236B Zachry Engineering Center
3131 TAMU
College Station, Texas 77843
979.845.4347
ebickel@tamu.edu

Optimal Sequential Exploration: A Binary Learning Model

Presented by:*
J. Eric Bickel
Industrial and Systems Engineering
Texas A&M University

Presented to:
DAAG 2006
Baltimore, Maryland

30-31 March 2006

*This is joint work with James E. Smith, Fuqua School of Business, Duke University

This paper is forthcoming in *Decision Analysis.* The paper and accompanying spreadsheet are on my website http://ie.tamu.edu/people/faculty/bickel/default.htm.

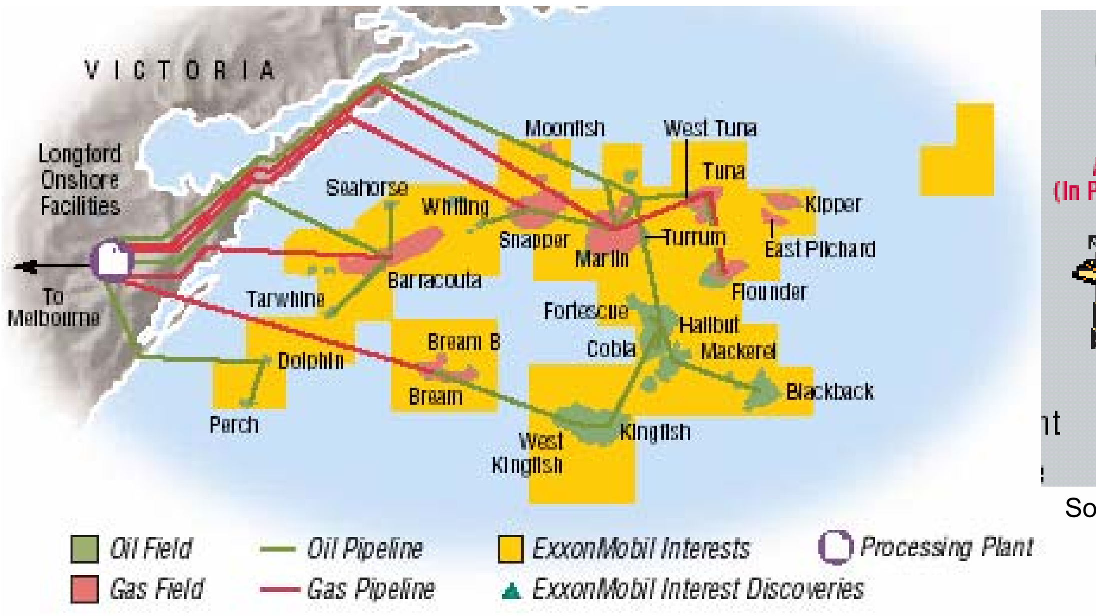
Agenda

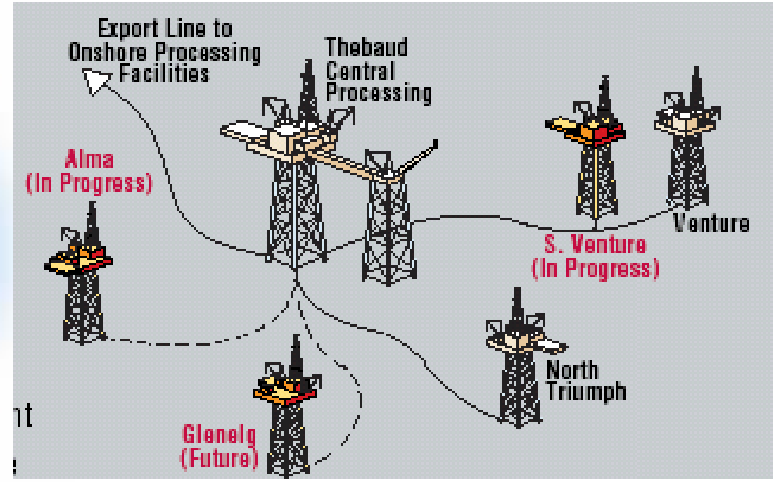
- The Challenge
- Determining the Joint Distribution
- Illustrative Example
- Conclusion (including life sciences example)

The motivation for this paper began with a consulting project to prioritize an oil and gas exploration program.

The Situation

- Ranking exploration drilling prospects and developing production plans.
- Multiple clusters comprised of several prospects each.
- The clusters were defined to include prospects that are geologically similar or prospects that would be co-developed.
  - What is the optimal drilling strategy?

