policies." University of Cincinnati Law Review 90, no. 2 (2021): 561-636.
31. Thayer, James Bradley. *A Preliminary Treatise on Evidence at the Common Law*. Boston: Little, Brown, and Company, 1898.
32. Wellman, Henry M. "Developing a Theory of Mind." In *The Wiley-Blackwell Handbook of Childhood Cognitive Development*, edited by Usha Goswami, 258-284. 2nd ed. Malden, MA: Wiley-Blackwell, 2011.