Differentiated Decision-Making as a Basis for Competitive Advantage

INFORMS Advances in Decision Analysis — Milan, Italy, June 19-21, 2019

Neil A. Hamlett, D.Sc., MBA

email: neil.hamlett@uncertainty-research.science https://www.linkedin.com/in/neil-hamlett-strategic-quant/

Abstract

Principles from Information Economics lead to a conjecture that three distinct modes exist for competitive differentiation based on information. Differentiated access affords informational asymmetries to competitors who enjoy it. Marketplace participants see information in evidence that competitors miss through differentiated insights. Differentiated ability to act occurs in superior abilities to transform information into utility. This can be characterized by a formalism with Decision Analysis at its core. Other disciplines including theory of evidence, information theory, and competency-focused frameworks for differentiation also contribute. The theory of evidence produces a more-complete characterization of uncertainty than traditional Bayesian techniques. Information theory provides for a rigorous, abstract approach to quantifying information and uncertainty. Mining a textbook case study exercises the formalism to quantify returns from information, and to analogously corroborate our conjecture.

1 Introduction

That we operate in an "Information-Age Economy"has been a cliché for a generation ([Shapiro1998]). Peter Drucker anticipated as early as the late 1950s the emergence of *knowledge work* ([Wartzman2014]). He subsequently foresaw that, "knowledge ⟨becomes⟩ a more-crucial economic resource than land, labor, or financial assets," ([Drucker1992]) leading to what he called a "post-capitalist society" ([Drucker1993]).

Porter and Millar ([Porter1985]) observed "the information revolution sweeping through our economy" in the mid-1980s. They provided heuristic descriptions of informational impacts in terms of existing value-chain frameworks for the analysis of strategy. Although this early work did not anticipate early-21st-century emergent operating models, likely strategic impacts were described. These included improved operational efficiencies, more-tailored responsiveness to market pulls, and adaptively shifting competitive scope.

Uncertainty is the opposite side of the information and knowledge coin ([Briggs2016]). It is an inescapable aspect of business environments. Organizations' approaches to uncertainty can lead either to differentiation or to demise ([Lynch2017]). Entirely new business models have emerged focusing on curation and exploitation of information and data ([Redman2015]). Such models that "have as their business the production and distribution of knowledge and information, rather than the production of things, ... have moved into the center of the economy" ([Drucker1993]).

Decision analysis (DA) is a core business competency guiding the transformation of information and uncertainty into economic value ([Spetzler2016]). The centrality of information's role in the economy amplify's DA's importance. If information is indeed a more-important business asset than land, labor, and financial assets historically have been, then DA's position perhaps approaches other more-prominent business-management disciplines such as operations and supply-chain management. The ability to parse information and uncertainty — and efficiently and effectively act thereupon — becomes an important determinant of competitiveness.

We undertake here the articulation of a formalism for the characterization of information's contribution to competitive differentiation. We begin with a conjecture arising from Information Economics. Specifically, three distinct modes of information-based advantage occur. Differentiated access arises from information others lack. Differentiated insights result from seeing things in the evidentiary corpora overlooked by others. Superior capabilities to extract value from information possibly available to others exemplifies differentiated ability to act. Organizations distinguishing themselves via information exploitation apply these modes individually or in combination to realize superior return on information.

Our formalism is grounded in a *Generalized Decision-Analysis Problem*. A novel formulation introduced here emphasizes attributing value to information. Specifically, the optimum decision produces the greatest expected net return of all options given a set of consistently-framed scenarios directly coupled by alternative courses of action and probabilistically-determined outcomes. These courses of action represent distinct decisions by active decision-makers in our scenarios. They are deterministic branches in contexts within which many aspects of outcomes are probabilistic.

Mechanically, our approach employs the *Theory of Evidence* (ToE), a generalization of the commonly-employed Bayesian-analysis approach to DA. Two aspects of ToE recommend themselves to our endeavor. First, traditional Bayesian theory under many circumstances can lead to overstatement of the degree of certainty underlying a scenario. Relaxing selected Bayesian axioms, ToE allows a mechanism of "partial belief", by which evidence can be treated as inconclusive. Secondly, traditional Bayesian theory can attribute preferential influence to prior probabilities. ToE combines evidence in a fashion that is more-symmetric.

We overlay our modes of differentiation onto strategy frameworks focused on organizational competencies. The Dynamic Capabilities framework by [Teece2009] and the Adaptivity-Advantage method by [Reeves2012] provide structure for exploring differentiating consequences of our modes of information-derived advantage. Each mode bears on distinct phases and components of these competency-based frameworks. This ties our information-differentiation modes to contexts of competitive differentiation.

A textbook case study related to petroleum-extraction decision illustrates the framework. A parameter study demonstrates the relationship between uncertainty — represented by either a Bayesian probability distribution or a ToE basic-belief assignment — and expected net-return from a consistently-framed decision scenario. Aspects of the case study corroborate the conjecture about modes of information-differentiation. Contrasts between expected-net returns employing Bayesian and ToE frameworks for uncertainty are apparent. This is corroborated by information-theoretic (entropy) comparisons. In particular, framing uncertainty using ToE leads lower expected-net returns. This is consistent with the assertion of greater perception of certainty when employing Bayesian approaches than with ToE.

This study defers consideration of key questions related to information-based differentiation. First, relative return on information provides the present criterion for informationally-derived advantage. This might not be sufficient for all scenarios. [Phlips1988] and [Laffont1989] examine market-equilibria conditions in information-driven markets. Extending our formalism to such competitive scenarios will produce a more-complete view.

Drucker's [Drucker1993] characterization of an information-intensive economy as "post-capitalist" represents another unaddressed complexity. Mainstream microeconomic theory explains a *material economy*, in which resource scarcity drives resource allocation. If information itself is a key asset, an *ethereal economy* alternatively occurs. Resource constraints in the ethereal economy tend to be talent-related rather than material. Moreover, information "products" can be simultaneously used by multiple customers without constraining the use by any individual.

How some microeconomic mechanisms play out in an ethereal marketplace is unclear. For example, price-elasticity, one of markets' principle mechanisms for self-regulation, results from degrees of scarcity and necessity in conjunction with each other. Scarcity, in particular, ceases to be a natural constraint in the case of ethereal goods. Our case study more-directly explores informational control over the material economy than over the ethereal.

Finally, our framework at this juncture remains epistemically imprecise. We knowingly conflate distinct entities such as knowledge, information, competency, and data. A more-rigorous epistemic framework will improve the descriptive power of the formalism.

2 Toolkit for Characterizing Differentiated Decision-Making

We introduce here our framework for characterizing opportunities for competitive advantage. It fuses together five distinct practices or disciplines. We explain their integration progressively as we introduce each. A coherent, holistic description of information's contribution to differentiation occurs. The DA discipline occupies a central, unifying role. These introductions also contain extensive surveys of prior relevant literature.

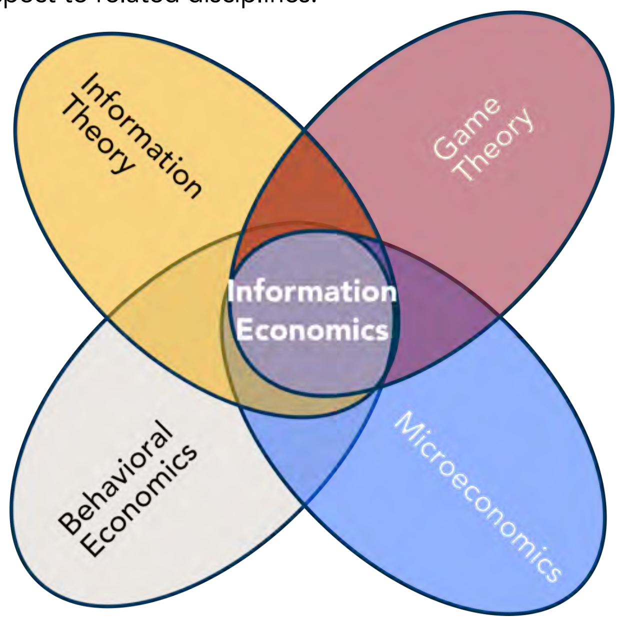
2.1 Information Economics

The field of Information Economics studies the attribution of value to information [Birchler1999, Brunnermeier2001]. It is most-commonly applied within the context of financial markets. It explains for example how bid-ask prices reveal market counterparties proprietary positions with respect to risk and other objectives. Common theories of financial-market operations, such as the Efficient-Market Hypothesis, are among the more-familiar applications of information economics.

The discipline of information economics resides at the nexus of four other disciplines. Figure 1 depicts this intersection. First, information theory provides mechanisms to precisely quantify information. This includes two loosely-related branches. Quantitative information theory was originally developed for communications theory [Shannon1948, Gallagher1968, Cover1991]. Borrowing concepts from the physical discipline of thermodynamics [Ben-Naim2008], it provides an abstract view of information as a quantity. Semantic information theory [Chomsky2014, Zhong2017] explains linguistic and psychological aspects of information.

Game Theory [Myerson1991, Papayoanou2010] explains counterparties' behaviors in competitive situations, often involving incomplete information. The "Elementary Game" assumes a prominent role in the treatment by [Birchler1999]. The elementary game postulates about counterparties' decisions whether or not to play a game based upon their informational states. Decisions are made based on expected outcomes in the presence of uncertainty.

Figure 1: Position of Information Economics with respect to related disciplines.



Concepts from microeconomics build upon information theory and game theory to explain equilibria conditions within markets involving multiple counterparties'. As usual, equilibrium is reached when bidding and asking counterparties' positions converge to mutually agreed-to price points. Information-economics theory here does not distinguish between the etherial and material markets referred to in the introduction.

Finally, behavioral economics has recently begun to influence information economics. One of the field's "founders" D. Kahneman describes in [Kahneman2011] the use of decision games resembling the Elementary Game to characterize participants' understandings of probabilities. Prospect theory [Kahneman1979] formally extends the resulting concepts into psychological explanations for decision makers' views of risk and utility.

Information modeling [Samuelson2004] features prominently within the information-economics toolkit. Information modeling employs concepts resembling those at the heart of decision analysis and theory of evidence, to be visited subsequently. First, it draws heavily from Bayesian theory, also foundational for decision analysis. Its representation of information dovetails well into ToE's framework. Both support ambiguity in states of belief. The machineries of the respective methods are well-positioned for integration and convergence.

Now a central information-economics principal leads to the fundamental conjecture of this study. Specifically,

> "The utility value of a piece of information is the difference between the expected utility a decision-maker can achieve by choosing the best action *conditional* on the information received, and the expected utility from the action that is best in the *absence* of the information. The money value is the amount of money that renders the decision-maker *indifferent* to both." [Birchler1999]

Our three modes of differentiation occur either explicitly or by implication. Differentiated access — possessing a piece of information that a counter party lacks — explicitly appears. A counter party with a *relevant* information advantage by extension enjoys an opportunity-realization advantage. This is information asymmetry.

Information-insights differentiation appears by implication. Both counterparties might share access to the same evidentiary corpus — or body of information — about the decision with which they are confronted. One however sees something in that evidentiary corpus that one or more counter party does not.

Michael Lewis' *Moneyball* [Lewis2004], a mid-2000s popular-nonfiction treatment American professional baseball, illustrates. The featured club employed a distinct model for attributing value to individual baseball players contributions. This differentiated insight conferred considerable advantage in playing competition.

Differences in insights from a common evidentiary corpus may arise from a variety of circumstances. The field of organizational epistemology (e.g., [vonKrogh1994, Seirafi2013]) systematically examines how organizations translate information into actionable knowledge. Prospect-theoretic considerations may inhibit — or sometimes even fortuitously enhance — an organizations' ability to recognize value opportunities. [Denrell2001] describes the "Hot-Stove" affect, under which risk-aversion inhibits organizations' abilities to learn. [Pisano2019] makes passing mention of how "innovative" organizations might cultivate counterproductive "knowledge" through undisciplined, indiscriminate learning.

Ability-to-act differentiation also appears explicitly. Different market participants enjoy different utility-realization capabilities. A distinct branch of business strategy studies organizations' abilities to effectively respond to new information. Teece's Dynamic Capabilities framework [Teece2009] and Reeves' Adaptivity Advantage [Reeves2012] exemplify. For instance, an organization might avail itself of an opportunity for differentiated return by responding first to information about an opportunity. Alternatively, another might realize superior return through responding more-efficiently.

2.2 Decision Analysis

Decision analysis (DA) is a business discipline focused characterizing the quality of decisions within the context of uncertainty [Spetzler2016]. Conceptually congruent with Information Economics, DA seeks to identify discrete actions yielding the optimum return in a precisely-framed scenario.

Our consideration of DA here begins by proffering a generalized statement of its central problem. Given this study's central conjecture, we emphasize attributing value to information driving decisions. We then review a methodology that enjoys near-paradigmatic status in the DA-practitioner community. Finally we give consideration to the implications of behavioral economics for DA.

2.2.1 The Generalized Decision-Analysis Problem

A generalized mathematical statement articulates the DA problem. Specifically,

$$\mathfrak{D}^{\text{opt}} = \arg\max_{\mathcal{D}} \Bigg\{ \Big(\sum_{\substack{x \in \mathscr{X} \\ y \in \mathscr{Y} \\ z \in \mathscr{Z}}} \mathscr{P}(x, y, z; \mathfrak{D})\, \mathscr{R}(z; \mathfrak{D}) \Big) - \mathscr{C}(\mathfrak{D}) \Bigg\}, \quad \text{(1a)}$$

$$\mathcal{R}^{\text{opt}} = \sum_{\substack{x \in \mathcal{X} \\ y \in \mathcal{Y} \\ z \in \mathcal{Z}}} \mathcal{P}\big(x, y, z;\, \mathfrak{D}^{\text{opt}}\big)\, \mathcal{R}\big(z;\, \mathfrak{D}^{\text{opt}}\big)\Big) - \mathcal{C}\big(\mathfrak{D}^{\text{opt}}\big), \text{ and} \tag{1b}$$

$$\mathcal{V}_{\text{info}} = \mathcal{R}^{\text{opt}} - \mathcal{R}^{]\text{nominal}}. \tag{1c}$$

Here, $\mathcal{D}$ denotes the set of all possible decisions within our framed scenario, $\mathcal{R}$ the expected net return conditioned on the particular set of decisions $\mathfrak{d} \in \mathcal{D}$, and $\mathcal{C}$ is the costs incurred with acting according to that particular decision set. The uncertainty is characterized by probability distribution $\mathcal{P}$ in terms random variables $\mathscr{X}$, $\mathscr{Y}$, and $\mathscr{Z}$. Explicitly,

$$\mathcal{P} : \mathscr{X} \times \mathscr{Y} \times \mathscr{Z} \mapsto [0, 1]\,.$$

In terms of our conjecture, variable $\mathscr{X}$ is the additional information obtained through selection some decision $\mathfrak{d} \in D$ at a cost of $\mathcal{C}\left(\mathfrak{d}\right)$. The random variable $\mathscr{Y}$ represents the additional information resulting from differentiated insights. And $\mathscr{Z}$ is the inherent phenomenology determining the returns. We attribute $\mathcal{V}_{\text{info}}$ information value by comparing the return $\mathcal{R}^{\text{opt}}$ arising from $\mathfrak{D}^{\text{opt}}$, the optimum decision, to some nominal baseline.

The mechanisms suggested by $\mathcal{P}$ indicate how uncertainty is resolved. $\mathcal{P}$ can be thought of as a joint probability distribution in the traditional Bayesian sense. We see in the next section that under ToE things become more-complicated.

Nonetheless, adding information associated with differentiated access or insights allows one to consider returns in terms of conditional probabilities, $\mathcal{P}\left(z \mid x, y\right)$ for example. If our phenomenological variable $\mathcal{Z}$ is not independent of the other variables,

$$\mathcal{P} \nvDash \left(\mathcal{Z} \perp \mathscr{X}, \mathscr{Y}\right),$$

then conditioning often has the affect of reducing its uncertainty. This in turn reduces the variability of the range of possible outcomes, $\mathcal{R}$. We more-rigorously consider affects of conditioning in a subsequent section on information quantification.

Now, (1a) might induce in an operations analysts the inclination to apply optimization-search techniques to its solution. Although not impossible, such an attempt might prove challenging. The decision-optimization rule (1a) is discretely non-convex. The presence of the decision factors $\mathfrak{D}$ introduce discrete, deterministic branches within the probability space. Analysis cases must be conditioned on each distinct decisional configuration represented by set of decisions $\mathfrak{D}$.