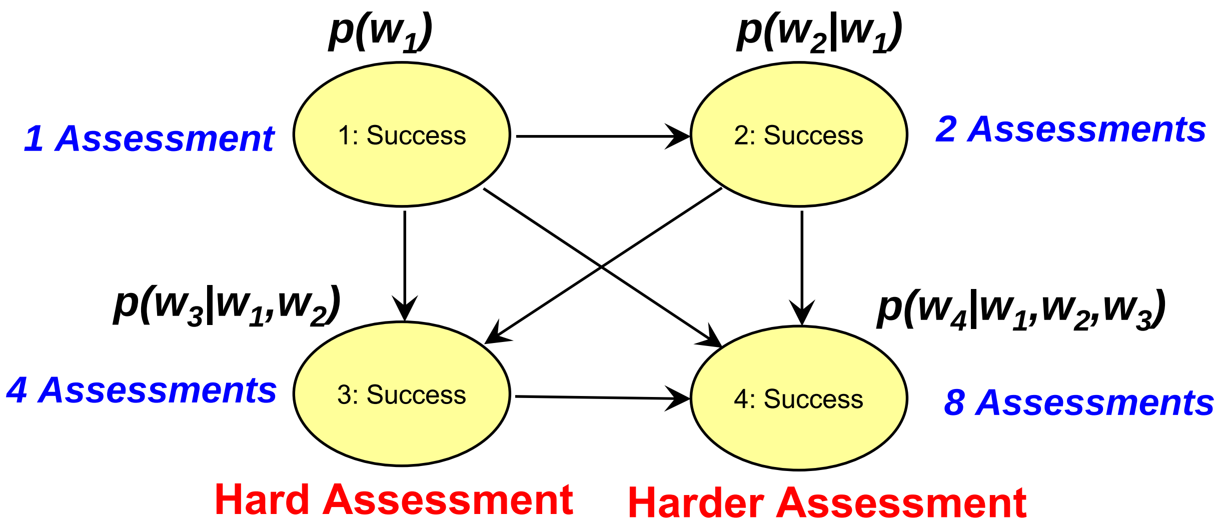




Source: ExxonMobil 2002 Annual Report

The client and details of the project have been disguised.

Even modestly sized clusters require many assessments—some of which are quite difficult.

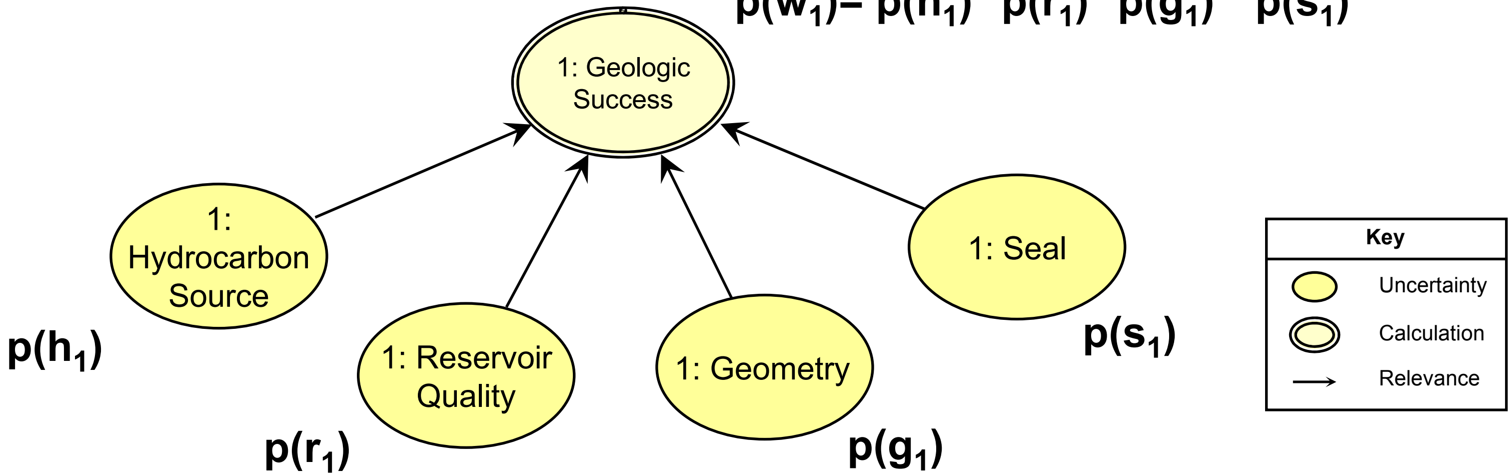


- Assessment of the joint distribution for *n* wells requires 2^n-1.
  - Four prospects would require 15 assessments.
  - Nine prospects would require 511 assessments.
- If we decompose the assessment into a series of conditional assessments, many will be heavily conditioned.
  - With 4 prospects the assessment of the 4th well will be conditioned on the outcome of 3 other wells and there will be $2^3=8$ different conditioning scenarios.
  - In the case of nine prospects the assessment for the 9th well will be conditioned on $2^8=256$ scenarios.

Yet, the problem is **even** harder than this!

The probability of geologic success is comprised of four factors.

