Probabilistic Modeling for Decision Support in Integrated Operations

Martin Giese, University of Oslo; Reidar B. Bratvold, SPE, University of Stavanger

Abstract

We report work carried out within the CODIO project on COllaborative Decision support for Integrated Operations. As one part of this project, we have designed a system to provide assistance in operational decisions based on real-time sensor readings in a typical scenario: while drilling close to the transition to a high-pressure formation, a gas influx is observed. The drilling team needs to decide whether to circulate, increase the mud weight, plug back, set a casing, etc.

After a brief description of the technology we used to model decision problems (Bayesian Networks, Influence Diagrams), we describe our case study and go on to apply the method of decision analysis to it. We discuss how the resulting influence diagram was tested using a simulation, and discuss challenges in applying the technology, as well as lessons learnt underway. We also give a list of desiderata to establish better decision making practices in the petroleum industry.

Introduction

One of the (sometimes tacit) assumptions behind Integrated Operations (IO) is that all the information provided by the technology will help to make better decisions faster. But for this to happen, information has to be presented in an accessible, useful, and decision relevant format. The work reported here was carried out within the CODIO project on COllaborative Decision support for Integrated Operations. CODIO proposes to use the tools of decision analysis to provide decision makers with logical and consistent recommendations based on the available data.

In this paper, we show how we have applied the method of decision analysis to a real case, representative of a typical scenario in drilling operations, and how we have evaluated the model. As we will show, decision analysis provides a tool to arrive at good decisions in a consistent way. An effort is required to do this however, and we will point out both challenges in the modeling process and problems based in the way decision making is treated in the industry today.

Modeling and Solving Decision Problems with Influence Diagrams

Decision trees are widely used in the oil & gas industry. A decision tree provides a graphical illustration of all the uncertainties, decisions, and payoffs associated with a decision situation. Although the application of decision trees is very intuitive for simple decision problems, the method loses its appeal for more complex decisions. A symmetric decision tree grows exponentially with the number of variables in the decision domain, and this often makes their use impractical. For those more complex decisions, the influence diagram (ID) can provide a level of insight and transparency which cannot be achieved with the equivalent decision tree.

The use of graphical structures for representing joint probability distributions, usually credited to Wright (1921, 1934) and Good (1961a, b), was first fully described by Howard and Matheson (1981). It came to computer science and statistics through the work of Kim and Pearl (1983), Lauritzen and Spiegelthaler (1988), and Pearl (1988). There have since been many significant developments in model building and analysis using graphical structures.

The influence diagram[1] is a compact graphical model representation for reasoning under uncertainty and was developed to aid in the structuring and analysis of complex decision situations. The ID offers two key benefits to a decision situation – structuring and probabilistic evaluation. The structuring part provides a graphical representation of the interrelationships among the decision basis elements (alternatives, information, and preferences), whilst the evaluation part evaluates the resulting decision structure and provides an assessment of the expected utility given the quantified decision basis.

Influence Diagram Components

---

[1] Graphical structures without decisions and payoffs are often called Bayesian Networks or Belief Networks whilst Influence Diagrams, Decision Networks, and Bayesian Decision Networks refer to graphical structures which include decisions and payoffs.

The representational framework of an influence diagram captures the probabilistic relationships among decisions (alternatives), key uncertainties (information), and utilities (values or preferences). The influence diagram model comprises four types of nodes: decision, chance, deterministic, and value. Decision nodes, depicted as rectangles, specify a set of alternatives. Chance nodes, depicted as ovals, represent the key uncertainties in the decision domain. Deterministic nodes, shown as double-bordered ovals, represent either constants or intermediate calculations. Value nodes, shown as diamonds, represent the criteria used to make the decisions. The arrows in an influence diagram represent conditional probabilistic relationships between the variables they connect. Their interpretations depend on the types of nodes they connect (Howard 1990).

Decision Analysis Cycle
In this section we illustrate how decision analysis and influence diagrams can be used effectively to follow a systematic approach to analyze oil & gas decisions.