Armed with the framework of (1a) – (1c), we can now think about differentiated returns associated with distinct modes of information-based differentiation. First consider the scenario of differentiated access. The increase in expected net returns arising from making some optimum set of decisions $\mathfrak{D}^{\text{opt}}$ straightforwardly becomes

$$\begin{aligned}\mathcal{R}^{\text{diff access}} - \mathcal{R}^{\text{nominal}} &= \\ &= \sum_{\substack{x \in \mathcal{X} \\ z \in \mathcal{Z}}} \mathcal{P}\big(x, z;\, \mathfrak{D}^{\text{opt}}\big)\, \mathcal{R}\big(z;\, \mathfrak{D}^{\text{opt}}\big)\Big) \\ &\quad - \sum_{z \in \mathcal{Z}} \mathcal{P}\big(z;\, \mathfrak{D}^{\text{nominal}}\big)\, \mathcal{R}\big(z;\, \mathfrak{D}^{\text{nominal}}\big)\Big) \\ &\quad - \mathcal{C}\big(\mathfrak{D}^{\text{opt}}\big) \\ &\quad + \mathcal{C}\big(\mathfrak{D}^{\text{nominal}}\big).\end{aligned} \tag{2}$$

At a marginal cost of $\mathcal{C}\left(\mathfrak{D}^{\text{opt}}\right) - \mathcal{C}\left(\mathfrak{D}^{\text{nominal}}\right)$ we realize an increase in marginal expected net-returns. Explicitly, our marginal return becomes

$$\mathcal{M}^{\text{diff access}} = \frac{\left(\begin{array}{l}\displaystyle\sum_{\substack{x \in \mathcal{X} \\ z \in \mathcal{Z}}} \mathcal{P}\big(x, z;\, \mathfrak{D}^{\text{opt}}\big)\, \mathcal{R}\big(z;\, \mathfrak{D}^{\text{opt}}\big)\Big) \\ \displaystyle - \sum_{z \in \mathcal{Z}} \mathcal{P}\big(z;\, \mathfrak{D}^{\text{nominal}}\big)\, \mathcal{R}\big(z;\, \mathfrak{D}^{\text{nominal}}\big)\Big)\end{array}\right)}{\mathcal{C}\big(\mathfrak{D}^{\text{opt}}\big) - \mathcal{C}\big(\mathfrak{D}^{\text{nominal}}\big)}.$$

A similar logic applies to differentiated insights. The difference being that margin for differentiated insights is undefined. Conditioning on $\mathcal{Z}$ does not cost us in the sense of incurring a cost associated with some decision $\mathfrak{d} \in \mathfrak{D}$.

The excess returns from differentiated ability to act arises from the return function $\mathscr{R}$ itself. Specifically,

$$
\begin{aligned}
&\mathscr{R}^{\text{diff ability}} - \mathscr{R}^{\text{nominal}} = \\
&= \sum_{\substack{x \in \mathscr{X} \\ z \in \mathscr{Z}}} \mathcal{P}(x, z; \mathfrak{D}^{\text{opt}})\, \mathscr{R}^{\text{differentiated}}(z; \mathfrak{D}^{\text{opt}})\Big) \\
&\qquad - \sum_{\substack{x \in \mathscr{X} \\ z \in \mathscr{Z}}} \mathcal{P}(x, z; \mathfrak{D}^{\text{opt}})\, \mathscr{R}^{\text{nominal}}(z; \mathfrak{D}^{\text{opt}})\Big).
\end{aligned}
$$

We assume here that both organizations operate given identical evidentiary corpora. The expend substantially similar amounts in acquisition of information. The difference arises from their ability transform information into utility.

2.2.2 A Common Decision-Analysis Method Employed within the Practitioner Community

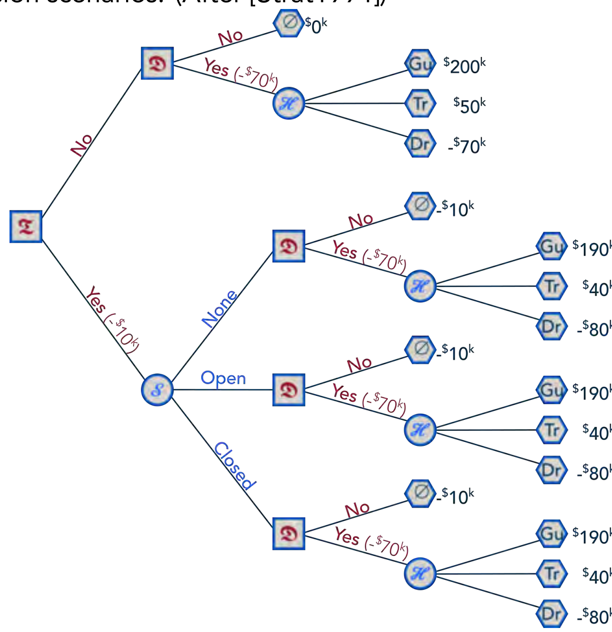
The DA-practitioner community methodologically realizes these concepts using techniques originally described separately by [Raiffa1970] and [Howard2016]. The associated "playbook" enjoys enjoy near-paradigmatic status. The common methodology focuses on decision-framing. Framing consists of analysis of distinct scenarios, which are coupled by discrete decisions and probabilistic chances.

Figure 2 illustrates. This example comes from a textbook petroleum-extraction case study to be considered in-detail, later. We focus here the structure represented by Figure 2 and explore its contents, later.

Decision trees are a commonly-used tool for framing of decisions. We construct decision trees from three distinct types of vertices (nodes). Square-shaped vertices represent deterministic branches in which representing a decision-maker's deliberate action. Figure 2 contains only binary — Yes-versus-No — decisions. This need not always be the case. The edges emanating from decision nodes are labeled with the decision options and with the *direct* costs associated each.

Round-shaped vertices are "chance" nodes. These represent discrete, probabilistic branches in our probability space. Graph edges emanating from chance nodes are labeled with the distinct state, and the probability of its occurrence. Hexagonal vertices are outcome vertices. These correspond to the final state in a

Figure 2: Illustration of decision-tree framing of decision scenarios. (After [Strat1994])



given scenario. Each is labeled with its corresponded expected net return.

Each left-to-right path through the tree — from initial decision to ultimate outcome — represents a distinct scenario. The scenarios are framed so that they are coupled together via deterministic decision nodes and probabilistic chance nodes. The left-most chance nodes are labeled with marginal probabilities. Proceeding left to right, each subsequent chance node is labeled with probabilities conditioned on the preceding chance-node branches. Employing classical Bayesian analysis, conditioning and marginalization occur according to standard practices in categorical-data analysis (e.g., [Agresti2002]). We discover in the next section that ToE analysis is more-complicated.

The costs and outcomes are represented according to Time-Value-of-Money (TVoM) principles (e.g., [Brealy2008]). Decision-framing requires that costs and returns be referenced to the same point in time. They therefore are net-present values (NPVs) referenced to point in time with meaningful business purposes.

Net-expected returns are aggregated right-to-left. Expected net returns at chance vertices are the probability-weighted NPVs of those adjacent to the right. Aggregation at decision nodes occurs by discretely selecting the right-adjacent vertex with the greatest corresponding net-expected return. To reiterate, the decision nodes introduce the discrete nonconvexity in the probability space of (1a).

2.2.3 Decision Analysis and Prospect Theory

The question might arise regarding the application of Behavioral Economics — Prospect Theory, in particular [Kahneman1979] — to decision analysis. Limited rationality of decision makers represents a topic of extensive contemporary interest. Why, for example, might or might one not factor prospect-theoretic considerations into (1a)? Four considerations are offered here.

First, one might consider attempting to incorporate an individual's or an organization's idiosyncratic utility functions into our return function $\mathcal{R}$. This in general would be difficult. Identification of idiosyncratic utilities would require challenging psychometric or sociometric analysis, for which the the net-improvement in informational value would be uncertain.

Second, within the analytical sciences, establishing objective performance bounds represents common practices. The use of the Cramer-Rao Lower Bound in statistical analysis represents an analogy (e.g., [Kay1993, Hogg2013]). Analysis following establishment of a theoretical best bound focuses on diagnosing root causes of variances. Addressable root causes are prioritized according to difficulty and significance. These could, of course, included prospect-theoretic challenges.

Thirdly, framing decisions in rational and objective terms constitutes the principal goal of DA. Kahneman modeled human cognition using two "systems" [Kahneman2011]. "System 1" characterizes an intuitive, reflexive mode of decision-making. "System 2" describes the deliberate, rational mode. Kahneman's essential conclusion was that human decision makers *are capable* of objective decision-making in "System-2" mode. This is however effortful and physiologically-demanding. DA is about forcing decision-makers into "System-2" modes for decisions that are sufficiently important.

Finally, defining a formalized methodology for comparative *return on information* represents the principle focus of our central conjecture and its exploration. Understanding obstacles to achieving objectively-identified, theoretically-best outcomes is related aspect. Many individuals and organizations no doubt encounter barriers arising from prospect-theoretic challenges. Diagnosing these — and if possible, remedying them as well — represents a potential outgrowth of this study. The diagnostic process requires separation of prospect-theoretic challenges from theoretically-optimum outcomes. We first establish the later using the framework of (1a)-(1c).

2.3 The Theory of Evidence

The Theory of Evidence (ToE) — also referred to a Dempster-Shafer Theory — seeks to represent uncertainty in a more-general fashion than the traditional Bayesian approach to probability [Shafer1976, Smets1994, Salicone2018]. It accomplishes this principally by expanding the domain over which probability masses are assigned. ToE can be formulated as a generalization of Bayes Theorem [Smets1993].

Originally articulated as a complete theory by Glen Shafer [Shafer1976], ToE was studied extensively as a probabilistic-reasoning framework for artificial-intelligence applications [Smets1994]. Its application to DA was also described [Strat1994, Smets2002]. The metrology community more-recently adopted a ToE framework as a replacement for the "accuracy-and-error" paradigm for characterizing uncertainty in scientific and engineering measurements [Salicone2018]. Commonalities with fuzzy logic have also been observed [Skalna2015], portending the potential for convergence of the two disciplines.

Finally, ToE seems apropos due to similarities it shares with information modeling [Samuelson2004] from Information Economics. A full reconciliation of the respective methods lies beyond the scope of the present study. Nonetheless, we lay the groundwork here for such a subsequent effort. The resulting payoff would be a tighter alignment of DA with Information Economics, leading to more-general approaches to information-valuation.

The following seeks to imbue an intuitive sense through illustrating the pattern of reasoning under ToE. We next concisely survey its mechanics. We then review approaches to its applications to decision analysis.

2.3.1 Probabilistic Reasoning under the Theory of Evidence

Consider Figure 3 for purposes of illustration. The graph represents a notional *Frame of Discernment* (FoD) corresponding to a two-dimensional, binary probability space. The FoD encompasses all propositions to which measure of belief can be attributed. Distinct propositions to which finite (nonzero) belief is attributed are *focal elements*. The *core* includes all such focal elements.

How does ToE differ from Bayesian theory? In the latter we can only attribute belief to the distinct "atomic" propositions. In Figure 3 this is constrained to

$$(\mathscr{X},\mathscr{Y}) = \Big\{\{(x,y)\},\{(\bar{x},y)\},\{(x,\bar{y})\},\{(\bar{x},\bar{y})\}\Big\}.$$

Within such a regime, complete uncertainty is expressed as

$$\mathcal{P}\Big(\{\xi,\eta\}\Big) = \frac{1}{|(\mathscr{X},\mathscr{Y})|} \quad \forall\,\{\xi,\eta\} \in (\mathscr{X},\mathscr{Y}).$$

For the illustration of Figure 3 — a binary, two-dimensional probability space — the cardinality $|(\mathscr{X},\mathscr{Y})| = 4$. Consequently, we assert that

$$\mathcal{P}\Big(\{\xi,\eta\}\Big) = \frac{1}{4} \quad \forall\,\{\xi,\eta\} \in (\mathscr{X},\mathscr{Y})$$

under complete uncertainty.

Reasoning under ToE admits that a given evidentiary corpus might not support any of the propositions in $(\mathscr{X},\mathscr{Y})$ with a degree of confidence anywhere near $\frac{1}{4}$. It therefore provides for a mechanism called *partial belief*. A closely-related concept of *vacuous belief* represents the most-extreme form of partial belief. Vacuous belief articulates a state of belief along the lines of, "One of the above but cannot tell which." Partial belief ranges across various degrees of vacuity, ranging from complete certainty that one of the "atomic" propositions is true, to the complete vacuity referred-to above. Smets [Smets1988, Smets1994] includes a provision for "None of the above", $\emptyset$. This lies beyond our present scope.

Returning to the illustration in Figure 3, the *basic belief assignments* are

$$\mathcal{P}\Big(\{\,(x,y)\,\}\Big) = \mathcal{P}\Big(\{\,(x,\bar{y})\,\}\Big) = 0.1,$$
$$\mathcal{P}\Big(\{\,(\bar{x},y)\,\}\Big) = \mathcal{P}\Big(\{\,(\bar{x},\bar{y})\,\}\Big) = 0.3, \text{ and}$$
$$\mathcal{P}\left(\begin{Bmatrix}(x,y)\,,(x,\bar{y})\,,\\(\bar{x},y)\,,(\bar{x},\bar{y})\end{Bmatrix}\right) = 0.2.$$

We thereby interpret the evidence as supporting the proposition "One of the above but cannot tell which" with probability of 0.2.

This brings us to a key aspect of reasoning under ToE. Our frame of discernment — the total set of propositions to which one might assign some degree of belief — is the power set $\wp$ of the atomic propositions. In our Figure 3 illustration, or *Bayesian* frame of discernment is

$$\mathscr{X}\times\mathscr{Y} = \begin{Bmatrix}(x,y)\,,(x,\bar{y})\,,\\(\bar{x},y)\,,(\bar{x},\bar{y})\end{Bmatrix}.$$

The ToE frame of discernment in contrast is specified by

$$\underline{\Omega}^{\mathscr{X},\mathscr{Y}} = \wp\Big(\mathscr{X}\times\mathscr{Y}\Big)$$

and includes all vertices in the directed graph in the Figure 3.

Now, the traditional Bayesian-probability scenario is simply a special case of ToE. [Smets1993] developed a *Generalized Bayesian Theorem* (GBT). The GBT extends the familiar framework of prior probabilities, likelihoods, and posterior probabilities to the more-general FoD $\underline{\Omega}$.

It is moreover accompanied by a *Conditional Cognitive Independence* framework for ascertaining the belief induced by "independent" observations. The concepts bear some resemblance to work in [Lindley1983], describing the reconciliation of probability distributions.

The Bayesian frame of discernment represents only one specialization of the ToE frame of discernment. Fuzzy logic employs a method in which the frame of discernment is monotonic [Chow1997, Skalna2015, Salicone2018]. That is,

$$\mathcal{X}_i = \{x_1,\dots,x_i\}, \text{ and}$$
$$\mathcal{X}_1 \subset \mathcal{X}_2 \subset \dots \mathcal{X}_n \equiv \mathscr{X}.$$

This framework is referred to as *Possibility Theory*. Its resemblance to ToE has led to movement towards the convergence of the two disciplines.

Figure 3: Illustrative probability-mass assignment over a frame of discernment under the Theory of Evidence (after [Jiroušek2018]).