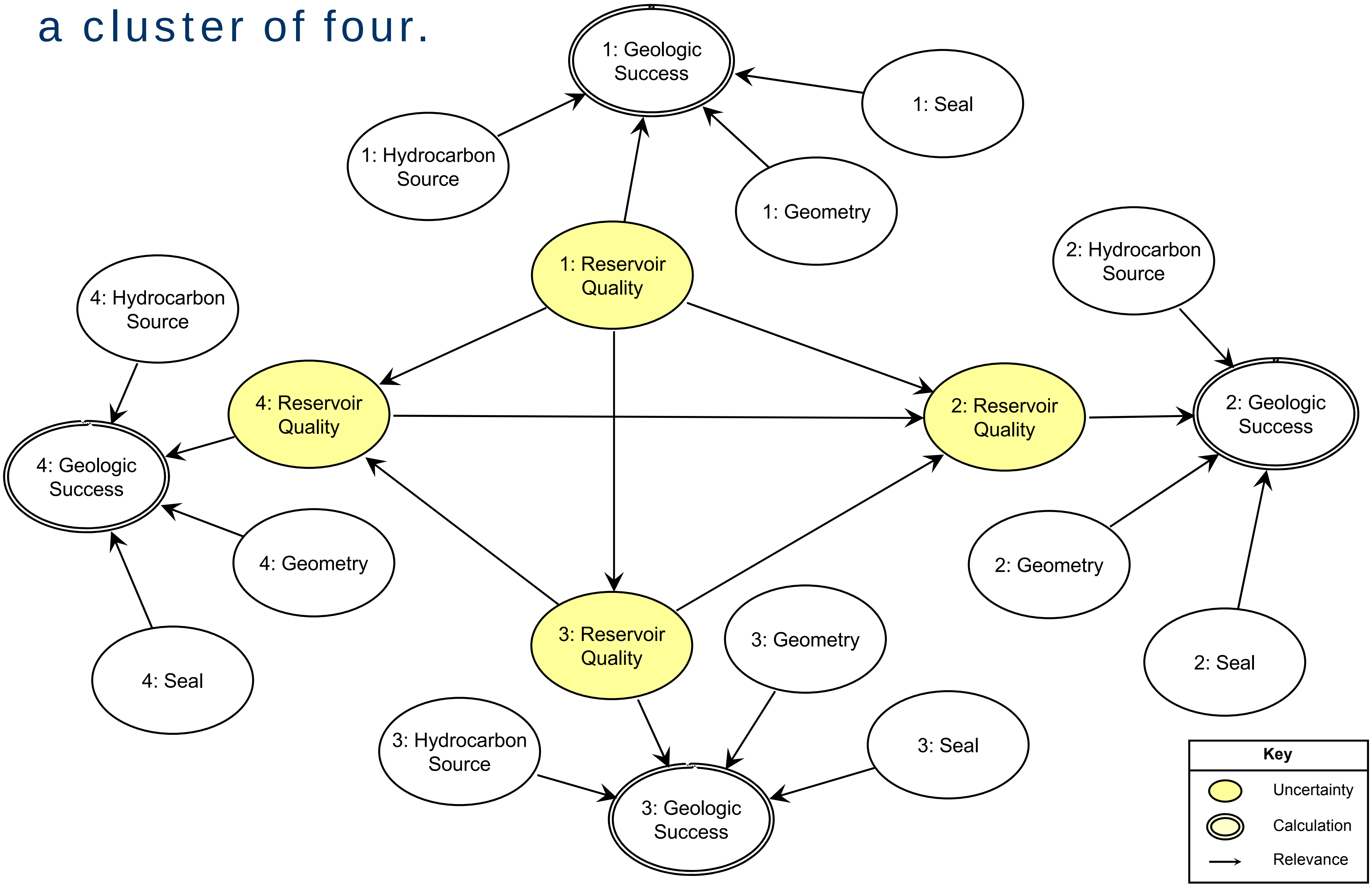
$$p(w_1) = p(h_1) * p(r_1) * p(g_1) * p(s_1)$$



- **Hydrocarbon Source**: Hydrocarbons in correct phase and quality were generated.
- **Reservoir Quality**: Reservoir rock of appropriate permeability and porosity is present.
- **Geometry**: Geometry of structure is as represented by seismic and slight changes would not jeopardize the accumulation of hydrocarbons.
- **Seal**: Seal exists with sufficient permeability to retain hydrocarbons.

The learning that takes place between prospects is in the context of these factors.

We have to perform these assessments for each component of geologic success—60 assessments for a cluster of four.



The client proposed a solution…

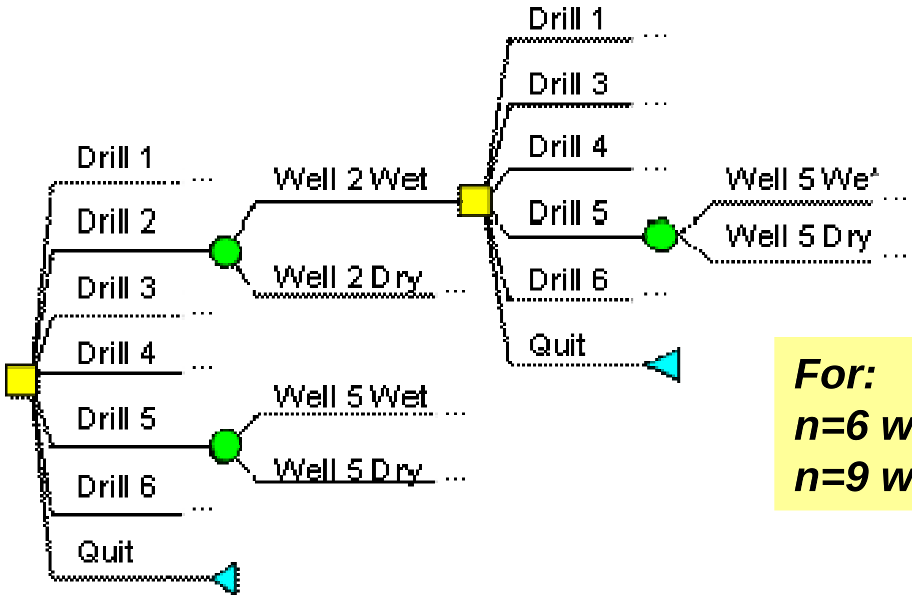
- Every time there is a success (failure) multiply the original prior by 1.05 (0.95).
- For example:

$$p(B_1|A_1\&) = 1.05\,p(B_1|\&)$$
$$p(B_1|A_2\&) = 0.95\,p(B_1|\&)$$

$$p(C_1|A_1B_1\&) = 1.05^2\,p(C_1|\&)$$
$$p(C_1|A_2B_1\&) = 0.95 \cdot 1.05\,p(C_1|\&)$$
$$p(C_1|A_1B_2\&) = 1.05 \cdot 0.95\,p(C_1|\&)$$
$$p(C_1|A_2B_2\&) = 0.95 \cdot 0.95\,p(C_1|\&)$$

- This procedure could yield probabilities greater than 1.
  - Therefore, the client suggested that probabilities be capped at 1.

The decision problem is equally complex.