To deal with complex real-world problems, decision analysis uses a cyclical approach (Figure 1). The starting point is a real decision situation which is too opaque to be evaluated intuitively. The purpose of the process is to provide a set of steps that transforms the complex and opaque decision situation into a transparent one where we have clarity of insight and clarity of action. In the framing phase, we gather data, generate alternatives, and identify values for making the decision. The ID is then used to build the structure of the decision situation where the values, alternatives, and information and their relevance are clearly identified. This is the first high-level model of the decision. In the deterministic analysis phase, we use the ID to perform deterministic sensitivity analysis to identify the crucial uncertainties in the problem. In the probabilistic analysis phase, the probability distributions of the crucial uncertainties are encoded. Using the ID, the alternatives are then evaluated. Finally, in the sensitivity analysis phase, the results obtained from the ID are examined, the alternatives, values, and information are revisited, and the model is reexamined. Based on this sensitivity analysis, we decide whether to act or to return to the beginning and start the cycle again. Following this iterative process results in the selection of the best policies and provides clear reasoning for a decision.
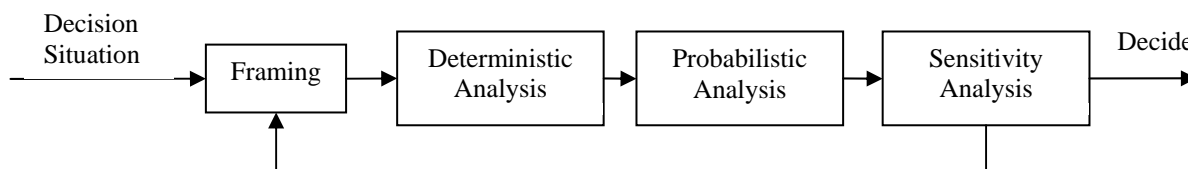

Figure 1 - Decision Analysis Cycle

Introductory Example
Consider the following decision situation (Rajaieyamchee and Bratvold 2009):

> An operator has a producing well with severe water cuts in the production that necessitate workover operations. He must decide either to replace the tubing and perforate (workover) or to drill a sidetrack parallel to the existing well path to maximize the oil production and minimize the water production. The cost estimates are known for each of them. But, he is uncertain about which producing zone (i.e., zone A or zone B) the water is coming from.

The decision problem can be represented by the influence diagram in Figure 2. It shows the probabilistic interrelationships between the deterministic nodes Perforate (A) and Perforate (B), the uncertainties Zone A, Zone B, Qwater, and Qoil, and the Sidetrack/Workover decision node. The arrows from Sidetrack/Workover into Perforate (A) and Perforate (B) deterministic nodes are functional dependence arrows, meaning that the decision made affects their outcomes. These deterministic nodes in turn alter the probability distributions of Qoil and Qwater depending upon which alternative is chosen. While the Qoil and Qwater are dependent on Zone A, Zone B, Perforate (A), and Perforate (B), they are independent of Sidetrack/Workover given Perforate (A) and Perforate (B). Finally, Net Present Value signified by NPV node deterministically depends on *Qoil*, and *Qwater* as well as the chosen action in *Sidetrack/Workover* decision.
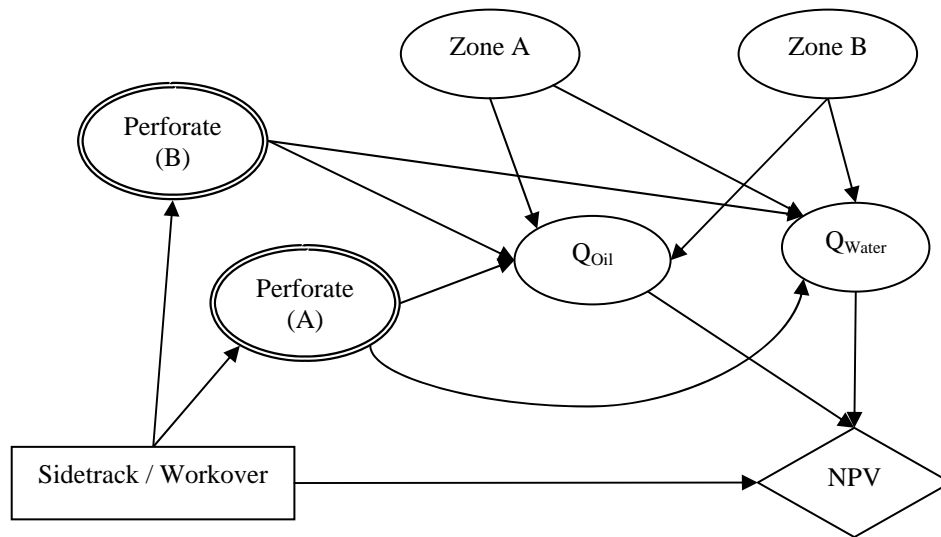
Figure 2 - An influence diagram for workover decision situation

Specifically, the arrows between chance nodes represent conditional probabilistic dependence. They are called relevance or influence arrows. The absence of a relevance arrow between two chance nodes is a much stronger statement of the expert's knowledge than the presence of an arrow. The former implies conditional independence, whereas the latter represents only potential conditional probabilistic dependence; i.e., the associated probability distribution of the child2 node may depend on the values of the parent node(s). For instance, consider the Zone A and Zone B nodes in Figure 2. The lack of a relevance arrow between them explicitly asserts that they are probabilistically independent. In contrast, the arrow between Zone A and Qoil implies that relevance may exist.