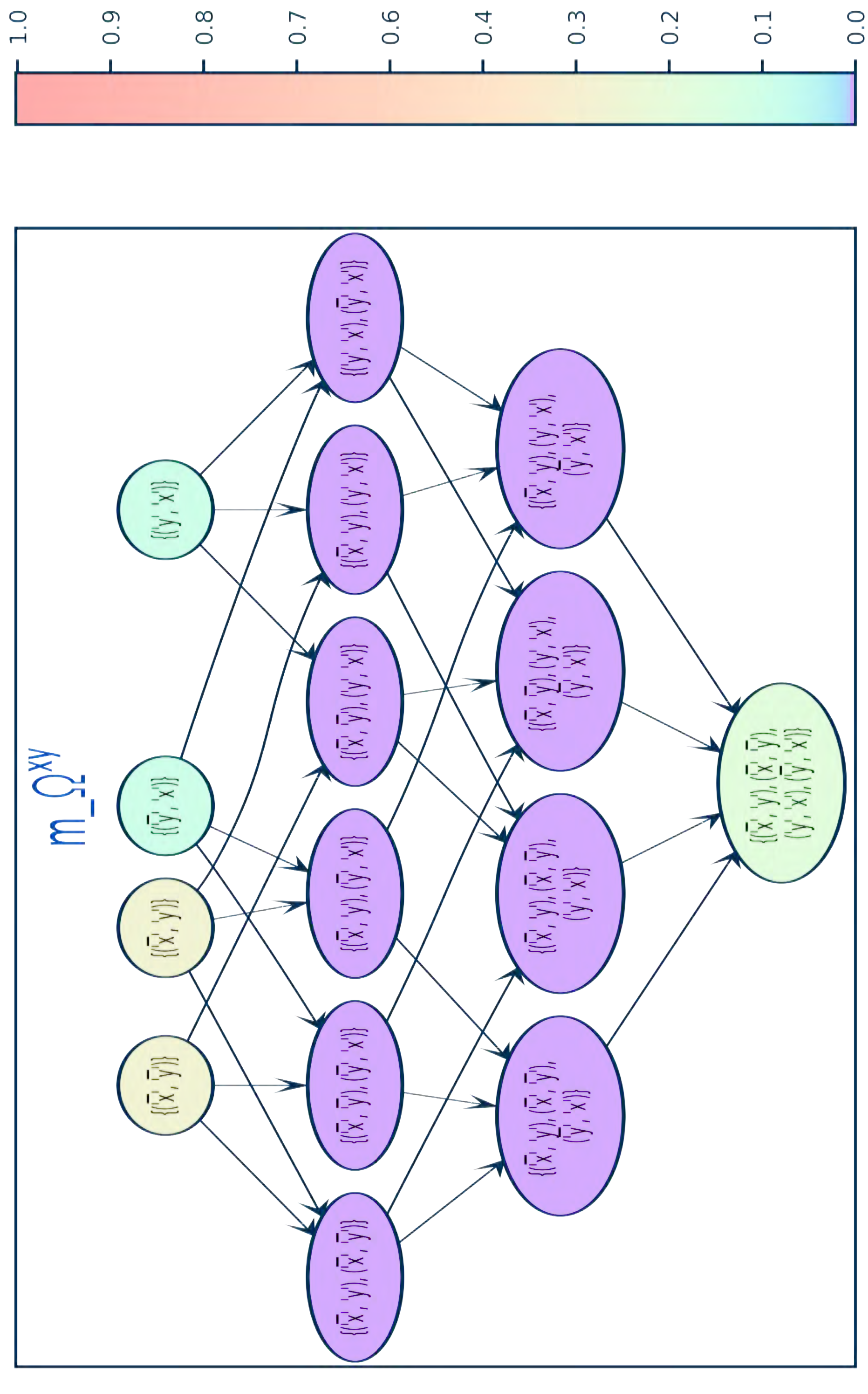
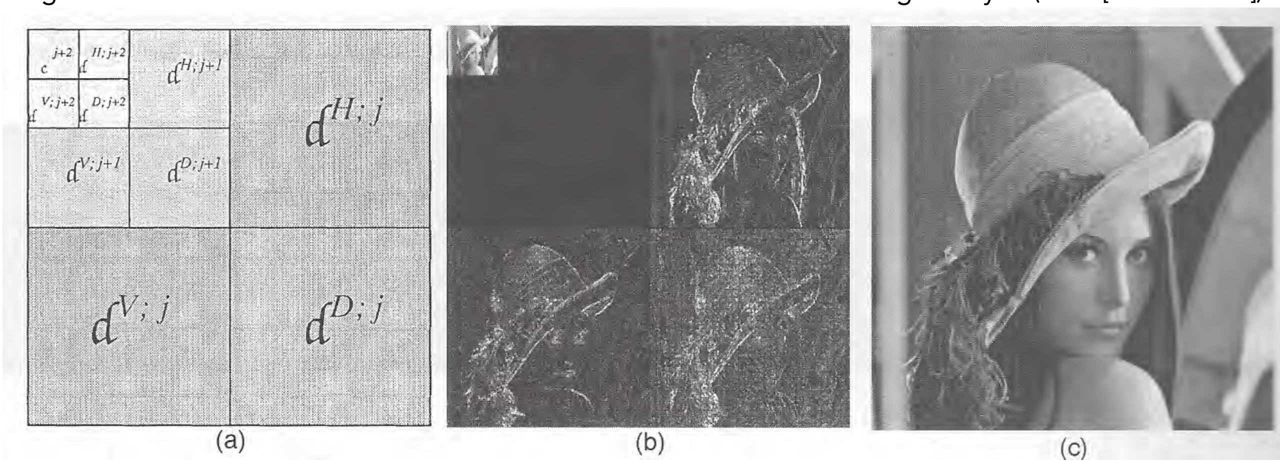


Figure 4: Frame-of-Discernment illustration based on multi-resolution image analysis (from [Prasad1997]).



Multi-resolution analysis offers another example [Mallat1989]. Multi-resolution decompositions combine a monotonic decomposition of the frame of discernment with a partitioning. The monotonic component is referred to as the *approximation* $\mathcal{A}$. Each approximation is further partitioned into another constituent approximation and a *detail* $\mathcal{D}$. Specifically, at the n^{th} scale,

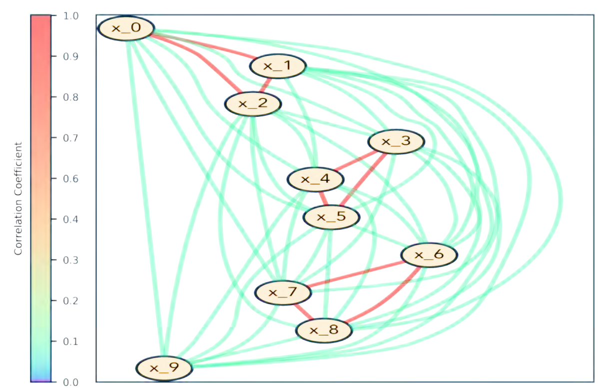
$$\mathcal{A}_{n+1} = \mathcal{A}_n \cup \mathcal{D}_n$$
$$\text{such that} \quad \mathcal{A}_n \cap \mathcal{D}_n = \emptyset.$$

Figure 4 contains a two-dimensional illustration from image-processing, a typical application [Prasad1997]. Figure 4(c) contains the original image. Figure 4(b) shows its multi-resolution decomposition. Figure 4(a) shows the partition scheme for the frame of discernment. Researchers [Tajeddini2019] have in fact combined multi-resolution analysis with ToE for noise-cancellation purposes.

How might evidence supporting partial belief might present itself in a practical context? Figure 5 illustrates using a graphical representation of a correlation matrix. Pair-wise correlation values are indicated by the heat-map color scale applied to the graph edges. The graph contains three cliques of three tuples of variables with correlations of 0.9.

Such instances can be interpreted as cases in which the measurements do not clearly resolve differences between the underlying measurands. Perhaps the measurands themselves are indistinct. Maybe the measure is not robustly order-preserving. In any case, distinguishing between discrete measured features might not be altogether justified.

Figure 5: Graphical representation of a correlation matrix for data with considerable collinearity.



In the case of Figure 5, assigning our beliefs to the highly-collinear cliques — $\{x_0, x_1, x_2\}$, $\{x_3, x_4, x_5\}$, $\{x_6, x_7, x_8\}$, $\{x_9\}$ — might be more-justified than attributing belief to the individual "atomic" propositions (features, variables). Principal-component analysis accomplishes this in a sense. Moreover, the unexplained variance (residuals) in a linear model might amount to belief vacuously attributed to the total feature set $\mathscr{X} = \{x_1, \ldots, x_9\}$; or to "none of the above", where $\mathscr{X} \ni x = \emptyset$.

2.3.2 The Mechanics of the Theory of Evidence

Extending Dempster's original method [Dempster1967] for the combination of evidence, Shafer [Shafer1976] described a set of four systematically-interrelated quantities for characterizing state of belief about propositions within the FoD. Both also described the combination of beliefs and evidence from distinct frames of discernment.

Four particular quantities reside at the theory's heart. We usually elicit a basic-belief assignment (bba) $m_{\underline{\Omega}}$ on our frame of discernment $\underline{\Omega} = \wp(\Omega)$. We populate it after the manner illustrated in Figure 3. Our bba satisfies the property

$$\sum_{\omega \in \underline{\Omega}} m_{\underline{\Omega}}(\omega) = 1. \tag{3a}$$

This is the ToE equivalent to the conventional Bayesian axiom

$$\sum_{\omega \in \Omega} \mathcal{P}(\{\omega\}) = 1,$$

where $\mathcal{P}$ is a probability-mass function on the "atomic" (singleton) focal elements of our FoD.

From $m_{\underline{\Omega}}$ we calculate three other quantities. The *Belief Function* $Bel_{\underline{\Omega}}$ is given by

$$Bel_{\underline{\Omega}}(\omega) = \sum_{\substack{\varpi \subset \omega \\ \varpi \neq \emptyset}} m_{\underline{\Omega}}(\varpi). \tag{3b}$$

We interpret $Bel_{\underline{\Omega}}$ as the *lower-probability* function. For a given proposition $\omega \in \underline{\Omega} = \wp(\Omega)$, $Bel_{\underline{\Omega}}(\omega)$ aggregates all of the belief attributed to ω itself, as well as to any subsets of ω. Moreover,

$$Bel_{\underline{\Omega}}(\Omega) = 1,$$

which indicates with certainty that any admissible proposition resides in the frame of discernment. [Smets1988, Smets1994] describe a ToE generalization by which this constraint is relaxed.

This justifies the direction of the edges of the directed graph in Figure 3. A basic-belief assignment attributed to some $\varpi \subset \omega$ does influence our total belief about ω. The converse is not however true. A belief attributed to $\omega \supset \varpi$ does not justify any conclusions about ϖ.

Our third quantity is referred to as a *Plausibility Function* $Pl_{\underline{\Omega}}$. Defined as

$$Pl_{\underline{\Omega}}(\omega) = \sum_{\omega \cap \varpi \neq \emptyset} m_{\underline{\Omega}}(\varpi), \tag{3c}$$

the plausibility function is interpreted as the *upper-probability* function. Plausibility accounts for the remaining belief supporting proposition ω after all of that refuting ω has been taken off the table.

Consider all of the evidence that refutes proposition ω. This can be expressed as $Bel_{\underline{\Omega}}(\neg\omega)$. This evidence is uncertain, and can only partially support refuting ω. In fact, we have as an identity

$$Pl_{\underline{\Omega}}(\omega) = Bel_{\underline{\Omega}}(\Omega) - Bel_{\underline{\Omega}}(\neg\omega) = 1 - Bel_{\underline{\Omega}}(\neg\omega).$$

Our final quantity, the *Commonality Function* $Q^{\underline{\Omega}}$ is defined by

$$Q_{\underline{\Omega}}(\omega) = \sum_{\substack{\varpi \supset \omega \\ \varpi \neq \emptyset}} m_{\underline{\Omega}}(\varpi). \tag{3d}$$

We interpret the commonality function as the belief attributable to some focal element ϖ that can be reallocated to other focal elements within the FoD. It has special application in the combination of beliefs on a compatible FoD according to Dempster's rule of combination. It is sometimes referred to as the "Fast-Fourier Transform of Belief Functions."

Finally, we can recover each of the above-four quantities — $m_{\underline{\Omega}}$, $Bel_{\underline{\Omega}}$, $Pl_{\underline{\Omega}}$, $Q_{\underline{\Omega}}$ — from any of the others. This is accomplished by a set of operations referred to as *Möbius Transforms*. We make use of one in particular,

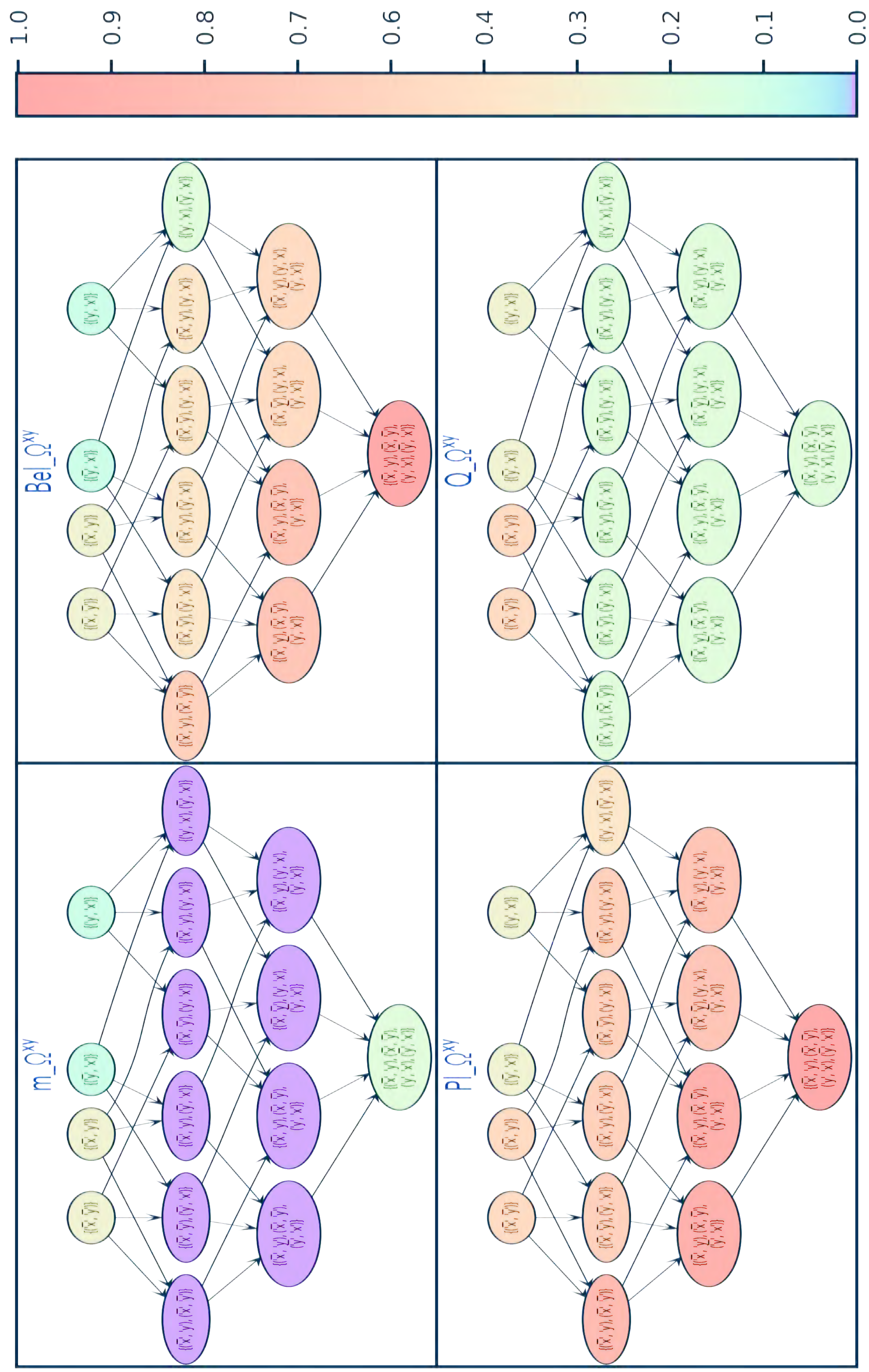
$$m_{\underline{\Omega}}(\omega) = \sum_{\varpi \subseteq \omega} (-1)^{|\omega|-|\varpi|} Bel_{\underline{\Omega}}(\varpi) \tag{3e}$$

in subsequent analysis. Here, $|\omega|$ denotes the cardinality — the number of set members — for focal element ω.

The fact that belief-attribution is preserved leads to a profound insight by which practical application of ToE to DA becomes possible. All of these operations merely *transfer* belief between focal elements. Belief is not destroyed or discarded. Figure 6 illustrates this for the four quantities defined above.

Having introduced the essential quantities by which ToE characterizes belief, we now summarize require operations essential to adapting ToE to the previously-described DA methodology. We specifically require marginalization, vacuous extension, and combining a conditional bba with a prior to get a joint bba. Vacuous extension is particularly important for comparing information of different evidentiary configurations. When

Figure 6: Illustrative Theory-of-Evidence (ToE) frame of discernment (after [Jiroušek2018]).



employing the information-theory concept of entropy, such comparisons must be based on a common FoD.

Marginalization is described in [Shafer1976] and [Almond1995]. Without loss of generality, we describe the operation here using the FoD illustrated in Figures 3 and 6. Our focal elements ω span two dimensions, (x, y). Moreover, $\Omega = \mathscr{X} \times \mathscr{Y}$. We represent ω as tuples, accordingly. For convenience, we define a projection operator, after [Almond1995], as

$$(\omega)^{\downarrow\mathscr{X}} = \{x : (x, y) = \omega \subseteq \underline{\Omega}\}.$$

In $(\omega)^{\downarrow\mathscr{X}}$, we basically retain the x's for which there is a corresponding $(x, y) \in \omega$.

We use this in marginalization. Our marginalized bba becomes

$$m_{\underline{\Omega}^{\downarrow\mathscr{X}}}(x) = \sum_{\substack{(\omega)^{\downarrow\mathscr{X}} = x \\ \omega \in \underline{\Omega}}} m_{\underline{\Omega}}(\omega). \tag{4}$$

Figure 7 illustrates the consequences of going from $m_{\underline{\Omega}}$ to $m_{\underline{\Omega}^{\downarrow\mathscr{X}}}$. The frame labeled "$m\_\Omega^{xy}$" reiterates our initial bba $m_{\underline{\Omega}}$ on FoD $\underline{\Omega}$ from Figure 3. The frame labeled "m_Ω^{x}" corresponds to $m_{\underline{\Omega}^{\downarrow\mathscr{X}}}$ on FoD $\underline{\Omega}^{\downarrow\mathscr{X}}$.