For:
n=6 we have 113,959 scenarios.
n=9 we have approximately 460 million

$$\text{Scenarios} = \sum_{i=0}^{n} ({}_nC_i)\, i!\, 2^i (n-i+1)$$

- Stage *i=0,1,…,n,* where *n* is the number of wells that have already been drilled.
- In stage *i* there are ${}_nC_i$ possible combinations of previously drilled sites.
- *i!* different orderings of the previously drilled wells.
- 2^i different outcomes of the *i* drilled wells
- *n-i+1* possible actions (undrilled sites + option to quit)

Related Research

Modeling dependence in exploration problems

- Recognized as an important problem in Newendorp (1975)
- Keefer (2004) – Underlying Event Model, studies accuracy using simulation

Using entropy methods with dependence

- MacKenzie (1994), Lowell (1994) – Construct joint distributions with pairwise constraints
- Abbas (2003, 2006) – Constructs joint distributions with pairwise constraints; studies accuracy using simulation

Dynamic programming models of oil and gas exploration:

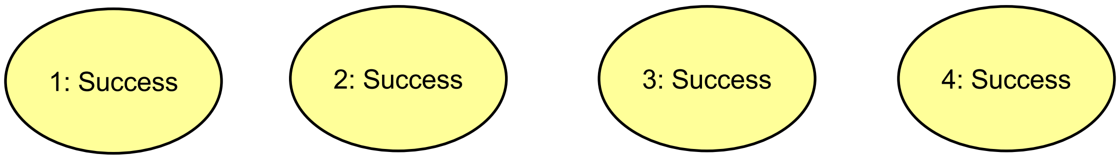
- Smith and McCardle (1998, 1999) – Real option/DA models focusing on price and production uncertainty

Agenda

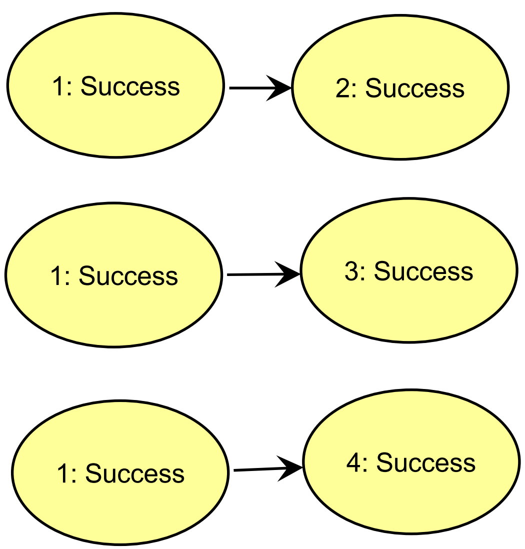
- The Challenge
- Determining the Joint Distribution
- Illustrative Example
- Conclusion (including life sciences example)

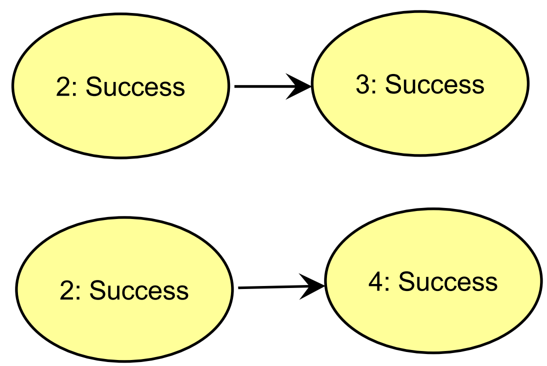
Instead of assessing the most difficult multi-conditioned uncertainties we focused on the marginals and each pair.

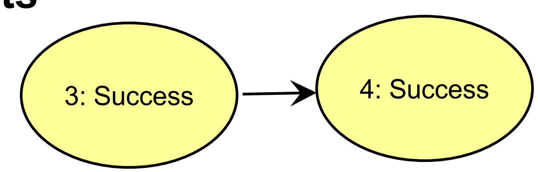
Marginal Assessments



Pairwise Conditional Assessments







Note: Pairwise conditional assessments must be consistent with the marginals.

If we could infer the full joint distribution from these assessments we could save five assessments per property (10 vs 15), or 20 per cluster. Most importantly, perhaps, we avoid the most difficult assessments.

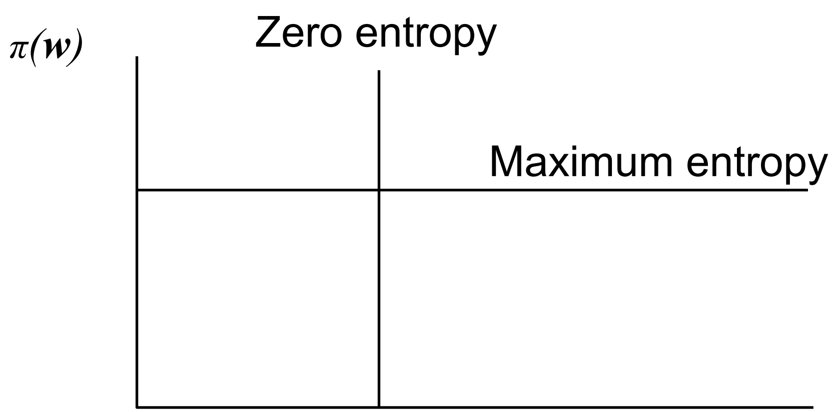
To what degree is the full joint distribution, constrained by these assessments?

The entropy of a probability distribution is a measure of the “amount” of uncertainty or its “flatness”.