Relevance is a mutual property; i.e., if node Zone A is relevant to node Qoil, then Qoil is also relevant to Zone A, which is the result of arrow reversal by applying the Bayes' theorem. This allows switching the direction of relevance arrows between the two chance nodes. To accomplish this, extra arrows from Zone B, Perforate (A), and Perforate (B) into Zone A must be added to ensure that Zone A and Qoil are in the same state of information. This manipulation transforms the influence diagram in Figure 2 into the alternative influence diagram model shown in Figure 3. Although the influence diagrams are consistent, they would differ in suitability for probability assessments. For most geoscientists or petroleum engineers, it would be easier to assess the conditional probabilities of Qoil given the knowledge at Zone A than assess the conditional probabilities of Zone A given the knowledge of the probabilistic distribution of Qoil. Therefore, the ID in Figure 2 is more useful for our decision problem. This illustrates that specifying the direction of arrows should be based on their suitability for probability assessment.

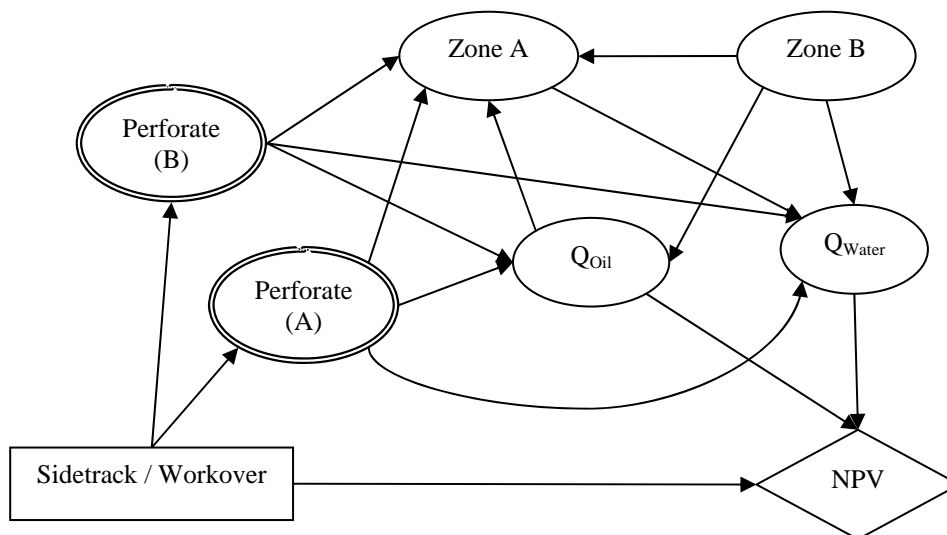**Figure 3 - Alternative influence diagram for the workover decision problem**

An ID can always be converted into a decision tree and in order to better appreciate the ID's compactness and transparency we can look at the equivalent decision tree for this problem as shown in Figure 4. In this case the resulting symmetric decision

---

[2] An arrow is said to originate at the *parent* and to terminate at the *child*.

tree has 128 end-points and by expanding the compact representation of the tree shown in Figure 4 into its full tree, we will lose its original intention of gaining insight and clarity into the decision problem
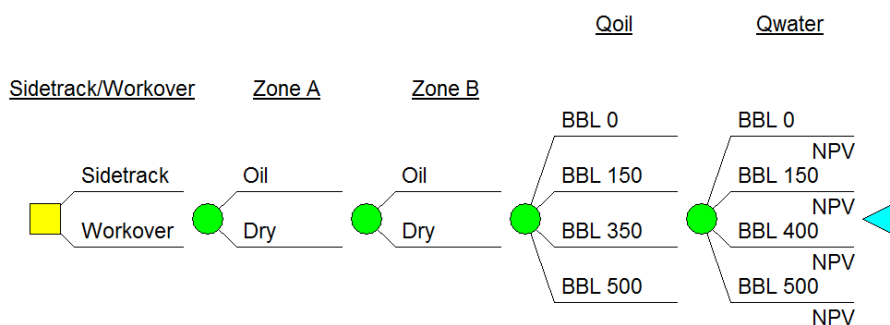

Figure 4 -Corresponding generic decision tree of the workover decision problem influence diagram model

It's been almost 30 years since IDs were introduced for supporting decision understanding and analysis. Although they have not been very popular in the oil & gas industry, they have been extensively used in other decision modeling situations including in the medical industry, Ecological and Biological studies, Aerospace, as well as in Marine Maintenance planning and risk-analysis, Transportation system, and in Software and Telecommunications.