Consider specifically the most-vacuous focal element, $\Omega = \left\{ \begin{matrix} (x, y), (\bar{x}, y), \\ (x, \bar{y}), (\bar{x}, \bar{y}) \end{matrix} \right\}$. The corresponding focal element in the marginalized FoD is $\Omega^{\downarrow\mathscr{X}} = \{ (x, \bar{x}) \}$. Similarly, $\{ (x, y), (\bar{x}, y) \}^{\downarrow\mathscr{X}} = \{x, \bar{x}\}$. Also observe how the bba is reallocated from $\underline{\Omega} \mapsto \underline{\Omega}^{\downarrow\mathscr{X}}$. We have

$$\begin{aligned} m_{\underline{\Omega}^{\downarrow\mathscr{X}}}(x) &= m_{\underline{\Omega}}\Big(\{ (x, y) \}\Big) + m_{\underline{\Omega}}\Big(\{ (x, \bar{y}) \}\Big) \\ &= 0.1 + 0.1 \\ &= 0.2, \end{aligned}$$

and

$$\begin{aligned} m_{\underline{\Omega}^{\downarrow\mathscr{X}}}(\bar{x}) &= m_{\underline{\Omega}}\Big(\{ (\bar{x}, y) \}\Big) + m_{\underline{\Omega}}\Big(\{ (\bar{x}, \bar{y}) \}\Big) \\ &= 0.3 + 0.3 \\ &= 0.6. \end{aligned}$$

This is belief-transferrance provides the basis for the Smets' *Transferable-Belief Model* [Smets1994, Smets2002], on which we base our ToE-based DA methodology.

We now seek to reverse the operation, through vacuous extension. We seek to map from $\underline{\Omega}^{\downarrow\mathscr{X}} \mapsto \underline{\Omega}^{\downarrow\mathscr{X}\uparrow\mathscr{Y}}$. During our preceding mapping from $\underline{\Omega} \mapsto \underline{\Omega}^{\downarrow\mathscr{X}}$ discarded any information in the $\mathscr{Y}$-dimension. Any extension back into that dimension must therefore be vacuous. Our belief-attribution cannot distinguish between y or $\bar{y}$. Beliefs attributed in $\underline{\Omega}^{\downarrow\mathscr{X}}$ to $\{x\}$ or $\{\bar{x}\}$ map in $\underline{\Omega}^{\downarrow\mathscr{X}\uparrow\mathscr{Y}}$ to $\{ (x, y), (x, \bar{y}) \}$ and $\{ (\bar{x}, y), (\bar{x}, \bar{y}) \}$, respectively. The frame labeled "$m\_\Omega^{xy\downarrow x\uparrow xy}$" in Figure 7 depicts this.

The belief-transferrance associated with extension appears notationally as

$$m_{\underline{\Omega}^{\downarrow\mathscr{X}\uparrow\mathscr{Y}}}(\omega) = \begin{cases} m_{\underline{\Omega}^{\downarrow\mathscr{X}}}(x) & \begin{array}{c} \text{if } \exists x \subseteq \mathscr{X} : \\ \omega = \{x\} \times \{y, \bar{y}\} \end{array} \\ 0 & \text{otherwise } \wedge \omega \subseteq \underline{\Omega} \end{cases}.$$

The information in $m_{\underline{\Omega}^{\downarrow\mathscr{X}\uparrow\mathscr{Y}}}$ exactly corresponds to that in $m_{\underline{\Omega}^{\downarrow\mathscr{X}}}$. It has merely been redistributed — transferred — to a *minimally-extended* FoD whose span now includes $\mathscr{Y}$. However, we can no longer resolve between focal elements in the $\mathscr{Y}$-dimension. Such extensions are alternatively referred to as *cylindrical.*

2.3.3 Decision Analysis Using the Theory of Evidence

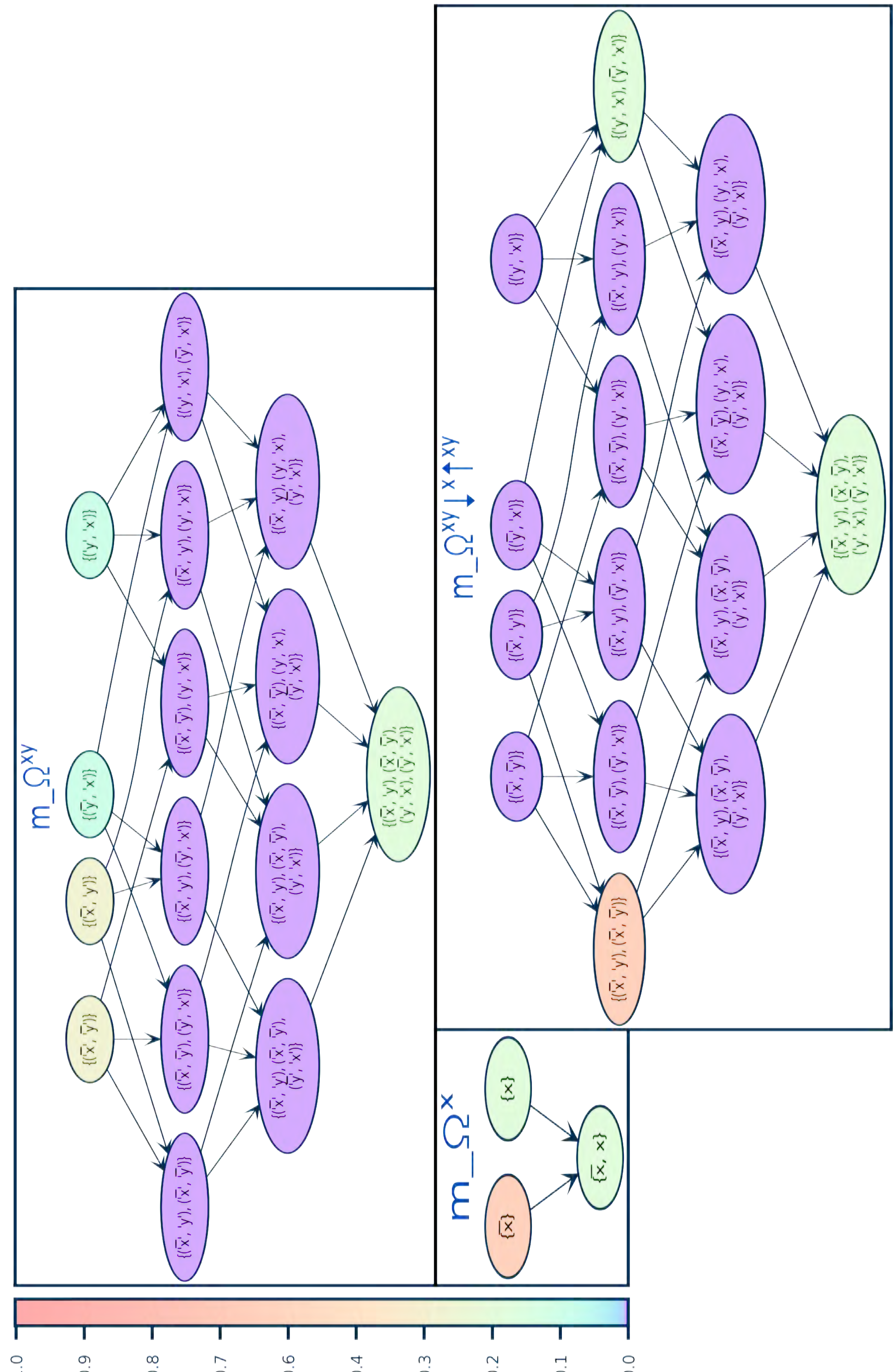
Although basic-belief assignment over a frame of discernment provide for a rich expression of uncertainty, the decision analysis employing the Howard-Raiffa methodology requires expression in terms of traditional Bayesian belief functions. Specifically, we must assign unit-norm sets of conditional and marginal probabilities to edges emanating from chance nodes in a decision tree. Figure 2 illustrates the pattern. Two approaches have been offered. We review these here.

Smets developed a Transferable-Belief Model (TBM) formulation of ToE for decision analysis [Smets1994, Smets2002]. The TBM formulation defines a *Pignistic Transformation* involving a *Betting Frame*. The betting fame is defined as

$$Bet_{P\Omega}(\omega) = \sum_{\substack{\omega \in \varpi \\ \varpi \subseteq \underline{\Omega}}} \frac{m_{\underline{\Omega}}(\varpi)}{|\varpi|}. \tag{5}$$

Smets developed the Pignistic Transformation from a set of six distinct desiderata. It is grounded in subjec-

Figure 7: Illustration of marginalization along projection $(\omega)^{\downarrow\mathscr{X}}$ followed vacuous projection $(\omega)^{\downarrow\mathscr{X}\uparrow\mathscr{Y}}$ (after [Jiroušek2018]).



tive probabilistic concepts, and emphasizes a *credal-pignistic* link.

Subsequently, Cobb and Shenoy [Cobb2006, Jiroušek2018] proposed an alternative approach. Their approach — the *Plausibility Transformation* — makes use of the Plausibility Function $Pl_{\underline{\Omega}}$. The argue that the plausibility transformation is "more-consistent with the Dempster-Shafer semantics of belief-function models." It moreover produces results under some circumstances that more-closely resemble expectations arising in scenarios involving Dempster's rule of belief-function combination [Dempster1967, Shafer1976]. The plausibility transformation ids defined as

$$Pl_{P^{\Omega}}(\{\omega\}) = \frac{Pl_{\underline{\Omega}}(\{\omega\})}{\sum_{\omega\in\underline{\Omega}} Pl_{\underline{\Omega}}(\{\omega\})}. \tag{6}$$

This is effectively the normalization of the Plausibility Function — or upper probability — over the singleton (atomic) focal elements in the frame of discernment.

The arguments about relative advantages of each technique appear arcane. Sufficient experience has not yet been accrued from which to characterize an empirically-derived distinction. The sequel proceeds using Smets' Pignistic transform (5).

2.4 Information Quantification

Attributing value to information features prominently among the objectives of this study. We therefore seek to quantify it in the most-fundamental fashion possible. For that purpose we turn to Information Theory.

Modern Information Theory was developed during the nascent stages of digital communications [Shannon1948, Gallagher1968]. Claude Shannon, its most-prominent pioneer, sought a theoretical upper bound on the volume of information that could be pushed through a noisy communications channel. He discovered that this can be characterized using quantitative mechanisms closely resembling those of the statistical-theromynamics discipline from physics [Ben-Naim2008]. Accordingly, Shannon termed the fundamental quantity *entropy*.

Within information-quantification contexts, entropy assumes at least two interpretations. First, it bears a meaning congruent with John Maynard Keynes *irreducible uncertainty* [Keynes1921]. This is the mirror image of Shannon's *channel capacity*. Cover and Thomas [Cover1991] derived a corollary theorem, a *data-processing inequality*. This theorem demonstrates that no amount of clever processing can extract information beyond a certain limit characteristic of a given data set.

The data-processing inequality is analogous to the Cramer-Rao Lower Bound (CRLB) [Kay1993, Hogg2013] on estimation-error variance from mathematical statistics. Within a measurement scenario, entropy can conversely be thought-of as the quantity of noise-free measurements — unaccompanied by any uncertainty — required to achieve perfect certainty about the state of a measurand.

We summarize the information-theoretic concepts required to justify our central conjecture about modes of information-derived differentiation. For instance, our generalized-decision problem (1a)-(1c) characterized the addition of differentiating information in terms of the opportunity make decision on conditioned probabilities instead of marginal ones. We develop the information-theoretic machinery to demonstrate this proposition.

2.4.1 Information-Quantification for Bayesian-Belief Functions

Consider a traditional probability distribution $\mathcal{P} : \mathbb{R}^n \mapsto [0, 1]$. We define its entropy as

$$\mathcal{H}(\mathcal{P}) = -\sum_{x\in\mathbb{R}^n} \mathcal{P}(x) \log(\mathcal{P}(x)). \tag{7}$$

The base of the logarithm is arbitrary. While studying *binary*-symmetric communications channels, Shannon found a base-two logarithm convenient. The differential equations of statistical thermodynamics lead to a natural logarithm.

Decision Analysis — as framed according to the Howard-Raiffa methodology — concerns itself with discrete, categorical probabilities. We are specifically interested in the marginal and conditional probabilities assigned to the graph edges emanating from chance nodes in a decision tree exemplified by Figure 2. This allows interesting considerations regarding the most-convenient logarithm base.

Recall that a state of complete uncertainty for an N-category discrete probability distribution is assigned a value of

$$\mathcal{P}(\omega_n) = \frac{1}{N} \quad \forall\, \omega_n \in \{\omega_1, \ldots, \omega_N\}$$

in a Bayesian context. The resulting entropy appears then as

$$\mathcal{H}(\mathcal{P}) = \log(N).$$

Observe then that if we select a logarithmic base of N, then we get $\mathcal{H}(\mathcal{P}) = 1$. This motivates the interpretation of entropy as the quantity of uncertainty-free propositions required to achieve compete certainty. If we receive $\mathcal{P}(\omega_\nu) = 1$ for some $\omega_\nu \in \{\omega_1, \ldots, \omega_N\}$, then $\mathcal{H}(\mathcal{P}) = 0$ results. Zero entropy corresponds to perfect certainty.

We now consider the entropy of joint and marginal distributions. Without loss of generality, consider a two-dimensional categorical distribution on $\underline{\Omega} = \mathcal{X} \times \mathcal{Y}$ such that $\mathcal{P} : \underline{\Omega} \mapsto [0,1]$. Our corresponding entropy appears as

$$\mathcal{H}(\mathcal{P}_{\underline{\Omega}}) = -\sum_{x\in\mathscr{X}}\sum_{y\in\mathscr{Y}} \mathcal{P}_{\underline{\Omega}}(x,y) \log\Big(\mathcal{P}_{\underline{\Omega}}(x,y)\Big).$$

Now consider another scenario $\mathcal{P}_I(x,y) = \mathcal{P}_{\mathscr{X}}(x)\,\mathcal{P}_{\mathscr{Y}}(y)$ in which x and y are statistically independent. This entropy simplifies to

$$\begin{aligned}\mathcal{H}(\mathcal{P}_I) &= -\sum_{x\in\mathscr{X}}\sum_{y\in\mathscr{Y}} \mathcal{P}_I(x,y) \log\Big(\mathcal{P}(x,y)\Big) \\ &= -\sum_{x\in\mathscr{X}} \mathcal{P}_{\mathscr{X}}(x) \log\Big(\mathcal{P}_{\mathscr{X}}(x)\Big) \\ &\quad -\sum_{y\in\mathscr{Y}} \mathcal{P}_{\mathscr{Y}}(y) \log\Big(\mathcal{P}_{\mathscr{Y}}(y)\Big)\end{aligned}.$$

From Jensen's inequality we can show that in general

$$\mathcal{H}(\mathcal{P}_\Omega) \leq \mathcal{H}(\mathcal{P}_{\mathscr{X}}) + \mathcal{H}(\mathcal{P}_{\mathscr{Y}}). \tag{8}$$

So, in the special case in which $\mathcal{P}_{\mathscr{X}}$ and $\mathcal{P}_{\mathscr{Y}}$ are marginals of $\mathcal{P}$ on $\mathscr{X} \times \mathscr{Y}$, respectively, then the marginalization operation often discards information. This observation is consistent with our conjecture that differentiated access to information and differentiated insights reduce the uncertainty underlying the decisional context. Decision-making without the benefit of differentiated access or differentiated insights amounts to using marginal distributions.

Finally, how do we compare information associated with distinct probability distributions? Direct comparison of entropy numbers provides a rough approach. Only comparisons of entropies evaluated on the same probability-distribution domain are valid, however. A comparison of the type in (8) is only valid if in fact $\underline{\Omega} = \mathscr{X} \times \mathscr{Y}$. It would not be valid to compare $\mathcal{H}(\mathcal{P}_\Omega)$ with $\mathcal{H}(\mathcal{P}_{\mathscr{X}})$ by itself. The two quantities are incompatible.

The *Kullback-Leibler* (KL) *distance* — also referred to as relative entropy — provides the most-direct comparison. The KL distance is defined as

$$\mathcal{D}(\mathcal{P} \parallel \mathcal{Q}) = \sum_{\omega\in\underline{\Omega}} \mathcal{P}(\omega) \log\left(\frac{\mathcal{P}(\omega)}{\mathcal{Q}(\omega)}\right). \tag{9}$$

This explicitly enforces the constraint that the distributions being compared have the same support. Most-conspicuously, the quantity becomes undefined for any ω such that $\mathcal{Q}(\omega)$ is zero but $\mathcal{P}(\omega)$ is finite.

2.4.2 Information-Quantification on Basic-Belief Assignments

Information-quantification on a ToE FoD is obviously more-complicated. Quite-recently, Jiroušek and Shenoy [Jiroušek2018] offered a definition for such an entropy. Their ToE entropy is comprised of two components. The Shannon Entropy defined in (7) is the first component. We evaluate the Shannon entropy in the singleton focal elements in our frame of reference. In the ToE context, the Shannon entropy is evaluated on the plausibility transformation $Pl_{P\Omega}$ defined in (6), and appears as

$$\begin{aligned}\mathcal{H}_S\Big(Pl_{P\Omega}\Big) = -\sum_{\{\omega\}\subset\mathscr{X}\times\mathscr{Y}} \Bigg[&Pl_{P\Omega}(\{\omega\}) \\ &\times \log\Big(Pl_{P\Omega}(\{\omega\})\Big)\Bigg].\end{aligned}$$

Jiroušek and Shenoy selected for their second component an entropy described by Dubois and Prade. This appears as

$$\mathcal{H}_d\Big(m_{\underline{\Omega}}\Big) = \sum_{\omega\in\underline{\Omega}} m_{\underline{\Omega}}(\omega) \log(|\omega|).$$

The total entropy associated with $m_{\underline{\Omega}}$ becomes the sum of these two quantities. Specifically,

$$\mathcal{H}\Big(m_{\underline{\Omega}}\Big) = \mathcal{H}_S\Big(Pl_{P\Omega}\Big) + \mathcal{H}_d\Big(m_{\underline{\Omega}}\Big).$$

Two things are noteworthy about $\mathcal{H}_d$. First, it returns values of zero for singleton focal elements, whose cardinality is unity. It therefore complements the Shannon entropy. A given focal element will either contribute a Shannon-entropy component, or a Dubois-Prade one.