Primer/Refresher on Information Entropy*

For a discrete distribution, information entropy is defined as:

$$H(\pi) = -\sum_{\mathbf{w}} \pi(\mathbf{w}) \ln(\pi(\mathbf{w}))$$

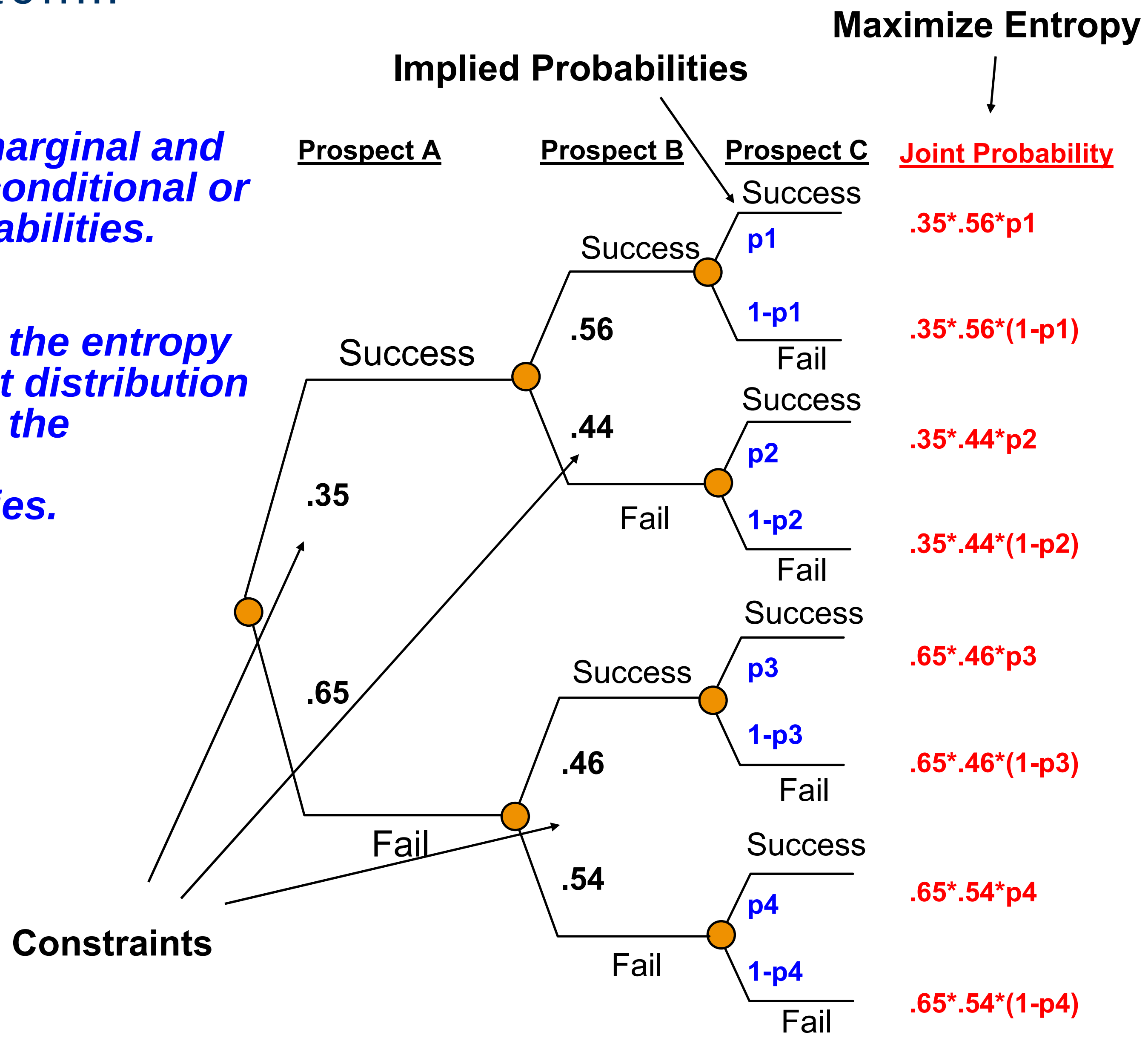


* For more see Shannon, C.E., 1948, “A Mathematical Theory of Communication,” *Bell System Technical Journal* (27), pp. 379-423.; Jaynes, E.T., Jaynes, “Information Theory and Statistical Mechanics,” *Physical Review*, pp. 108-171; Jaynes, E.T. *Probability Theory: The Logic of Science*, 2003.

Our approach…

1. *Assess marginal and pairwise conditional or joint probabilities.*

2. *Maximize the entropy of the joint distribution subject to the assessed probabilities.*



Agenda

- The Challenge
- Determining the Joint Distribution
- Illustrative Example
- Conclusion (including life sciences example)

An illustrative example involving six wells.

Well	Prob. of Success (p_i)	Expected Values: Given Success (s_i)	Expected Values: Given Failure (f_i)	Intrinsic Value
1	0.35	60	-35	-1.75
2	0.49	15	-20	-2.85
3	0.53	30	-35	-0.55
4	0.83	5	-40	-2.65
5	0.33	40	-20	-0.20
6	0.18	80	-20	-2.00