The Case Study

In the CODIO project, we have applied a decision analytic approach to operational decisions during a drilling operation. An ID model was constructed using the information contained in a typical drilling program about the anticipated geology; like pore pressure, fracture pressure, rock hardness at different depths, etc., as well as knowledge about the equipment used; like drill bit, gas gauges, etc. This ID, which provides a visual representation of the structure and the context of the drilling operation, is then fed with real-time sensor readings concerning estimated depth, mud weight, mud flow, gas readings, and so on, to obtain probabilistic knowledge of important unknowns like depth of formation boundaries, actual down-hole pore pressure, etc. Using this probabilistic knowledge, one can estimate the utility[3] of different operational decisions, like continuing to drill, stopping to circulate, changing mud weight, setting a casing shallower than planned in the drilling program, etc.

Influence Diagrams have historically been applied mainly to decision problems that are concerned with a fixed, finite sequence of decisions known before hand, and where it makes sense to talk about a total utility of the outcome after all decisions have been made. We applied IDs to a slightly different kind of decision problem, namely one of operational decision making during a continuous process. Decisions about aborting a drilling operation, changing a parameter like mud weight, etc., can be made at any point in time, and it is not natural to identify a fixed finite number of decision points. However, as will be illustrated, the ID is a useful tool also in this setting.

The Scenario

We tested the decision analytic approach using an influence diagram model on a real case that occurred on the Norwegian continental shelf in 2002. The drilling program for this job prescribed drilling to a certain depth, which was not far from an anticipated high-pressure formation, before setting a casing. The depth of the transition to this particular high-pressure formation is highly uncertain due to vague seismic signals. Common practice is to set the casing at a safe distance from the high-pressure formation; several 100ft prior to the expected depth. However, this would have meant an earlier start of the open-hole section to be drilled next, and accordingly a lower pressure-tolerance of that open-hole section. Since drilling proceeds through a high-pressure zone, a new casing would eventually have to be set, which would again have to be somewhat higher than that in the drilling program. Eventually, starting with the usual casing depth, the reservoir could only have been reached with one more casing than required by the drilling plan in our case. In other words, setting the casing further down provided the possibility of reaching the reservoir depth with one casing less than usual.

The early part of the drilling operation was executed according to plan, but shortly before reaching the casing depth, a small amount of gas appeared and eventually dissipated. The drilling team did not assign much importance to this small gas peak, and continued to drill another stand. After that, a considerable gas influx was observed, which could not be contained by increasing the mud weight, and which eventually required plugging back and side-tracking, setting back the operation by several days.

---

[3] A *utility* is a numerical rating assigned to every possible outcome a decision maker may be faced with. For a risk-neutral decision maker who is using monetary value as the decision metric, the utility is equal to the monetary value.

Given the amount of gas observed after drilling the last stand, the drilling team had to follow the standard procedures for dealing with such a situation. The most interesting operational decision happened when the first, smaller gas peak appeared, and this decision is the focus of our discussion.

An unexpected gas influx can be due to a high-pressure zone occurring earlier than anticipated, requiring a casing setting. It could also be due to a small gas pocket that can easily be contained by increasing the mud weight, a much quicker and cheaper option. One could also stop and circulate in the hope that one is just seeing a small amount of connection gas[4] that will quickly dissipate. The question is how to decide without knowing which is the case, and how to account for the risks at hand; e.g., of fracturing if we increase the mud weight too much. Since the gas influx may indicate that the formation boundary is significantly higher than the depth indicated in the drilling program, the formation's fracture pressure might also be different from the values in the drilling program. The drilling operation has to deal with a number of uncertain events and it is impossible to say for sure what caused the gas influx, so the decision needs to be taken in the face of these uncertainties. This requires us to assess the probabilities of each of the causes, along with other relevant but uncertain factors, computing the expected utilities for each combination of these and each of the possible actions, and then taking the action that has the maximum expected utility (Bratvold and Begg 2010). This is the approach prescribed by decision analysis, and this is what IDs allow us to compute.