The acceptance of the log-cardinality term is also interesting. Its resemblance to Smets' Pignistic transformation (5) seems noteworthy. It is consistent with the *Principle of Minimal Commitment* to which he devotes considerable attention in a description [Smets1993] of a key piece of machinery underlying the TBM methodology.

We also observe that a ToE equivalent to the KB distance is not obviously practical. Consider, for example, the illustration in Figure 7. The bba $m_{\Omega^{\downarrow \mathscr{X} \uparrow \mathscr{Y}}}$ might be considered approximately equal to the case considered in (8). This is particularly valid if $\mathcal{P}_{\mathscr{X}}(x) = \sum_{y \in \mathscr{Y}} \mathcal{P}_{\Omega}(x,y)$ and $\mathcal{P}_{\mathscr{Y}}(y) = \sum_{x \in \mathscr{X}} \mathcal{P}_{\Omega}(x,y)$. If we took $\mathcal{P}_{I}(x,y) = \mathcal{P}_{\mathscr{X}}(x)\, \mathcal{P}_{\mathscr{Y}}(y)$, the outer product of the marginals, this approximately resembles marginalization followed by vacuous extension.

Under such circumstances, the Dubois-Prade entropy would take values of zero on focal elements for which belief had been transferred during the operations. These zero-valued contributions to the overall entropy have the affect of overstating the degree of certainty. Caution is therefore advised when considering such a comparison.

2.5 Competency-Focused Strategy Frameworks

Our basic conjecture that information — and by extension decision-making— affords opportunities for competitive advantage insinuates us into the realm of strategy. The strategy science community tends to define its realm narrowly. The discipline traditionally focused on market positioning and resources [Martin2015].

Researchers more-recently proffered updated definitions about what constitutes a strategic decision. Leiblein, *et al*, [Leiblein2018] argue for complexity criteria including interdecisional, interactor, and intertemporal characteristics. Van den Steen [VanddenSteen2016] suggests along the same lines criteria including centrality, importance, capability-development, and competitiveness. Fully aligning decision-making to those factors awaits future consideration.

We do briefly consider here competency-oriented frameworks for competitiveness. Figure 8 illustrates two that are particularly apropos. These include David Teece's *Dynamic-Capabilities* framework [Teece2009] and the *Adaptivity-Advantage* methodology by Martin Reeves [Reeves2012]. Teece characterizes Dynamic Capabilities in terms of three-stage lifecycle. We use this lifecycle as the unifying structure for evaluating the relevance of each of the information-differentiation modes from our central conjecture.

We overlay Reeves' methodology — complementary to that of Teece — over the lifecycle. This alignment of frameworks, combined with their complementary perspectives, allow us to consider how the conjectured information-differentiation modes contribute in each phase.

As conveyed by the fade-in/fade-out representation in Figure 8, the lines of demarcation between the competitive-position-lifecycle stages are fuzzy, as opposed to crisp. Cycle-stage durations and progressions may vary by industry, also. A framework by Suarez, *et al*, [Suarez2005] attempts to characterize industry contexts by degrees and natures of ecosystem volatilities. In the most-dynamic, Teece's three stages may actually be continuously recurring simultaneously. This complexities remain for subsequent consideration.

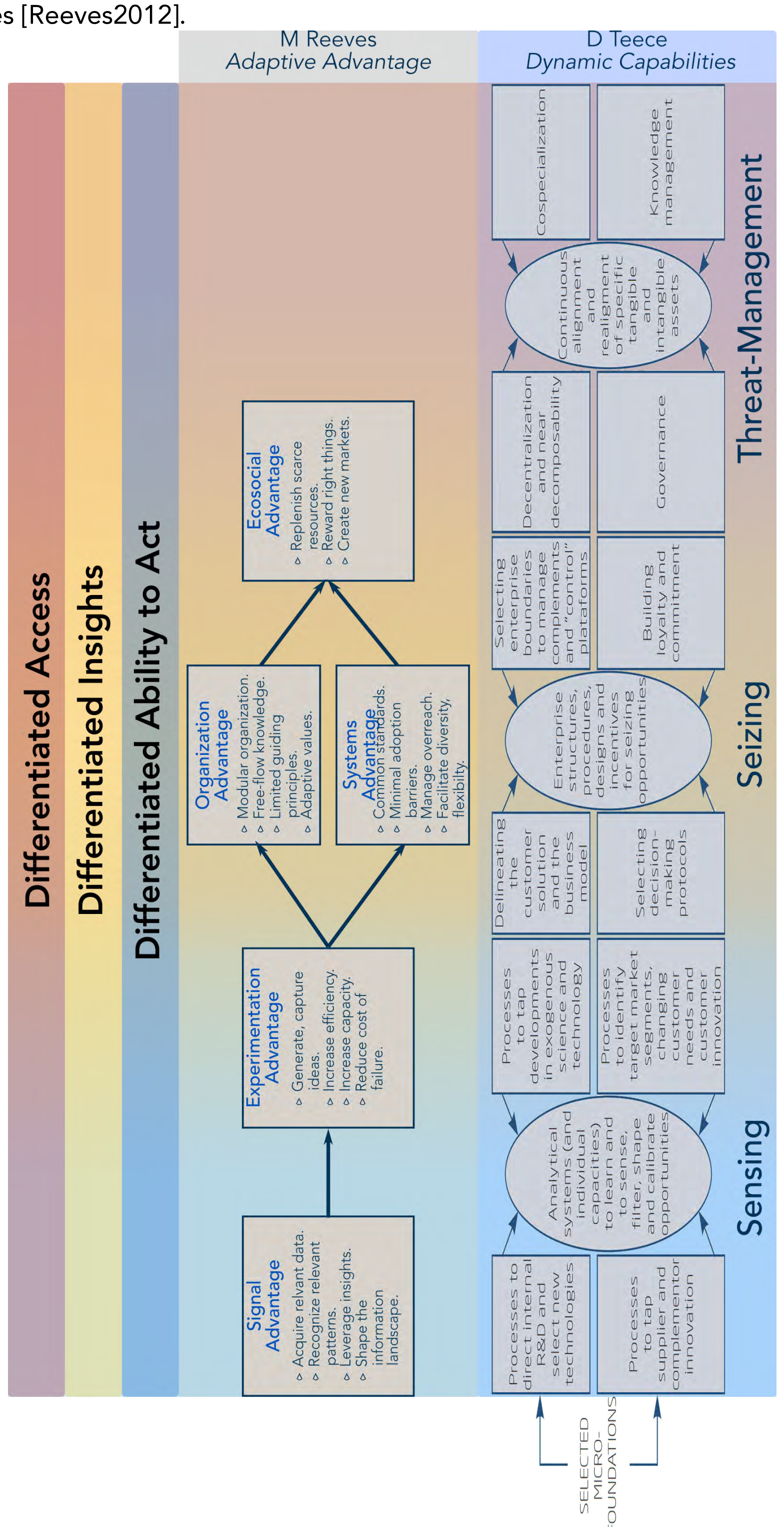
An attempt to assign distinct conjectural modes of information to individual competitive-lifecycle phases appears in Figure 8. We therefore cursorily describe how each mode contributes to each phase. [Hamlett2019] covers this topic in greater detail.

2.5.1 Sensing Stage

Teece's *Sensing Stage* fundamentally focuses on identifying opportunities to assert new market positions and identifying new resource needs. New market positions could pertain to material or etherial business objectives. We align Reeves' *Signal Advantage* to the sensing stage. The framework describes cultivation of organizational-epistemic competencies for sensing, rather than the act of sensing itself.

Activities in the sensing stage of the competitive lifecycle place a premium on differentiated insights and differentiated access, as well. The insights mode seems particularly important in creating new markets. Would-be disrupters ([Christensen2004, Christensen2015]) attempting to encroach upon incumbent market space see and act upon the latters'

Figure 8: Overlay of information-differentiation modes onto competency-oriented strategy frameworks of Teece [Teece2009] and Reeves [Reeves2012].



vulnerabilities. [Bryce2007] contains a playbook for disrupters and other insurgents.

Market-space incumbents usually enjoy differentiated access based on their established positions. Powerful, new information-centered operating models have emerged since Porter [Porter1985] and Drucker [Drucker1993] separately described the progressive centrality of information's importance to the economy. Platform operators ([VanAlstyne2016, Iansiti2017]) harvest deep insights about market-transaction counterparties through the formers' roles as orchestrators of transactions. They enjoy the benefit of formidable information asymmetries. This translates into a greater likelihood of recognizing new marketplace opportunities than would-be insurgents experience.

Connected-systems business models — sometimes referred to as *Internet of Things* — also create pronounced information asymmetries for incumbents ([Porter2014, Porter2015]). These models — with their internal and expectation realizations from an organizational perspective — involve extensive instrumentation of products, pipelines, and systems. This generates voluminous data for mining. The associated information is typically proprietary, affording access advantages to owners of the systems that generate it.

2.5.2 Seizing Stage

Teece [Teece2009] defines the *Seizing Stage* from an incumbent's perspective. It involves agilely reconfiguring existing resources in order to exploit a new opportunity. [Reeves2012] summarizes distinguishing characteristics allowing organizations to successfully pull this off.

Informationally, ability to act involves a strong organizational-learning component in addition to adaptivity and dynamism. Perhaps all of these attributes exist in concert. Considerable research describes inhibitors to such learning. Drennel and March [Denrell2001] studied risk-aversion — a "hot-stove" affect — as an inhibitor to learning. Pisano [Pisano2019] makes passing mention of indiscriminate "learning" of the wrong things by "innovative" cultures. Groysberg, *et al*, [Groysberg2018] constructed a framework characterizing organizational cultures. Different cultural styles, obviously, better lend themselves to seizing than others.

2.5.3 Threat-Management Stage

Marketplace positions, once captured, must be defended. Threat-management involves aspects of both the sensing and the seizing stages. Sensing provides insights into vulnerabilities and new opportunities. It involves overcoming biases such as confirmation biases and endowment affects [Kahneman2011] to formulate a clear view of the environment. Inhibitors such as those described by [Denrell2001] and [Pisano2019] become vulnerabilities in themselves. Vulnerabilities exploitable by would-be disruptors ([Christensen2004, Christensen2015]) and other insurgents ([Bryce2007]) must be identified and remediated.

Ability-to-act differentiation takes on two aspects. The first is competent management of the existing market position ([Sadun2017]). Operational effectiveness — particularly in complex contexts — remains difficult to replicate.

Reconfiguring existing resources in response to a threat is however often difficult for incumbents to manage. This can involve painful decisions to cannibalize existing revenue streams [Wessel2012]. From a behavioral-economics perspective, this requires overcoming endowment affects.

3 Case Study

We now employ a simple decision-analysis case study to demonstrate concepts related to our central conjecture. First analyzed by Strat [Strat1994], the case involves analyzing a decision about a petroleum-extraction project. Strat established a baseline using a traditional Bayesian DA approach. He then applied a ToE-based approach. Smets [Smets2002] subsequently revisited the problem using his TBM methodology.

Here we reproduce and extend the previous work. Our extensions look at the case from different angles in order to corroborate our central conjecture about differentiation modes based on information. This includes looking deeper into the results to understand specifically how additional information shifts the differentiations between outcomes. We conduct a parameter study to demonstrate the relationship between the value of information and its underlying uncertainty. We show how reduced uncertainty leads to higher informational value.

Our extension of Smets' results from the Transferable-Belief Model (TBM) provide a basis for a point-comparison with the Bayesian approach. Specifically, we compare entropies according to Jiroušek's and Shenoy's approach [Jiroušek2018] according to a common frame of discernment. We thereby demonstrate that the ToE representation leads to a lower degree of certainty. We again seek to illuminate how that difference shows up in the distributions of final outcomes.

The parameters of the case study appear to be stylized and contrived. They nonetheless provide the mechanism to develop logical machinery that can subsequently be applied to more-realistic problems. They moreover support thought experiments by which we validate the premises of our central conjecture.

3.1 Decision Analysis Based on the Traditional Bayesian Method

We are considering an oil-extraction project. We must decided whether or not to drill a new oil well. We face two dimensions of uncertainty related to $\mathscr{S}$, the state of the oil field and $\mathscr{H}$ the productivity of an oil will drilled there. Three yield states $\mathscr{H}$ — Dry well with no production, a Trickle well producing a modicum of output, and a Gusher producing high volumes — are possible. The net-present value (NPV) of revenues from these states are \$0, \$120,000, and \$270,000, respectively. A well costs \$70,000 to drill.

We may at a cost of \$10,000 elect to perform a seismic survey. Such a "test" reveals information about the geological structure $\mathscr{S}$ of the field. The likely yield depends in part on this structure. A field lacking structure — "None" — is likely to be porous. Petroleum crude is unlikely to have accumulated under such circumstances. Accumulation may have occurred if the structure is is "Open". A "Closed" structure is conducive to a petroleum reservoir.

So, we face two decisions: Whether to test $\mathfrak{T} = \{\text{Yes}, \text{No}\}$, and whether to drill $\mathfrak{D} = \{\text{Yes}, \text{No}\}$. We may make our decision $\mathfrak{D}$ with or without knowledge of $\mathscr{S}$, the oil-field structure. The adaptation of our generalized optimum decision rule (1a) becomes therefore

$$(\mathfrak{D}^{\text{opt}}, \mathfrak{T}^{opt}) = \\ = \text{arg}\max_{(\mathfrak{D},\mathfrak{T})} \Bigg\{ \Big(\sum_{\substack{h \in \mathscr{H} \\ s \in \mathscr{S}}} \mathscr{P}(h, s;\, \mathfrak{T})\, \mathscr{R}(h;\, \mathfrak{D}) \Big) \\ - \mathscr{C}(\mathfrak{D}, \mathfrak{T}) \Bigg\}.$$

Observe that our return is independent of the state $\mathscr{S}$ of the oil field. Moreover, it only depends on decision $\mathfrak{D}$ whether or not to drill. If we test $\mathfrak{T} \ni \mathfrak{t} = \text{Yes}$, then our probability distribution $\mathscr{P}(h, s;\, \mathfrak{D}, \mathfrak{t} = \text{Yes})$ remains a joint distribution in $\mathscr{H} \times \mathscr{S}$. We get to make our decision using probabilities $\mathscr{P}(h;\, \mathfrak{D}, \mathfrak{t} = \text{Yes} \mid s)$ conditioned on $\mathscr{S}$.

If we do not test — $\mathfrak{T} \ni \mathfrak{t} = \text{No}$ — then $\mathscr{P}(h, s;\, \mathfrak{D}, \mathfrak{t} = \text{No}) = \mathscr{P}(h;\, \mathfrak{D}, \mathfrak{t} = \text{No})$. That is, our probability distribution $\mathscr{P}$ becomes independent of the oil-field structure $\mathscr{S}$. Our outcome probabilities for $\mathscr{H}$ are based on the marginal $\mathscr{P}(h) = \sum_{s \in \mathscr{S}} \mathscr{P}(h, s)$. Our prior information-theory discussion leads us to expect the corresponding reduced degree of certainty will lead to lower expected net returns.

Also observe that the probability distribution $\mathscr{P}$ is independent of the decision whether or not to drill $\mathfrak{D}$. It characterizes probabilistic outcomes $\mathscr{H}$, which in turn influence the return $\mathscr{R}$. In our general case (1a) - (1b) we abstracted our set of decisions into a single, opaque set. In a specific scenario, we partition our decision set into subsets influencing our knowledge of the probability state, and those merely influencing outcomes. In some cases distinct decisions may fit into both categories. Our scenario here is clearer when we separate them out.