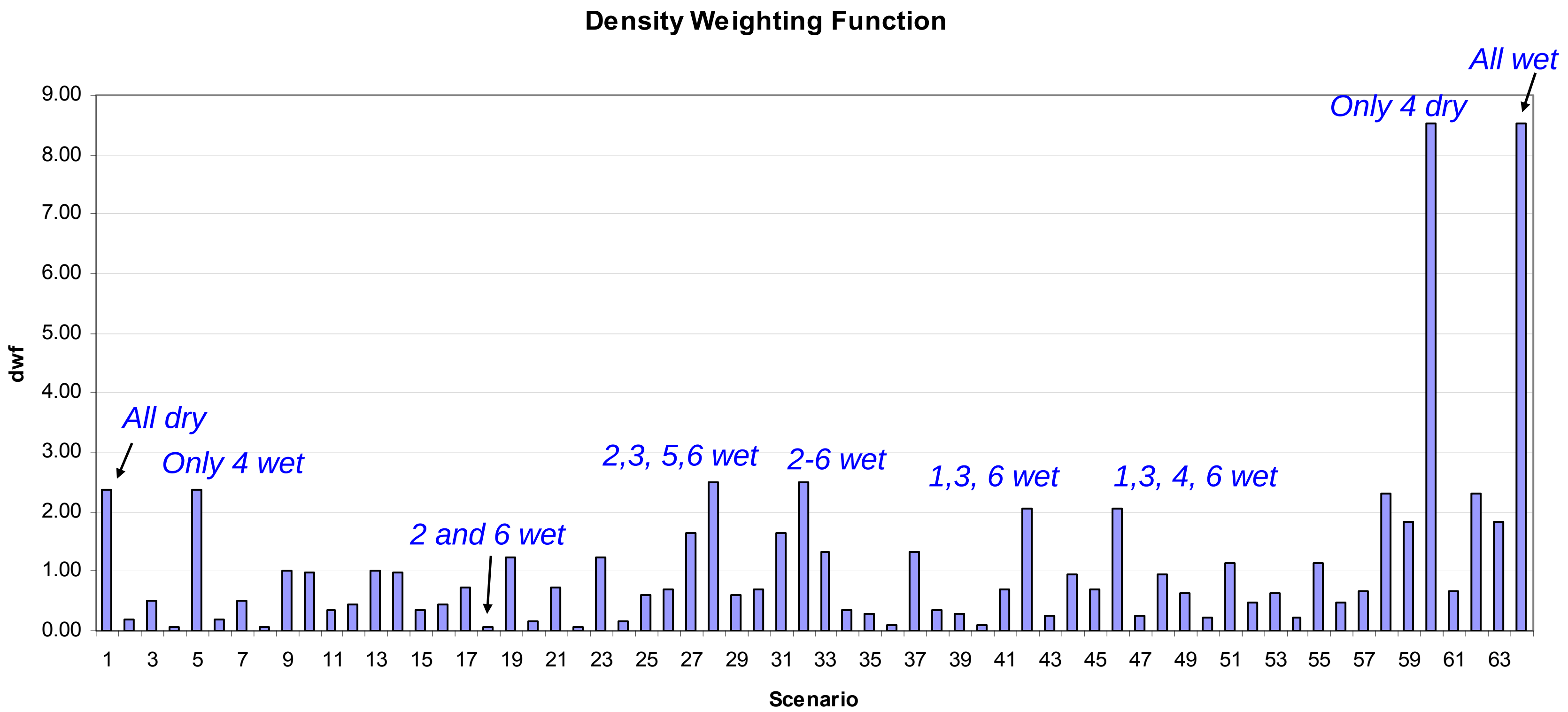
In the absence of learning, no wells should be drilled. But the client believed the wells were dependent: Can this be exploited?

i \ j	Direct Conditional Assessments P(j wet|i wet): 1	2	3	4	5	6	Marginal (p_i)
1		0.59	0.63	0.83	0.39	0.31	0.35
2			0.65	0.83	0.55	0.24	0.49
3				0.83	0.42	0.31	0.53
4					0.33	0.18	0.83
5						0.26	0.33
6							0.18

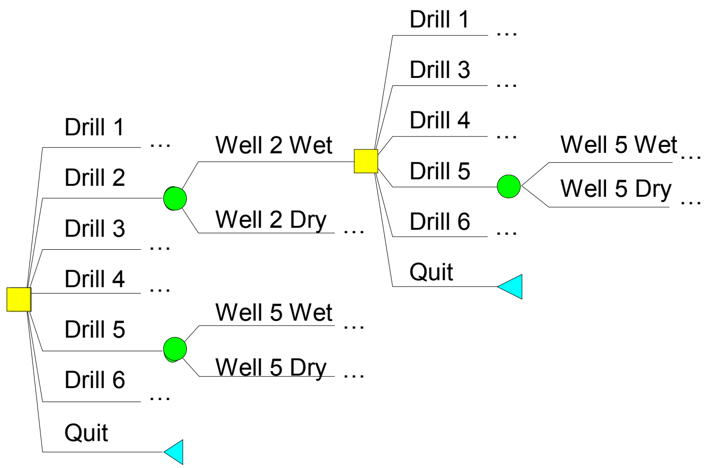
i \ j	Implied Correlation Matrix (ρ_{ij}): 1	2	3	4	5	6
1		0.147	0.147	0	0.094	0.248
2			0.236	0	0.459	0.153
3				0	0.203	0.359
4					0	0
5						0.146
6						

The optimal solution is similar in spirit to the client's suggestion; we just need to find the correct weights.

Joint Distribution = Independence Distribution (Product of Marginals) x Density Weighting Function



We can find the optimal drilling order by solving a dynamic program:



- This tree can be simplified by recognizing the recombining structure of the tree:
  - If we observe 1 to be dry and 3 wet, the future (conditional) probabilities and cash flows are the same regardless if we drilled well 1 or 3 first.
  - This a consequence of Bayesian updating: posterior distributions do not depend on the order in which information is received.

Number of Scenarios Considered

	Non-recombining Tree	Recombining Tree
• 6 wells:	113,959	729
• 9 wells:	≈460 million	≈ 79 thousand

Dynamic programming solves this decision problem by beginning at the end of the recombining decision tree and choosing optimal path as we work backwards.

For example, let's look at the case where we have drilled wells 1-4 and 6.