We will discuss the decision-analytical approach to this kind of operational problem, and show how IDs can help to structure the decision situation, provide insight and recommend a choice. We will *not* discuss what would have been the best course of action in this scenario given knowledge available only in hindsight as this is somewhat "unfair" and invites hindsight bias (Bratvold and Begg 2010). Rather, we will discuss how to achieve clarity of thinking and, eventually, how to reach the optimal decision during the operation given the information available at that time.

Available Data and Decision to take
We were able to use all data included in the drilling program, including geological data available about the field before the particular drilling operation. We also had access to detailed log data from the operation, which we could use to replay the scenario, providing our ID model with the data that would have been available if the ID model had been available during the actual operation.

We can distinguish between two types of information that are relevant for this decision problem. One kind, which we will call *configuration data*[5], represents estimates and expectations about the operation before it starts, and is taken from the drilling program, cost models, risk assessments, etc. The other kind, which we will call *real-time data,* is obtained during the operation and consists of sensor data and any other relevant observations.

The configuration data relevant for our case study included estimates of

- depths of geological boundaries, in particular the transition to the considered high-pressure formation

- pore pressure, fracture pressure, and rock hardness (compressive strength) at various depths,

- the annulus volume up to the depth where our investigation starts, allowing to estimate the lag time of gas carried up in the drilling fluid,

- the influence of the mud flow rate on down-hole equivalent circulating density,

- the relation between rock hardness, weight on bit, RPM, and rate of penetration,

- the probability of requiring an extra casing dependent on the casing depth,

- all relevant costs and benefits connected to various operations under various circumstances, including cost of repairing damage made, cost incurred by reduced production of the finished well, etc.

A complete description of the decision problem at hand has to contain not only the expected values for this configuration data, but also a quantification of the uncertainty attached to the numbers. It is for instance not enough to know that geologists anticipated the transition to a high-pressure zone to start at a depth of 5800ft TVD. One also needs a quantification that says whether this means between 5799ft and 5801ft, or between 5700ft and 5900ft. To be more precise, a probability distribution needs to be given that quantifies the possibility that the zone starts at 5700ft, 5800ft, 5900ft, etc. Such a quantified assessment of the uncertainty can be specified by 10/90-quantiles, a standard deviation, or in a number of other ways, but it is absolutely required. While these uncertainties are not usually stated along with the base case values of the cited configuration data, and experts often are struggling to put a number on them when asked, decision analysis commonly applies elicitation methods (Bratvold and Begg 2010) which help the experts assessing these probabilities. In decision analysis we are only concerned

---

[4] When one stand has been drilled, a connection needs to be made, i.e. the block is removed from the drilling pipe while another stand is attached to it. While making the connection, mud pumps have to be turned off for some minutes, which leads to a reduced downhole pressure. This can in turn cause a gas influx that usually disappears when the pumps are turned back on. Such a gas influx is referred to as *connection gas*.

[5] In the context of Bayesian inference, which is the mathematical algorithm embedded in IDs, the configuration data is called *the prior*.

with subjective probabilities whereby probabilities reflect a person's or a team's knowledge about the event in questions. Probability is a state of knowledge, and not a state of things, and there is no "true" or "correct" probability. The only requirement is that the assigned probability truly reflects the expert's knowledge (Bratvold and Begg 2010). Furthermore, sensitivity analysis is an essential element of any decision modeling activity and will provide valuable insights as to whether the recommended decision is robust (it will stick) given reasonable variations in the input parameters including the probabilities.

In this case real-time data available in the form of log data included

- measured depth and true vertical depth

- RPM, torque, weight on bit, hook load, rate of penetration (ROP)

- mud flow, stand pipe pressure, trip tank and active system volumes

- gas in mud (top side)

RPM, weight on bit, and ROP can be used to obtain a crude estimate of the solidity of the rock being drilled. For instance, a drop in ROP while RPM and weight on bit remain unchanged can indicate that drilling has reached a harder formation.