Our two returns — distinguished by whether or not we conduct the geophysical survey $\mathfrak{T}$ — become from (1b)

$$\mathscr{R}^{\text{test}} = \sum_{\substack{h \in \mathscr{H} \\ s \in \mathscr{S}}} \mathscr{P}(h, s;\, \mathfrak{t} = \text{Yes})\, \mathscr{R}(h;\, \mathfrak{D}^{\text{opt}}) \Big) \\ - \mathscr{C}(\mathfrak{D}^{\text{opt}}, \mathfrak{t} = \text{Yes})$$

and

$$\mathcal{R}^{\text{no test}} = \sum_{\substack{h \in \mathcal{H} \\ s \in \mathcal{S}}} \mathcal{P}(h; , \mathfrak{t} = \text{No})\, \mathcal{R}(h; \mathfrak{D}^{\text{opt}})\Big) - \mathcal{C}(\mathfrak{D}^{\text{opt}}, \mathfrak{t} = \text{No}).$$

After (1c), the value $\mathcal{V}_{\text{test result}}$ of the information from the seismic survey appears as

$$\mathcal{V}_{\text{test result}} = \mathcal{R}^{\text{test}} - \mathcal{R}^{\text{no test}} - \mathcal{C}(\mathfrak{D}^{\text{opt}}, \mathfrak{t} = \text{Yes}). \quad (10)$$

Recall that our key question is "How much is the information worth?" We therefore subtract the cost of its acquisition out of our calculation of $\mathcal{V}_{\text{test result}}$. Our return $\mathcal{R}^{\text{test}}$ given $\mathfrak{t} = \text{Yes}$ is calculated net $\mathcal{C}(\mathfrak{D}^{\text{opt}}, \mathfrak{t} = \text{Yes})$, which includes the decision-acquisition cost. To explicitly see the information's value, we segregate out its acquisition cost. Also, we obtain an expected marginal return on information as

$$\mathcal{M}_{\text{test result}} = \frac{\mathcal{R}^{\text{test}} - \mathcal{R}^{\text{no test}}}{\mathcal{C}(\mathfrak{D}^{\text{opt}}, \mathfrak{t} = \text{Yes}) - \mathcal{C}(\mathfrak{D}^{\text{opt}}, \mathfrak{t} = \text{No})}. \quad (11)$$

Finally, the presence of a "Test/No-Test" decision $\mathfrak{T}$ makes this resemble a differentiated-access scenario. Were this a competitive situation, one market participant might select $\mathfrak{T} \ni \mathfrak{t} = \text{Yes}$, while another operates in a scenario characterized by $\mathfrak{T} \ni \mathfrak{t} = \text{No}$. *If* information from the test produces a scenario in which conditioning $\mathcal{P}(h; \mathfrak{D}, \mathfrak{t} = \text{Yes} \mid s)$ is materially advantageous to with respect to the marginal $\mathcal{P}(h; \mathfrak{D}, \mathfrak{t} = \text{No})$, then participants electing to test enjoy an informational asymmetry.

Table 1 gives the probabilistic structure of the case study. This is a joint probability structure. The resulting conditional probabilities appear in Table 2. We see for this scenario, the conditional probabilities in Table 2 differentiate much more than than the joints. For another perspective, Table 3 contains the outer product of the marginals from Table 1. This begins to reveal the relative loss of information due to marginalization.

We can quantitatively characterize this in two ways. First, we have the Kullback-Leibler distance (9). The KL distance between distributions in Table 1 and Table 3, the outer product marginals of the former, is 0.144 nats. We use the natural logarithm, here, for entropy calculations.

Table 1: Joint probability distribution $\mathcal{P}(h, s)$ relating oil-field geophysical structure $\mathcal{S}$ to drilling outcomes. (From [Strat1994].)

	Structure, $\mathcal{S}$			
Outcome, $\mathcal{H}$	None	Open	Closed	Marginal
Dry (Dr)	0.30	0.15	0.05	0.50
Trickle (Tr)	0.09	0.12	0.09	0.30
Gusher (Gu)	0.02	0.08	0.10	0.20
Marginal	0.41	0.10	0.24	1.00

Table 2: Conditional probability distributions $\mathcal{P}(h \mid s)$relating oil-field geophysical structure to drilling outcomes. (After [Strat1994].)

	Structure, $\mathcal{S}$		
Outcome, $\mathcal{H}$	None	Open	Closed
Dry (Dr)	0.7317	0.4286	0.2083
Trickle (Tr)	0.2195	0.3429	0.3750
Gusher (Gu)	0.0488	0.2286	0.4167
$\sum_h \mathcal{P}(h \mid s)$	1.00	1.00	0.24

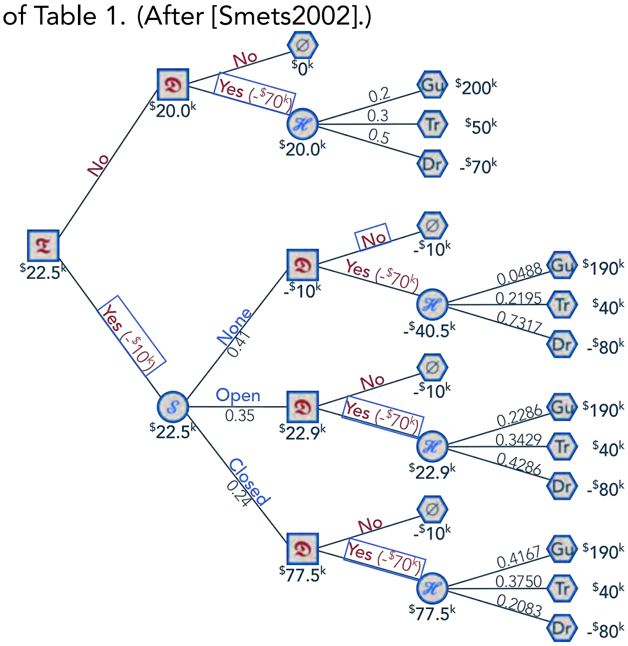
Alternatively, we might consider the sums of the diagonal elements of the two tables. This is loosely related to the categorical-statistical concepts of sensitivity and sensitivity. The sum of the diagonals in Table 1 is 0.52, while the sum is 0.358 for Table 3. Consequently Table 1 contains 1.45 times more of its population in the diagonals than Table 3. We employ these in our subsequent parameter study in an attempt to gain insight into information and value.

Figure 9, now, depicts the decision-tree framing according to the Howard-Raiffa methodology. This is Figure 2 with the probabilities applied to the graph vertices emanating from the chance nodes, and the expected net returns propagated back to the original decision node. Our first decision is $\mathfrak{T}$, whether conduct the seismic survey. We apply marginal utilities to graph edges emanating from the left-most, circu-

Table 3: Outer-product of marginals on $\mathcal{H}$, and $\mathcal{S}$ for $\mathcal{P}(h, s)$ given in Table 1.

	Structure, $\mathcal{S}$			
Outcome, $\mathcal{H}$	None	Open	Closed	Marginal
Dry (Dr)	0.205	0.175	0.12	0.50
Trickle (Tr)	0.123	0.105	0.12	0.30
Gusher (Gu)	0.082	0.07	0.048	0.20
Marginal	0.41	0.10	0.24	1.00

Figure 9: Decision-tree framing for the oil-drilling case study employing Strats' Bayesian uncertainty structure of Table 1. (After [Smets2002].)



lar chance nodes. For the $\mathfrak{T} \ni \mathfrak{t} =$ No scenario we operate without information about the structure $\mathcal{S}$ of the oil field. Consequently, we apply $\mathcal{P}(h)$ from the marginal column in Table 1 to that $\mathcal{H}$ chance node. For the $\mathfrak{T} \ni \mathfrak{t} =$ Yes scenario apply $\mathcal{P}(s)$ from Table 1's marginal row to the $\mathcal{S}$ chance node.

Conditional probabilities get applied to the $\mathcal{H}$ chance nodes to the right of the $\mathcal{S}$ chance node. We propagate the expected net returns right-to-left as probability-weighted net returns. Our "Drill/No-Drill" decision nodes $\mathfrak{D}$ are the conditional decisions. We elect to drill or not to drill based on the expected net returns of the nodes to the right of each. The decision to drill $\mathfrak{D} \ni \mathfrak{d} =$ Yes represents the optimum decision in all cases except for one. If the seismic survey reveals no structure, the likelihood of a "Dry" well is is high enough for us to have an expected loss. We therefore select $\mathfrak{D} \ni \mathfrak{d} =$ No.

What then is the value of the information obtained from the seismic survey $\mathfrak{T}$? We see that the expected net return given testing is $\mathcal{R}^{\text{test}} = \$22,500$. The return without testing is $\mathcal{R}^{\text{no test}} = \$20,000$. The optimum decision then is to test and proceed with drilling if $\mathcal{S} \ni s \neq$ None. From (10), the information value $\mathcal{V}_{\text{test result}} = \$12,500$. Paying \$10,000 to receive \$12,500 appears to be a bargain. Moreover, the expected marginal return on information from the test from (11) becomes

$$\mathcal{M}_{\text{test result}} = \frac{\$2,500}{\$10,000} = 25\%$$

for a geophysical-survey cost of \$10,000. In this case, the return is quite high.

A parameter study allows us to explore how the value of information relates to uncertainty. To proceed, we need a joint distribution $\mathcal{P}(s,h)$ over which we can exercise simple control. We set aside considerations here about phenomenological accuracy. This is a thought experiment.

How do we accomplish this? One approach involves constraining one of our marginals and specifying all elements in $\mathcal{P}(s,h)$ using a single parameter. Table 4 provides just such a mechanism. We sweep our α factor over a range of values. For each value of α we solve for the $\{a_1, a_2 a_3\}$ using the linear equation

$$\begin{pmatrix} 1 & \alpha & \alpha^2 \\ \alpha & 1 & \alpha \\ \alpha^2 & \alpha & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix}. \tag{12}$$

For this analysis we want $\alpha \ll 1$. When α is very small, the matrix is mostly-diagonal. This amounts to a case in which the expected outcome conditioned on a geological state $\mathcal{P}(h \mid s)$ will be high for one h for each s. Our test is very-informative. As α approaches unity, conditioning on s looses its ability to discriminate between states in $\mathcal{H}$.

To elaborate, for $\alpha \ll 1$ our uncertainty is dominated by the marginals $\{m_1, m_2, m_3\}$. This represents our irreducible uncertainty ([Keynes1921]) for this scenario. The uncertainty in $\mathcal{P}(h,s)$, which has substantial discriminate power. As α becomes larger, the uncertainty in $\mathcal{P}(h,s)$ comes to dominate. When it becomes sufficiently large, uncertainty in the marginals $\{m_1, m_2, m_3\}$ becomes diminsishingly significant.

Figure 10 shows just such an analysis. Employing the scheme from Table 4, we sweep α from 0.1 to 0.5 in steps of 0.025. A quick validity check shows us values from Strat's point study — K-B distance $\approx$ 0.11, ratio of diagonals $\approx$ 1.45, $\mathcal{V}_{\text{test result}} \approx \$2,500$ — fall close to these curves.

Secondly, these results corroborate our conjecture about the relationship between uncertainty and value

Table 4: Approach to synthesizing $\mathscr{P}(s,h)$ for purposes of parameter study.

	Structure, $\mathscr{S}$			
Outcome, $\mathscr{H}$	None	Open	Closed	Marginal
Dry (Dr)	a_1	$\alpha\, a_2$	$\alpha^2\, a_3$	$m_1 = 0.5$
Trickle (Tr)	$\alpha\, a_1$	a_2	$\alpha\, a_3$	$m_2 = 0.3$
Gusher (Gu)	$\alpha^2\, a_1$	$\alpha\, a_2$	a_3	$m_3 = 0.2$

of information. The unequivocal trend in this analysis holds that increasing certainty in $\mathscr{P}(h,s)$ increases the value of the associated information. Reduced uncertainty increases the predictability of outcomes. All parameters point in that direction.

Additionally, the point-parameters from the baseline case defined by Table 1 are characterized by a relatively high degree of uncertainty. The associated parameters fall in the lower-left-hand region of the space for which calculations in Figure 10 appear. We assume that the given probability structure is stylized to illustrate the methodology.

We see at the boundaries of the parameter study that the value of the information can swing negative. What might be occurring here? It seems that the information from the geophysical survey is of such low quality that its use is counterproductive. The information serves to avoid the case in which $\mathscr{S} \ni s =$ None. Some probability of a favorable outcome exists in this scenario. If we cannot however distinguish between favorable and unfavorable outcomes with sufficient precision, then using the information is counterproductive.

3.2 Decision Analysis Using Theory of Evidence

We now turn to the ToE-based framing of the decision case study. As is obvious from the previous section, ToE analysis is considerably more-complex than Bayesian analysis. This should be clear from equations (3a)-(3e). That we assign partial belief should also contribute to this recognition. The payoffs however arise from a more-complete characterization of uncertainty. Marrying ToE to information modeling from Information Economics offers promise of tighter linkage to valuation.

We now walk through for ToE decision analysis a procedure that resembles that performed for the Bayesian analysis above. We begin with basic-belief assignments (bbas) on our frame of discernment. We must transform our bbas into Bayesian probabilities. We accomplish that using the pignistic transform (5). This gives us the bba-derived Bayesian probabilities with which we can populate the decision tree. We attempt moreover to characterize the degree of uncertainty and to compare it with the Bayesian approach.

Table 5 contains our starting point. This contains the bba assignments on our frame of discernment. We elicit this information either through subject-matter experts or through analytical approaches when possible. As with the real world, ambiguity is inescapable. This elicitation approach accounts for that.

Now, the contents of Table 5 are conditional bbas. The are "conditioned" on the geological structure $\mathscr{S}$. Adopting Smets' notation, $m^{\mathscr{H}}[s](h)$ denotes the degree of belief attributed to the proposition that outcome $h \subseteq \mathscr{H}$ is correct given evidence that the geophysical structure is characterized by $s \subseteq \mathscr{S}$. Note that both h and s are set objects within their respective frames of discernment. We see in Table 5 that many of our beliefs are vacuously attributed to combinations of propositions.

For example, $m^{\mathscr{H}}[\{\text{None}\}](\{\text{Dr},\text{Tr}\}) = 0.4$ indicates that, were we presented with certain evidence of the geological structure $\mathscr{S} \supseteq s = \{\text{None}\}$, we would belief that the likelihood is 0.4 that the outcome will be either a "Dry" well or a "Trickle". Our evidence does not support the proposition that one or the other is more or less likely. Unless we received additional evidence, we at best can merely say that one or the other is valid.

Given that Table 5 contains conditional bbas, we need additional information in order to obtain the marginals and the joint. The "prior" belief about the geological state $\mathscr{S}$ is similarly given to us vacuously. This appears in Table 6.

These values were originally provided by Strat [Strat1994], who also gave us the Bayesian joint probability distributions $\mathscr{P}(h,s)$ from Table 1. One surmises that Tables 5 and 6 were deliberately concocted to offer a ToE-equivalent to a set of Bayesian probabilities. "Fuzzing" things up seems like a plausible proposition about the author's intent. That the values of the "prior" bbas in Table 6 correspond to the marginal distributions in Table 1 probably was not coincidental.

Getting the marginal bbas on $m^{\mathscr{H}}[m^{\mathscr{S}}]$ shown in Table 5 involves a nontrivial exercise. Smets de-

Figure 10: Parameter study demonstrating relationship between Bayesian-uncertainty structure and value of information.

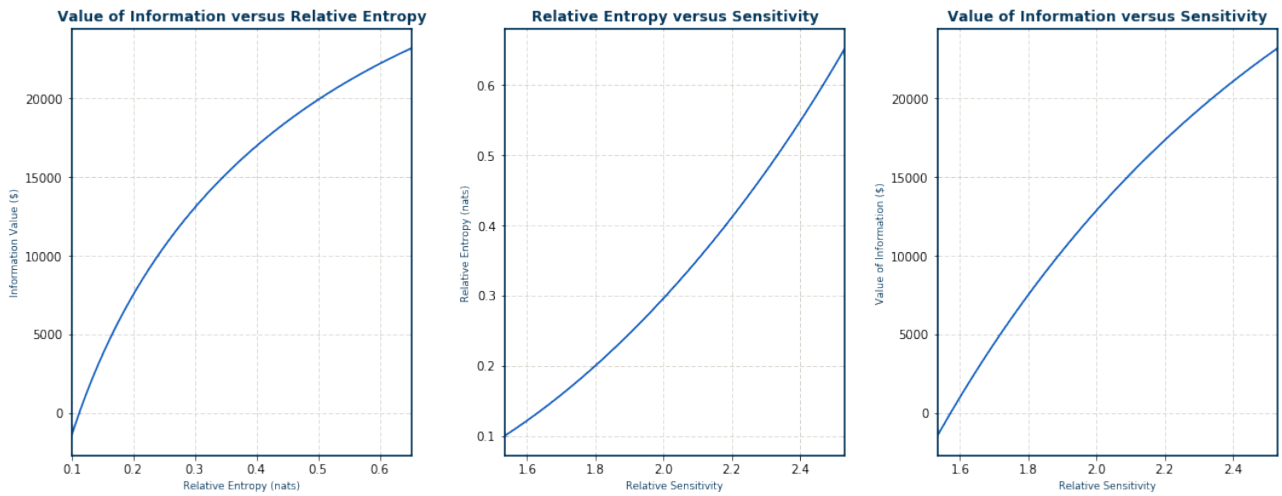


Table 5: Conditional basic-belief assignments $m^{\mathscr{H}}[s](h)$ for Theory-of-Evidence Transferable-Belief Method (TBM). (From [Strat1994].)