						For Well 5		
1	**2**	**3**	**4**	**5**	**6**	P(Wet)	P(Dry)	v5(ω)
1	1	1	1	-	0	0.58	0.42	14.70
1	1	1	1	-	1	0.65	0.35	18.81

\$14.70 = 0.58 * \$40 + 0.42 * (-\$20)
\$18.81 = 0.65 * \$40 + 0.42 * (-\$20)

We should drill well 5 if we end up in either of these situations.

Now, let's back up before we have drilled wells 5 and 6…

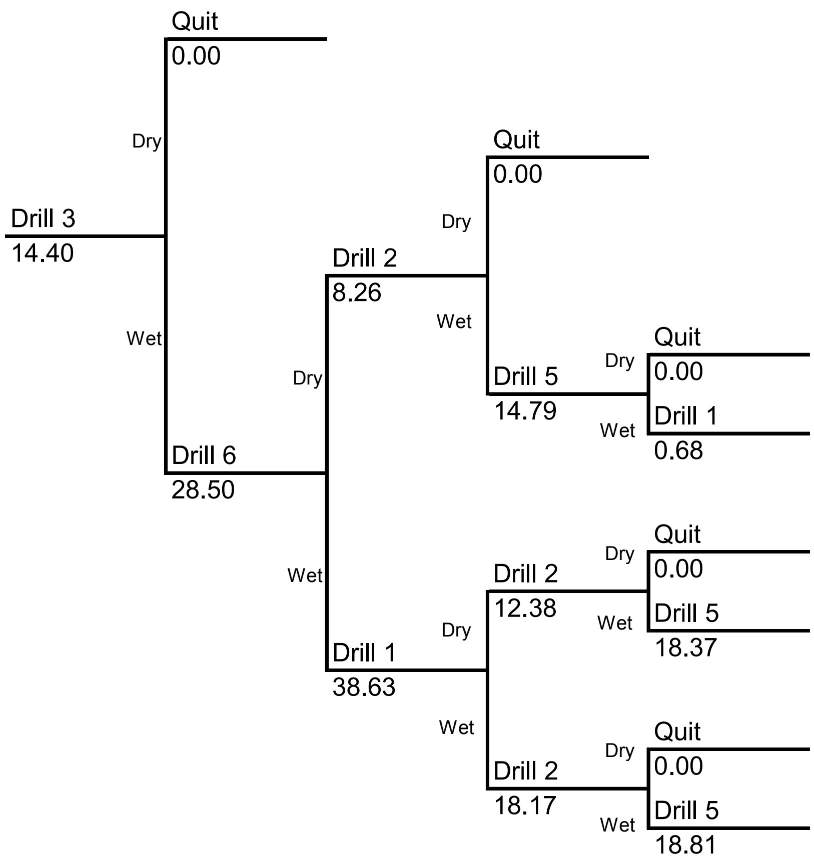
						For Well 6		
1	**2**	**3**	**4**	**5**	**6**	P(Wet)	P(Dry)	v6(ω)
1	1	1	1	-	-	0.48	0.52	44.40

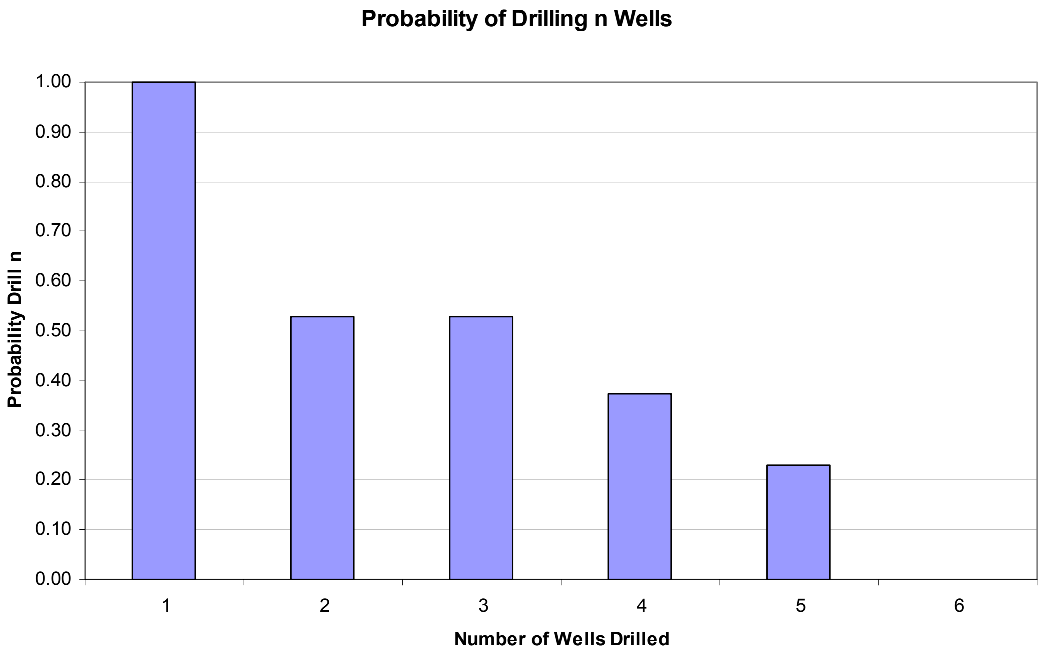
\$44.40 = 0.48 * (\$80 + **\$18.81**) + 0.52 * (-\$20 + **\$14.70**)

We would perform a similar analysis for well 5 and obtain a value of \$44.29; therefore, we should drill well 6.

We continue in this way until we have solved for the optimal action in each state.