The decision to be taken at any point in time during a drilling operation is what to do next. While there are many possibilities, we singled out the most relevant actions in the case study:

A1: Continue drilling

A2: Stop and circulate for a set amount of time

A3: Increase mud weight and circulate for a set amount of time

A4: Set a casing now, possibly plugging back first

Even though the details of operational decisions vary with developing technology, such as extensive down-hole measurements and wired pipe, our case study can be seen as representative for the decisions to be made during a drilling operation: they are real-time decisions, based on real-time, continuous sensor data, and they influence a continuous process. Some of the data involved, like rate of penetration or weight on bit, provides fairly immediate information of the down-hole reality, whereas others, like gas or cuttings observed on the platform reflect down-hole happenings with a delay dependent on mud circulation.

Applying Decision Analysis to the Case Study

We will now illustrate how decision analysis can be applied to an operational decision like that of our case study. We can begin with the simplest of influence diagrams describing a single decision to be taken dependent on some observations, given in Figure 5.
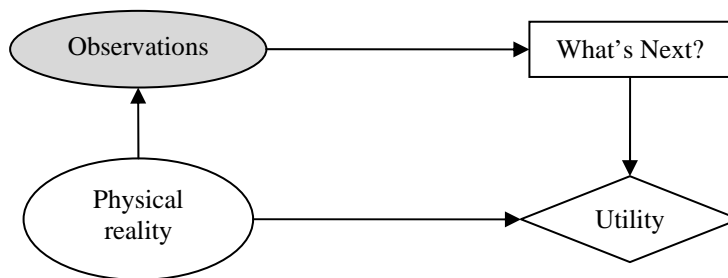

Figure 5 - The simplest of influence diagrams

We will use gray nodes are the ones corresponding to *observable* variables; i.e., those variables whose values are known from sensor readings and other observations. Conversely, we will refer to non-observable variables as being *hidden*. The hidden variable "Physical Reality," represents unknown events like the actual geology, pore pressure, the actual position of the bit, etc. The decision node "What's Next?" represents the decision among the various operational alternatives. Finally, the "Utility" node represents the stake holder's utility, for instance the net present value that results from the whole operation.

The diagram expresses that we only have the observations as input to make our decision, and not the actual physical realities. However, the latter have a connection to our observations. The utility is dependent on our actions and the physical reality.

This generic diagram does not yet help very much in making the actual decision, since it lacks fine structure in both the observations and the physical realities, and the connections between them and the utilities. We can analyze the decision problem further by first decomposing the nodes representing observations and those representing physical reality.

Since the presence of gas in the well and its provenance is a central aspect of the case, we start by observing the following causal chain:

- At any depth, there is some pore pressure, which depends on the depth as well as the geology surrounding the well.

- At any moment, we have a measurement of the depth that includes some error.

- At any moment, there is some equivalent circulating density (ECD) down-hole, which can be estimated using mud weight and flow.

- If the pore pressure is larger than the ECD, gas will flow into the hole.

- Such gas will be circulated out and measured with a delay given by the lag time, which can be estimated using the mud flow and the annulus geometry.

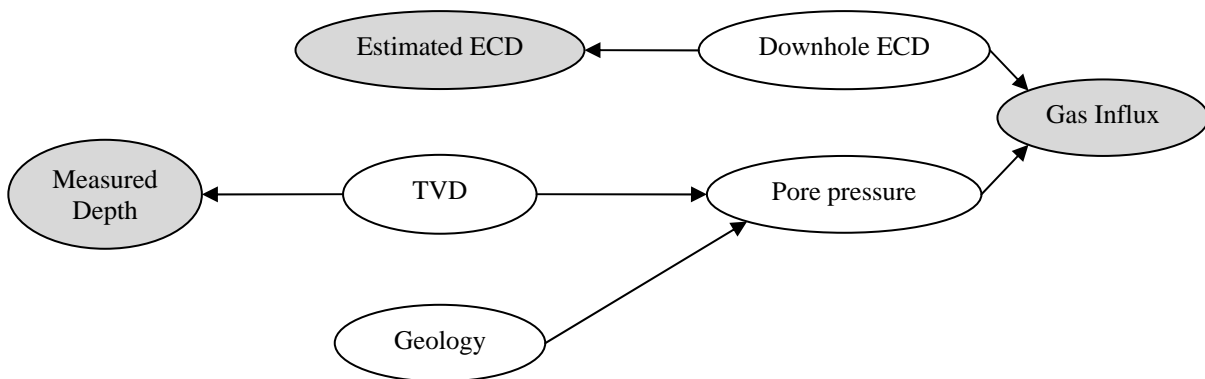This connection can be illustrated by the influence diagram in Figure 6.