	Structure, $\mathscr{S}$			
Outcome, $\mathscr{H}$	$m^{\mathscr{H}}[\{\text{None}\}]$	$m^{\mathscr{H}}[\{\text{Open}\}]$	$m^{\mathscr{H}}[\{\text{Closed}\}]$	Marginal, $m^{\mathscr{H}}[m^{\mathcal{S}}]$
{Dr}	0.45	0.10	0.0	0.2385
{Tr}	0.0	0.0	0.0	0.0
{Gu}	0.0	0.0	0.25	0.0
{Dr, Tr}	0.40	0.30	0.0	0.2885
{Dr, Gu}	0.0	0.0	0.0	0.0330
{Tr, Gu}	0.0	0.30	0.30	0.0050
{Dr, Tru, Gu}	0.15	0.0	0.45	0.4350
$\sum_{h \subseteq \mathscr{H}} m^{\mathscr{H}}[s](h)$	1.0	1.0	1.0	1.0

Table 6: "Prior" basic beliefs $m^{\mathcal{S}}(s)$ about the geological state of the target oil well.(From [Strat1994].)

Structure, $\mathcal{S}$	$m^{\mathcal{S}}(s)$
{None}	0.5
{Open}	0.0
{Closed}	0.0
{None, Open}	0.3
{None, Closed}	0.0
{Open, Closed}	0.2
{None, Open, Closed}	0.0
$\sum_{s\subseteq\mathcal{S}} m^{\mathcal{S}}(s)$	1.0

scribed the technique at length in [Smets1993] describing his *Disjunctive Rule of Combination*. These values can be obtained through a four-step method implied by [Smets2002, p. 47]. We first translate each $m^{\mathscr{H}}[s](h)$into its corresponding Belief Function $Bel^{\mathscr{H}}[s](h)$ using (3b). Secondly, we obtain an intermittent quantity

$$Bel[\mathcal{X}](h) = \prod_{s\in\mathcal{X}} Bel^{\mathscr{H}}[s](h).$$

The quantity $Bel[\mathcal{X}](h)$ is essentially a joint belief function on sets of structural states instead of distinct ones. We next get the marginal belief function by applying our "prior" $m^{\mathcal{S}}(s)$,

$$Bel^{\mathscr{H}}\left[m^{\mathcal{S}}\right](h) = \sum_{\mathcal{X}\subseteq\mathcal{S}} m^{\mathcal{S}}(\mathcal{X})\ Bel^{\mathscr{H}}[\mathcal{X}](h)\ \forall h\subseteq\mathscr{H}. \tag{13}$$

We finally recover $m^{\mathscr{H}}\left[m^{\mathcal{S}}\right](h)$ from $Bel^{\mathscr{H}}\left[m^{\mathcal{S}}\right](h)$ using the Möbius-transform inversion (3e).

Further analysis of [Smets1993] may yield more-direct transformations. Also, simply multiplying $m^{\mathcal{S}}(\mathcal{X})\ Bel^{\mathscr{H}}[\mathcal{X}](h)$ in (13) without evaluating the summation produces the "joint" belief function $Bel^{\underline{\Omega}}(h,s)$ on the frame of discernment $\underline{\Omega} = \mathscr{X}\times\mathcal{S}$. The summation over $\mathcal{X}\subseteq\mathcal{S}$ accomplishes a marginalization along the lines of (4).

A reiteration of our notation will probably help to clarify things. We said previously that $m^{\mathscr{H}}[s](h)$ represents the belief assigned to some proposition about our outcome associated with a focal element $h\subseteq\mathcal{H}$ given certain evidence that the geological structure is given by $s\subseteq\mathcal{S}$.

We in fact did not receive certain evidence about the geological structure. We received a basic-belief assignment $m^{\mathcal{S}}(s)\quad\forall s\subseteq\mathcal{S}$. Our conditional bba $m^{\mathscr{H}}\left[m^{\mathcal{S}}\right](h)$ then represents our belief about the proposition for a given outcome $h\subseteq\mathcal{H}$ given the aggregate evidence represented by $m^{\mathcal{S}}(s)$.

Figure 11 provides graphical depictions of belief structure on our outcome frame of discernment $\mathscr{H}$. It includes for each column in Table 5 the bbas $m^{\mathscr{H}}[s](h)$, the corresponding belief functions $Bel^{\mathscr{H}}[s](h)$, and the plausibility functions $Pl^{\mathscr{H}}[s](h)$. Its study is instructive for many reasons. First, one enjoys a visual glimpse of how the beliefs are allocated. We see how in $Bel^{\mathscr{H}}[s](h)$, beliefs provided initially by $m^{\mathscr{H}}[s](h)$ flow from most-specific towards most-vacuous. For example, $m^{\mathscr{H}}[\{\text{None}\}](\{\text{Dr}\})$ influences our beliefs $Bel^{\mathscr{H}}[\{\text{None}\}](\{\text{Dr},\text{Gu}\})$. But belief attribution does not flow in the opposite direction.

Plausibility functions are even-more interesting. In calculating plausibilities according to (3c) we look for combinations such that $\omega\cap\varpi\neq\emptyset$ for $\omega,\varpi\subseteq\mathscr{H}$. In Table 5 we attribute finite basic belief $m^{\mathscr{H}}[\{\text{None}\}](\{\text{Dr},\text{Tr}\})$. In the plausibility calculation, this flows to $Pl^{\mathscr{H}}[\{\text{None}\}](\{\text{Dr}\})$ because $\{\text{Dr},\text{Tr}\}\cap\{\text{Dr}\}\neq\emptyset$.

Now returning to our objective of framing a our decision, we must translate each of the bbas in Tables 5 and 6 into Pignistic Transforms using (5). These appear in Tables 7 and 8. These are the Bayesian probabilities derived from our ToE representation for application to the decision tree.

The probabilities on $\mathscr{H}$ in table 7 for the distinct structural states {None}, {Open}, and {Closed} serve the same purpose as the conditional probabilities from Table 2 do in the Bayesian framing. The $Bet_{P\mathscr{H}}[m^{\mathcal{S}}](s)$ functions like the marginal $\mathcal{P}(S)$. And the $Bet_{P\mathcal{S}}(s)$ quantities from Table 8 are equivalent to the Bayesian marginals $\mathcal{P}(h)$.

Figure 12 shows the framing using the probabilities of Tables 7 and 8. Some similarities exist between Figure 12 and Figure 9. In both cases, the optimum decision $\mathfrak{D}\ni\mathfrak{d}^{\text{opt}}=$ Drill. The value of information is approximately the same, also. The value of information $\mathcal{V}_{\text{test result}}\approx\$11,860$ for the ToE analysis, which compares rather closely to $\mathcal{V}_{\text{test result}}\approx\$12,500$ for the Bayesian approach. The expected marginal return on information

$$\begin{aligned}\mathcal{M}_{\text{test result}} &= \frac{\$1,855}{\$10,000}\\ &= 18.6\%\end{aligned}$$

Figure 11: Conditional and marginal belief structures based on basic-belief assignments in Table 5.

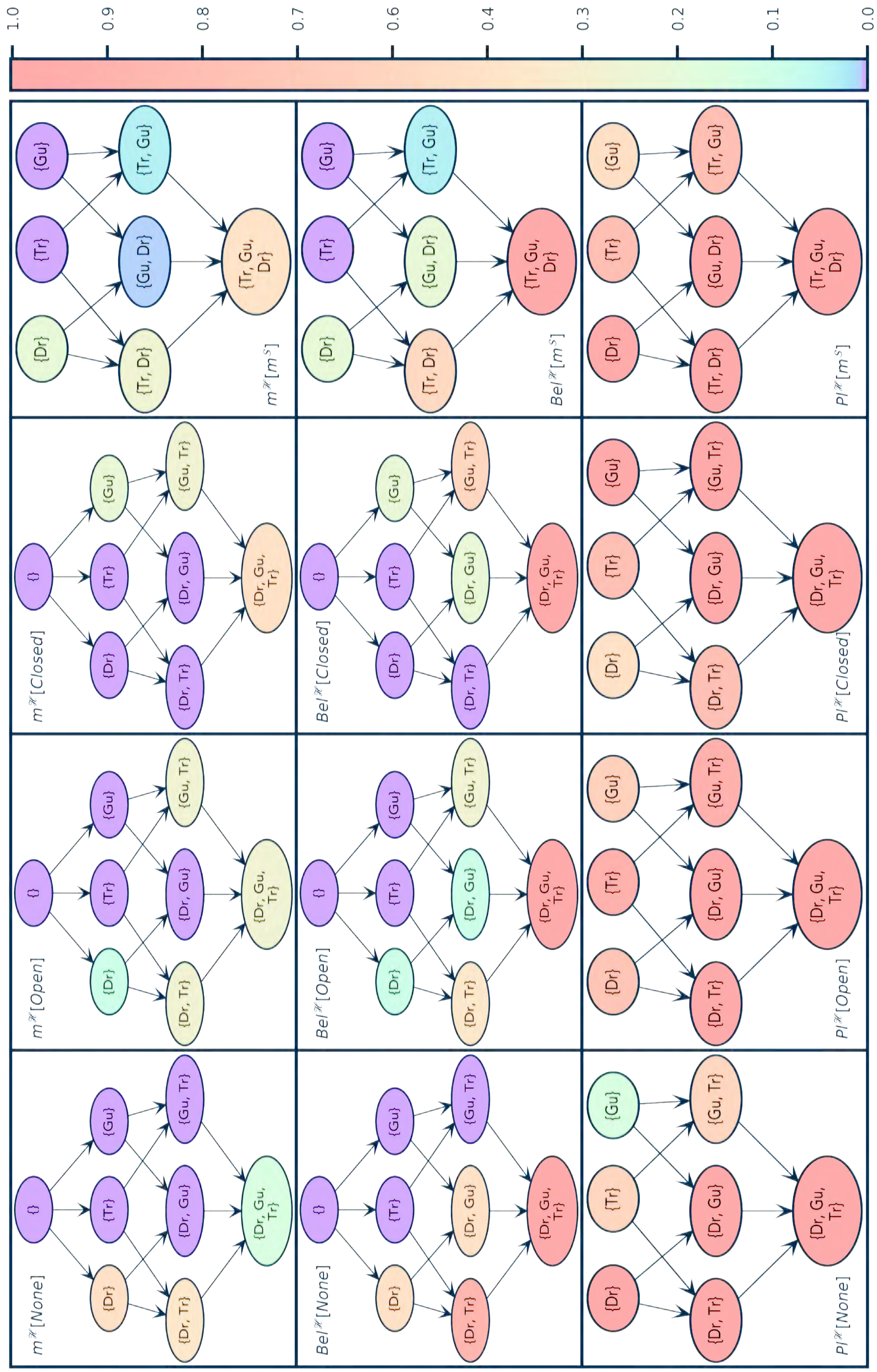


Table 7: Pignistic transforms $Bet_{P\mathscr{H}}[s](h)$ of basic-belief assignments on $\mathscr{H}$ from Table 5. (After [Smets2002]).

Outcome $\mathscr{H}$	Structure, $\mathscr{S}$ {None}	{Open}	{Closed}	$m^{\mathscr{S}}$
{Dr}	0.70	0.35	0.15	0.5303
{Tr}	0.25	0.25	0.55	0.1640
{Gu}	0.05	0.40	0.30	0.3058
$\sum_h$	1.0	1.0	1.0	1.0

Table 8: Pignistic transforms $Bet_{P\mathscr{S}}(s)$ of basic-belief assignments on $\mathscr{S}$ from Table 6. (After [Smets2002]).

Structure, $\mathscr{S}$	$Bet_{P\mathscr{S}}(s)$
{None}	0.65
{Open}	0.25
{Closed}	0.1
$\sum_s$	1.0

still seems quite high at a geophysical-survey cost of $10,000.

The expected net return however is much lower in the case of the ToE analysis compared to the Bayesian analysis. In the latter we expect for $\mathscr{R}^{\text{opt}}$ to yield $22,500. The ToE analysis only yields an expected net return of $12,850.

What is occurring here? The ToE yields much-lower degrees of certainty.

An exercise in KL-distance calculation allows for exploration of this hypothesis. We seek to compare the Bayesian-probability structure of $Bet_{P\mathscr{H}}[s](h)$ with that for $\mathscr{P}(h,x)$ in our Bayesian decision framing. Our $Bet_{P\mathscr{H}}[s](h)$ quantities from Table 7 are effectively conditional probabilities conditioned on $Bet_{P\mathscr{S}}(s)$ from Table 8. Converting $Bet_{P\mathscr{H}}[s](h)$ gives us $Bet_{P\mathscr{H}\times\mathscr{S}}(h,s)$. This "joint" distribution is on a compatible frame of discernment with that for $\mathscr{P}(h,x)$.

A few lines of python logic produces Table 9 containing values for $Bet_{P\mathscr{H}\times\mathscr{S}}(h,s)$. Direct comparison of this table with $\mathscr{P}(h,s)$ in Table 1 immediately reveals a more-pessimistic outlook by the former. Overall, the pignistic transforms indicate a 25% lower chance in getting a "Gusher" well. We are considerably less-likely to find conditions favorable for extraction, also.

As for our Kl-Distance calculation, $\mathcal{D}\Big(\mathscr{P} \parallel Bet_{P\mathscr{H}\times\mathscr{S}}\Big) =$

Figure 12: Decision-tree framing of case-study decision using Theory-of-Evidence Transferable-Belief Method (TBM). (After [Smets2002].)

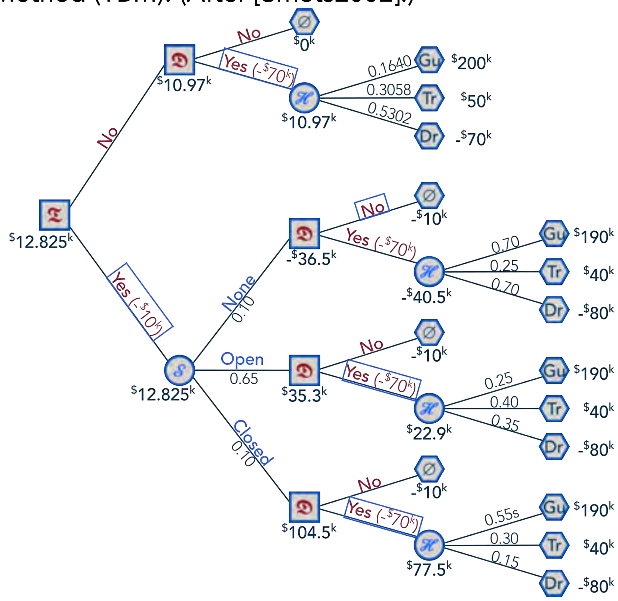


Table 9: "Joint" probability distribution $Bet_{P\mathscr{H}\times\mathscr{S}}(h,s)$ obtained from combining $Bet_{P\mathscr{H}}[s](h)$ from Table 7 with $Bet_{P\mathscr{S}}(s)$ from Table 8

Outcome, $\mathscr{H}$	Structure, $\mathscr{S}$ None	Open	Closed	Marginal
Dry (Dr)	0.4550	0.0875	0.015	0.5575
Trickle (Tr)	0.1625	0.1000	0.055	0.2925
Gusher (Gu)	0.0325	0.0625	0.030	0.1500
Marginal	0.65	0.25	0.10	1.00

-0.153. This distance is greater than $\mathcal{D}\Big(\mathcal{P} \parallel \mathcal{P}_{\mathcal{X}}\mathcal{P}_{\mathcal{Y}}\Big) = -0.111$, comparing from the Bayesian framing the outer product of the marginals to the original joint distribution $\mathcal{P}\left(h, x\right)$ of Table 1. That distances with respect to $Bet_{P\mathcal{H}\times S}$ and $\mathcal{P}_{\mathcal{X}}\mathcal{P}_{\mathcal{Y}}$ both take negative signs indicates that they contain less information than $\mathcal{P}\left(h, x\right)$. The comparison distributions exceed $\mathcal{P}\left(h, x\right)$ where it is small, and are lower than it for values at which it is larger. Assuming that $\mathcal{P}\left(h, x\right)$ is relatively informationally-rich, both contain less information, more uncertainty. Moreover, $Bet_{P\mathcal{H}\times S}$ contains less information than $\mathcal{P}_{\mathcal{X}}\mathcal{P}_{\mathcal{Y}}$.

All of this is consistent with the proposition that the Bayesian framing asserts a greater degree of certainty than the ToE. Whether one is more "accurate" is difficult to tell. Following the metrology community's philosophy, asking about "accuracy" can represent a dubious line of inquiry. We should instead focus on clarifying our actual belief about our condition of uncertainty. The ToE decision-framing leaves us less-certain than the Bayesian framing.

3.2.1 Interpreting our Case Study in the Context of our Central Conjectures

Beyond exercising DA machinery, seek from our case study corroboration of our central conjecture about distinct modes by which information-differentiation yields competitive differentiation. We can interpret aspects of our case study to corroborate our conjecture analogously. More-direct corroboration awaits more-realistic case studies.

Deciding whether or not to perform the geophysical survey adds information in the same way that differentiated access does in the more-general sense. We argue in (1a) - (1c) that differentiated access allows decision-making using conditional probabilities when competitors are working with marginals. Moreover, this additional information should reduce the uncertainty. Introduction of information from the geophysical survey certainly contributes in that manner.

What of differentiated insights? Our generalized formulation (1a) - (1c) treats differentiated insights as another independent dimension on which to condition our joint probability distribution $\mathcal{P}$. Our case study only offers one such opportunity. Differentiated insights might alternatively reduce uncertainty in another fashion. Employing a different mechanism, our "synthetic" joint distributions constructed by (12) certainly does this.