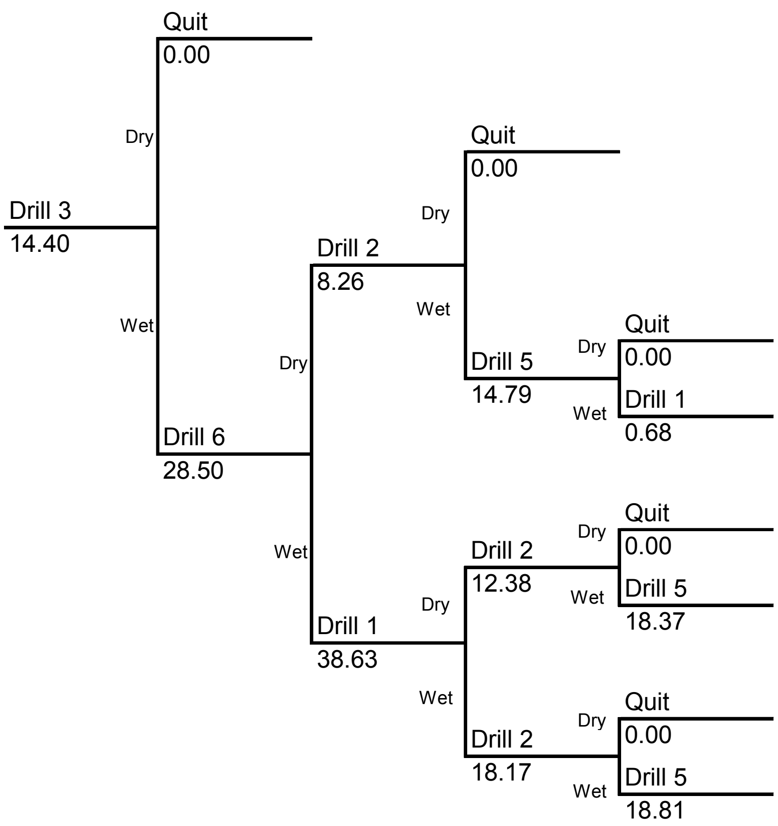
Example: The solution to the DP exploits dependence and leads to an expected value of $14.40 million.





Discount rate = 1% per drilling period ≈ 10% per year

The optimal policy confounds most simple "rules" for sequencing.



- The optimal solution does not:
  - Drill the well with highest probability of success first (4).
  - Drill the cheapest well first (2, 5, or 6).
  - Drill smallest expected loss first (4).
  - Drill the well with highest intrinsic value first (5).
  - Drill the potentially largest well (6) first.
- Impossible to solve using index policies:
  - A small increase in the value of success at well 5, makes it optimal to drill well **2** first, rather than 3.

		Expected Values		
Well	**Prob. of Success** (p_i)	**Given Success** (s_i)	**Given Failure** (f_i)	**Intrinsic Value**
1	0.35	60	-35	-1.75
2	0.49	15	-20	-2.85
3	0.53	30	-35	-0.55
4	0.83	5	-40	-2.65
5	0.33	40	-20	-0.20
6	0.18	80	-20	-2.00

Agenda

- The Challenge
- Determining the Joint Distribution
- Illustrative Example
- Conclusion (including life sciences example)

This methodology could be applied in any search/exploration setting.

P(compound will succeed in tumor type j | succeed in tumor type i)

i\i	1 Ovarian	2 Cervical	3 H&N	4 Breast	5 Prostate	6 NSCLC	*Marginal*
1 Ovarian		.18	.25	.50	.40	.25	*.08*
2 Cervical			.35	.35	.20	.30	*.12*
3 H&N				.30	.25	.35	*.25*
4 Breast					.40	.30	*.30*
5 Prostate						.30	*.20*
6 NSCLC							*.25*
Marginal	*.08*	*.12*	*.25*	*.30*	*.20*	*.25*	

Note: Illustrative example provided by Jack Klober (Director of Portfolio Management, J&J).

Conclusions:

- The proposed method provides a flexible and practical approach for formulating and solving sequential exploration problems.
  - Pairwise assessments seem minimal for this purpose
  - Easy to extend the DP to include synergies among prospects, e.g., cost savings resulting from shared infrastructure
  - Spreadsheet available for up to 6 wells
    - Takes 1-2 second to find optimal probability distribution using Solver
    - Solution of DP is essentially instantaneous
- Focus on oil and gas exploration, but similar problems arise in R&D as well
- Pieces of approach can be used independently or in other contexts:
  - DP model can be used with joint distributions constructed in some other way
  - Joint distributions could be used in simulation models or other decision models

Appendix: The Dynamic Programming Formulation

- Let $\boldsymbol{\omega} = (\omega_i, \ldots, \omega_n)$ denote the state where $\omega_i = \begin{cases} 1 \text{ if well } i \text{ is wet} \\ 0 \text{ if well } i \text{ is wet} \\ \text{"}-\text{" if well } i \text{ hasn't been drilled} \end{cases}$

- The initial state is $\boldsymbol{\omega} = (\text{-,-,-,-,-,-})$.

- DP Recursion:

  $v(\boldsymbol{\omega}) = 0$ if all wells have been drilled in state $\boldsymbol{\omega}$

  Otherwise:

P(well i is wet)

P(well i is dry)

$$v(\boldsymbol{\omega}) = \max_i \left[p(\omega_i = 1)(s_i + \delta v(\boldsymbol{\omega}_i^1)) + p(\omega_i = 0)(f_i + \delta v(\boldsymbol{\omega}_i^0)), 0 \right]$$

Value given state ω

Continuation value given well i is wet (dry).

where: s_i is expected future value of successful well,
f_i is expected future value of failed well, and
δ is a single period discount factor