Figure 6 - An ID for a gas influx

The "physical reality" node has been decomposed into nodes for the TVD of the bit, the geology surrounding the wellbore, and the pore pressure at the bottom of the wellbore. Three concrete observable variables represent the measured bit depth, measured gas influx, and the estimated downhole ECD. The diagram shows the relationship between these values. In order to do calculations, the uncertainties need to be quantified. In particular, the probability of various geologies needs to be given, as well as the probability of observing a given pore pressure at a given depth for a given geology, etc. Having assessed all the probabilities, as well as an actual measured depth and estimated ECD, and the fact that gas is observed, the backward reasoning typical of IDs allow us to infer that the pore pressure must have been higher than the ECD, and hence the geology might be different than assumed in the drilling plan.

Providing a complete quantification of "the geology" is difficult as it is hard to assess all the probabilities of different pore pressures at different TVDs for the different geologies. Moreover, it is only the local geological situation near the current drilling depth that is of importance for the gas influx. Upon conferring with a petroleum geologist, we decided to simplify our ID using additional assumptions. The main uncertainty in this context is the depth of the high pressure formation transition. The uncertainty range of this depth is a few hundred feet. However, if we knew the distance from the transition; i.e., the difference between TVD and the transition depth, we could give a much more precise estimate of the pore pressure. Furthermore, for the purpose of this scenario, we chose to neglect other aspects of the geology. This gives us the ID in Figure 7.
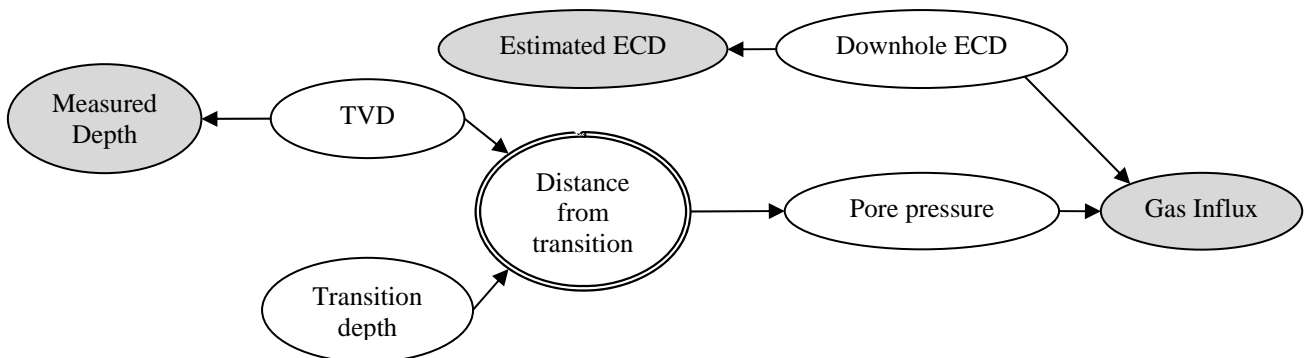

Figure 7 - Factoring out the distance from a formation transition

Note that the "Distance from transition" node is what is known as a deterministic node, since it merely computes the difference of two other variables with no uncertainty added. We mark deterministic nodes with a double outline. The estimated pore pressure at a given distance from a formation transition is now comparatively easy to obtain from geologists and the information in drilling programs. We have "factored" the uncertainty of the pore pressure into two parts: the distance from the transition depth and the remaining uncertainty given that distance. This is a transformation which we applied in the light of stated simplifying assumptions that can be discussed with domain experts.

Next, we can add to our ID that a gas influx can also be caused by a gas pocket, and assess the probability of a gas pocket at any given depth. We won't go into the details, but the principle is the same: the causal chain is modeled, and the ID's Bayesian inference algorithm infers what the updated probabilities of the configuration data must be given the measured real time data.

Having taken gas measurements into account, we can move on to model rock solidity. Decomposing the uncertainty of rock solidity, we can now say that