Finally, differentiated ability to act is accounted for in our return function $\mathcal{R}$. We did not demonstrate here any variation in $\mathcal{R}$. As formulated in (1a) - (1c), $\mathcal{R}$ provides a mechanism — unexercised here — to demonstrate differentiated utility-realization efficiency. Other attributes of differentiated competency are required for a more-complete framework. This might include agility, predictability, and others. Our conjectured information-differentiation modes are presently articulated in operational terms. Operational differentiation is competitively important [Sadun2017]. Realizing full benefits of the formalism's

4 Conclusions and Further Directions

A conjecture that three distinct modes of informationally-derived modes of competitive differentiation follows directly from a basic principle of Information Economics. These modes include differentiated information access, differentiated insights from information that may be commonly available, and differentiated ability to act on information. Organizations can exercise these modes individually, or in concert with each other.

These modes can be represented using a mathematical formalism finding its roots in Decision Analysis. It arises from extending decision analysis with concepts and methods from Information Economics, the Theory of Evidence, Information Theory, and competency-focused frameworks for competitive differentiation. The formalism appears very general. It should be extensible to a variety of distinct operating models and strategic positions.

The formalism's use of the Theory of Evidence imbues it with power to more-richly characterize uncertainty than arises from traditional Bayesian-based decision-analysis frameworks. This leads to less-optimistic estimates of expected net returns. These lower estimated values seem consistent with the proposition that traditional Bayesian theory can lead to understatement of uncertainty. Use of the Theory of Evidence occurs at the expense of descriptiveness. Visualization and explanation require further work.

Extending the formalism into a generalized graphical-belief model along the lines of [Almond1995] represents an obvious next step. We demonstrate here the most-important computational aspects of such a methodology. What remains is an algorithmic framework to build decision trees based on specification of arbitrary lists of chance, decision, and outcome nodes, and from a list of graph edges by which they are associated.

The Theory-of-Evidence facet of the formalism is well-positioned for convergence with other related disciplines. Information Modeling [Samuelson2004] from Information Economics represents one such opportunity. Integration of the two methodologies would strengthen its information-valuation features.

Theory of Evidence is moreover converging towards fuzzy-logic theory (e.g, [Skalna2015]). This would further enrich our formalism's ability to more-extensively characterize uncertainty. As exemplified in our case study, specific quantities — most-conspicuously, expected net returns — are specified as discrete, crisp values. These quantities are actually better-characterized as distributions. So-specifying them using fuzzy logic would provide a more-complete view of the state of uncertainty underlying a scenario. This is however accompanied by the burden of more-transparent explanation and visualization.

Finally, closer-alignment with concepts from strategy will benefit the formalism's utility. The criteria for what makes a framework strategic are often narrow (e.g., [VanddenSteen2016, Leiblein2018]).

References

[Agresti2002] Agresti A (2002) *Categorical Data Analysis*, second ed, (J. Wiley & Sons, Hoboken, NJ).

[Amram1999] Amram M, Kulatilaka N (1999) *Real Options: Managing Strategic Investment in an Uncertain World* (Harvard Business School Press, Boston, MA).

[Almond1995] Almond R (1995) *Graphical Belief Modeling* (Chapman & Hall, London, UK).

[Ben-Naim2008] Ben-Naim A (2008) *Statistical Thermodynamics Based on Information: A Farewell to Entropy* (World Scientific, Hackensack, NJ).

[Birchler1999] Birchler U, Bütler M (1999) *Information Economics* (Routledge, New York, NY).

[Briggs2016] Briggs W (2016) *Uncertainty: The Soul of Modeling, Probability, & Statistics* (Springer International Publishing AG, Switzerland).

[Brunnermeier2001] Brunnermeier M K (2001) *Asset Pricing under Asymmetric Information: Bubbles, Crashes, Technical Analysis, and Herding* (Oxford University Press, New York, NY).

[Bryce2007] Bryce D J, Dyer J H (2017) Strategies to Crack Well-Guarded Markets. *Harvard Business Review* May 2007, https://hbr.org/2019/04/5-myths-about-strategy.

[Cattani2018] Cattani G, Sands D, Porac J, Greenberg J (2018) Competitive Sensemaking in Value Creation and Capture. *Strategy Science* 3(4):632-657.

[Chomsky2014] Chomsky N (2014) *Aspects of the Theory of Syntax* (MIT Press, Cambridge, MA).

[Christensen2004] Christensen C M, Anthony S D, Roth E A (2004) *Seeing What's Next ? Using Theories of Innovation to Predict Industry Change* (Harvard Business School Press, Boston, MA).

[Christensen2015] Christensen C M, Raynor M E, McDonald R (2015) What Is Disruptive Innovation? *Harvard Business Review* December 2015, https://hbr.org/2015/12/what-is-disruptive-innovation.

[Cobb2006] Cobb B R, Shenoy P P (2006) On the plausibility transformation method for translating belief function models to probability models. *International Journal of Approximate Reasoning* 41:314-330.

[Cover1991] Cover T M, Thomas J A (1991) *Elements of Information Theory* (Wiley-Interscience, New York, NY).

[Chow1997] Chow S Y, Teicher H (1997) *Probability Theory: Independence, Interchangeability, Martingales* (Springer, NY).

[Dempster1967] Dempster A P (1967) Upper and Lower Probabilities Induced by a Multivalued Mapping. *The Annals of Mathematical Statistics* 38(2):325-339.

[Denrell2001] Denrell J, March J G (2001) Adaptation as Information Restriction: The Hot Stove Effect. *Organization Science* 12(5):523?538.

[Dixit1994] Dixit A K, Pindyck R S (1994) *Investment Under Uncertainty* (Princeton University Press, Princeton, NJ).

[Drucker1992] Drucker P F (1992) The New Society of Organizations. *Harvard Business Review* September-October 1992, https://hbr.org/1992/09/the-new-society-of-organizations.

[Drucker1993] Drucker P F (1993) *Post-Capitalist Society* (Harper-Collins eBook, NY, 2009).

[Drucker1999] Drucker P F (1999) Knowledge-Worker Productivity: The Biggest Challenge. *California Management Review* 41(2):79-94.

[deFinetti2017] de Finetti B (2017), *Theory of Probability: A Critical Introductory Treatment* (John Wiley & Sons, 2017, West Sussex, UK). (Originally published *Teoriea Delle Probabilita*, Wiley, 1970.)

[Freeling1983] Freeling A N S, Sahlin N I Combining Evidence. *Evidentiary Value: Philosophical, Judicial, and Psychological Aspects of a Theory* (C W K Gleerups, Lund, Sweeden).

[Gallagher1968] Gallagher R G (1968) *Information Theory and Reliable Communication* (John Wiley & Sons, New York).

[Groysberg2018] Groysberg B, Lee J, Price J, Cheng J Y J (2018) The Leader's Guide to Corporate Culture. *Harvard Business Review* January-February 2018, https://hbr.org/2018/01/the-culture-factor.

[Grant1996] R. M. Grant (1996) Toward a knowledge-based theory of the firm. *Strategic Management Journal* 17:109-122.

[Hamlett2019] Hamlett N A (2019) Competitive Advantage Based on Information (unpublished working paper).

[Hogg2013] Hogg R V, McKean J w, Craig A T (2013) *Introduction to Mathematical Statistics*, seventh ed, (Pearson, Boston, MA).

[Howard2016] R. A. Howard, A. E. Abbas, *Foundations of Decision Analysis*, Upper Saddle River, NJ: Pearson, 2016.

[Iansiti2017] Iansiti M, Lakhani K R (2017) Managing Our Hub Economy. *Harvard Business Review* September-October 2017, https://hbr.org/2017/09/managing-our-hub-economy.

[Jiroušek2018] Jiroušek R, Shenoy P P, A (2018) A new definition of entropy of belief functions in the Dempster-Shafer theory. *International Journal of Approximate Reasoning* 92:49-65.

[Kahneman2011] Kahneman D (2011) *Thinking Fast and Slow* (MacMillan, NY).

[Kahneman1979] Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. *Econometrica* 47(2)263-292.

[Kay1993] Kay S M (1993) *Fundamentals of Statistical Signal Processing: Estimation Theory* (Prentice Hall, Englewood Cliffs, NJ).

[Keelin2016] Keelin T W (2016) The metalog distributions. *Decision Analysis* 13(4)243-277.

[Keynes1921] Keynes J M (1921) *A Treatise on Probability* (MacMillan and Co., Ltd., London).

[Kochebnderfer2015] Kochenderfer M J (2015) *Decision Making Under Uncertainty* (MIT Press, Boston, MA).

[vonKrogh1994] von Krogh V, Roos J, Slocum K (1994) An Essay on Corporate Epistemology. *Strategic Management Journal* 15(S2):53-71.

[Kramosil2001] Kramosil I (2001) *Probabilistic Analysis of Belief Functions* (Kluwer Academic/Plenum, NY).

[Laffont1989] Laffont J J, Bonin J P, Bonin H (1989) *The Economics of Uncertainty and Information*, (Harvard Business School Press, Cambridge, MA).

[Leiblein2018] Leiblein M J, Reuer J J, Zenger T (2018) What Makes a Decision Strategic? *Strategy Science* 3(4):558-573.

[Lewis2004] Lewis M (2004) *Moneyball: The Art of Winning an Unfair Game* (W. W. Norton & Co, New York, NY).

[Lindley1983] Lindley D (1983) Reconciliation of probability distributions. *Operations Research* 31(5):866-880.

[Lynch2017] Lynch G S (2017) *Uncertainty Advantage: Leadership Lessons for Turning Risk Outside-In* (Archway Publishing, Bloomington, IN).

[Mallat1989] Mallat S G (1989) Mulstiresolution approximations and wavelet orthonormal bases of $\mathbb{L}^2(\mathbb{R})$ *Transactions of the American Mathematical Society* 315(1):69-87.

[Martin2015] Martin R L (2015) Strategy Is About Both Resources and Positioning. *Harvard Business Review* April 27, 2015, https://hbr.org/2015/04/strategy-is-about-both-resources-and-positioning.

[Myerson1991] Myerson R B (1991) *Game Theory* (Harvard University Press, Cambridge, MA).

[Nickerson2018] Nickerson J, Argyres N (2018) Strategizing Before Strategic Decision Making. *Strategy Science* 3(4):592-605.

[Papayoanou2010] Papayoanou P (2010) *Game Theory for Business: A Primer in Strategic Gaming* (Probabilistic Publishing, USA).

[Pettigrew2016] Pettigrew R (2016) *Accuracy and the Laws of Credence* (Oxford University Press, Oxford, UK).

[Pisano2019] Pisano G P (2019) The Hard Truth About Innovative Cultures. *Harvard Business Review* January-February 2019, https://hbr.org/2019/01/the-hard-truth-about-innovative-cultures.

[Phlips1988] Phlips L (1988) *The Economics of Uncertain Information*, (Cambridge University Press, Cambridge, UK).

[Porter1985] Porter M E, Millar V E (1985) How Information Gives You Competitive Advantage. *Harvard Business Review* July 1985, https://hbr.org/1985/07/how-information-gives-you-competitive-advantage.

[Porter1996] Porter M (1996) What Is Strategy? *Harvard Business Review* November-December 1996, https://hbr.org/1996/11/what-is-strategy.

[Porter2014] Porter M E, Heppelmann J E (2014) How Smart, Connected Products Are Transforming Competition. *Harvard Business Review* November 2014, https://hbr.org/2014/11/how-smart-connected-products-are-transforming-competition.

[Porter2015] Porter M E, Heppelmann J E (2015) How Smart, Connected Products Are Transforming Companies. *Harvard Business Review* October 2014, https://hbr.org/2015/10/how-smart-connected-products-are-transforming-companies.

[Prasad1997] Prasad L, Iyengar S S (1997) *Wavelet Analysis with Applications to Image Processing* (CRC Press, Boca Raton, FL).

[Brealy2008] Brealey R A, Myers S C, Allen F (2008) *Principles of Corporate Finance* ninth ed (McGraw-Hill, NY).

[Raiffa1970] H. Raiffa, Decision Analysis (1970) *Introductory Lectures on Choices Under Uncertainty* (Addison-Wesley, Reading, MA).

[Redman2015] Redman T C (2015) 4 Business Models for the Data Age. *Harvard Business Review* May 20, 2015, https://hbr.org/2015/05/4-business-models-for-the-data-age.

[Reeves2011] Reeves M, Deimler M (2011) Adaptability: The New Competitive Advantage. *Harvard Business Review* July-August 2011, https://hbr.org/2011/07/adaptability-the-new-competitive-advantage.

[Reeves2012] Reeves M, Deimler M (2012) *Adaptive Advantage: Winning Strategies for Uncertain Times* (Boston Consulting Group, Boston, MA).

[Salicone2018] Salicone S, Prioli M (2018) *Measuring Uncertainty within the Theory of Evidence* (Springer International Publishing AG, Cham, Switzerland).

[Samuelson2004] Samuelson L (2004) Modeling Knowledge in Economic Analysis. *Journal of Economic Literature* 42(2):397-403.

[Seirafi2013] Seirafi K (2013) *Organizational Epistemology* (Physica-Verlag, Heidelberg).

[Shafer1976] Shafer G (1976) *A Mathematical Theory of Evidence* (Princeton University Press, Princeton, NJ).

[Shannon1948] Shannon C E (1948) A Mathematical Theory of Communication. *Bell Systems Technical Journal* 27:379-423.

[Shapiro1998] Shapiro C, Varian H R (1998) *Information Rules: A Strategic Guide to the Network Economy* (Harvard Business Review Press, Boston, MA).

[Skalna2015] Skalna I, *et al* (2015) *Advances in Fuzzy Decision Making: Theory and Practice* (Springer, Cham, SW).

[Smets1988] Smets P (1988) Belief Functions. *Non-Standard Logics for Automated Reasoning* (Smets P, Mamdani, EH, Dubois D, Prade H (Eds), Academic Press, San Diego, CA).

[Smets1993] Smets P (1993) Belief functions: The disjunctive rule of combination and the generalized Bayesian theorem. *International Journal of Approximate Reasoning* 9:1-35.

[Smets1994] P. Smets and R. Kennis (1994) The Transferable Belief Model. *Artificial Intelligence* 66:191-234.

[Smets2002] Smets P (2002) Decision making in a context where uncertainty is represented by belief functions. *Belief Functions in Business Decisions*, Srivastava R P and Mock (Eds) (Physica-Verlag, Heidelberg) 18-61.

[Spetzler2016] Spetzler C, Winter H, Meyer, J. *Decision Quality: Value Creation from Better Decisions* (Wiley, NY).

[Sadun2017] Sadun R, Bloom N, Van Reenen J (2017) Why do we undervalue competent management? *Harvard Business Review*, September-October, 2017.

[Strat1994] Strat TM (1994) Decision analysis using belief functions. *Advances in the Dempster-Shafer Theory of Evidence*, Yager R R, Fedrizi M, Kacprzyik J (Eds), (Wiley, NY) : 275-330.

[Suarez2005] Suarez F, Lanzolla G (2005) The Half-Truth of First-Mover Advantage. *Harvard Business Review* April 27, 2015, https://hbr.org/2005/04/the-half-truth-of-first-mover-advantage.

[Teece2009] Teece D (2009) *Dynamic Capabilities* (Oxford University Press, New York, NY).

[Tajeddini2019] Tajeddini M A, Aalipour A, Safarinejadian B (2019) Decision fusion method for bearing-faults classification based on wavelet denoising and Dempster-Shafer Theory *Iran Journal of Science and Technology, Transactions on Electrical Engineering* 43:295-305.

[Ulrich2004] Ulrich C (2004) Valuation of IT Investments Using Real Options Theory. *Business & Information Systems Engineering* 5(5):331-341.

[VanAlstyne2016] Van Alstybne, M. W. Parker, G. G., Chouldary, S. P., (2016) Pipelines, Platforms, and the New Rules of Strategy. *Harvard Business Review*, April 2016, https://hbr.org/2016/04/pipelines-platforms-and-the-new-rules-of-strategy.

[VanddenSteen2016] Van den Steen E (2016) A formal theory of strategy. *Management Science* 63(8):2616-2636.

[Wartzman2014] Wartzman R (2014) What Peter Drucker Knew About 2020. *Harvard Business Review*: October 16, 2014, https://hbr.org/2014/10/what-peter-drucker-knew-about-2020.

[Wessel2012] Wessel M, Christensen C M (2012) Surviving Disruption. *Harvard Business Review* December 2011, https://hbr.org/2012/12/surviving-disruption.

[Zhong2017] Zhong Y (2017) A Theory of Semantic Information, Presented at the IS4SI 2017 Summit DIGITALIZATION FOR A SUSTAINABLE SOCIETY, Gothenburg, Sweden, 12?16 June 2017. *MDPI Proceedings*.

[Zhu2016] Zhu F, Furr N (2016) Products to Platforms: Making the Leap. *Harvard Business Review* April 2016, https://hbr.org/2016/04/products-to-platforms-making-the-leap.