- The rock at any distance from the formation boundary has some compressive strength
- The compressive strength, the weight on bit, RPM, and the state of the bit together determine the rate of penetration.
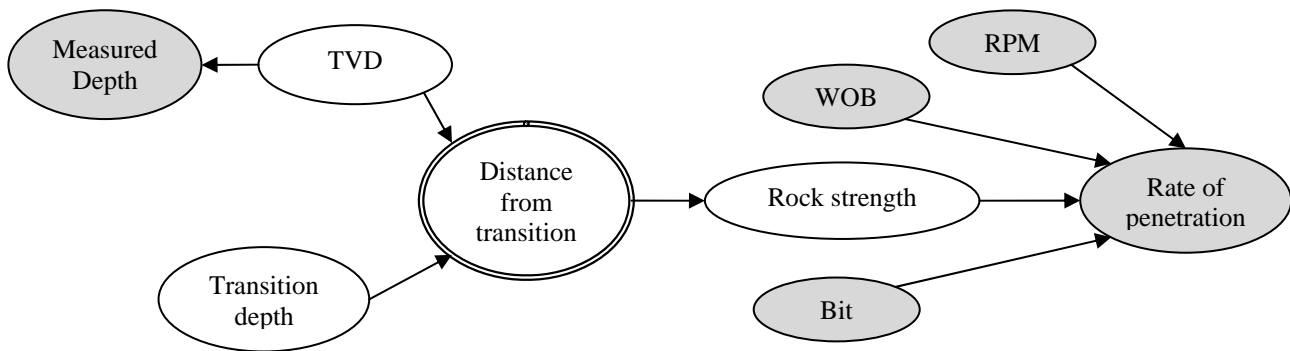
This is expressed by the diagram in Figure 8:

Figure 8 - An ID for rock strength

This diagram is now merged with the one for the gas influx to give the ID in Figure 9.
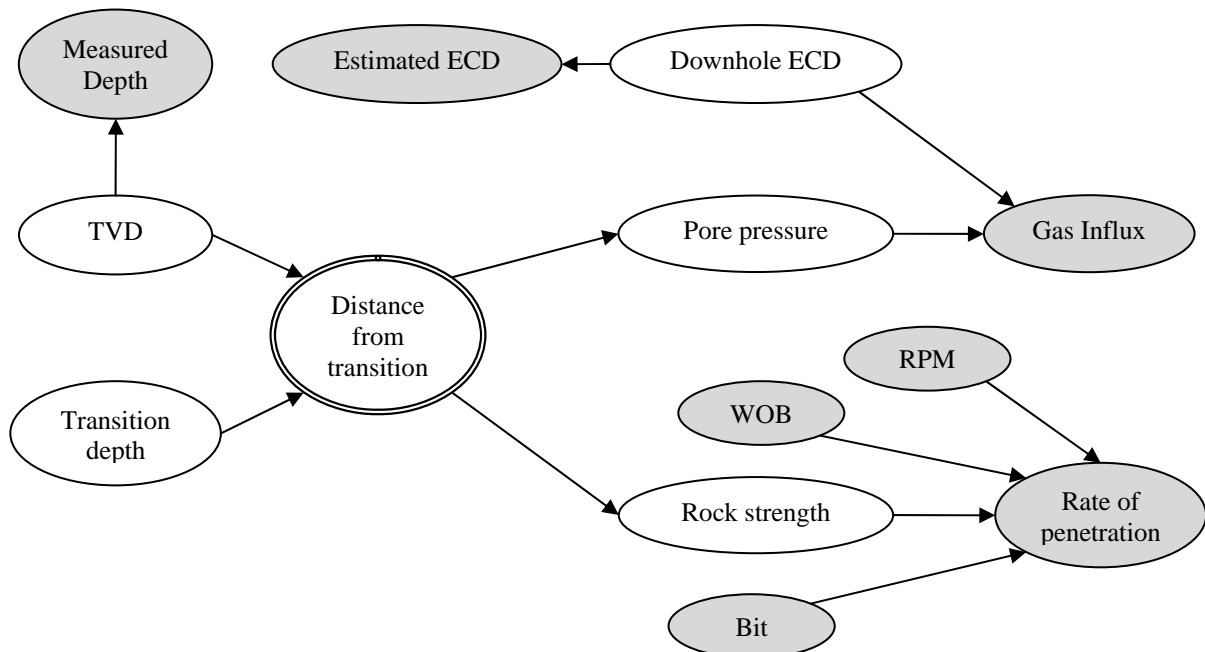

Figure 9 - A combined ID for gas influx and rock strength

With this configuration, information pertaining to rock hardness and to gas can be combined to give an improved assessment of the geological uncertainties. The same method can be used to include further observations like cuttings or downhole sensor readings. The principle remains the same. By proceeding in this way, we ultimately obtain a fine grained representation of the relevant uncertainties, the observable variables, and the connections between them. Moreover the structure is such that it is easily possible to quantify the relations between the variables for a particular drilling operation, enabling calculations. In any decision modeling situation, there is always the question: What is the optimal level of detail in the event definition? The answer is always: the level at which the experts finds it most meaningful to quantify the uncertainties. In a case like this where there is measured real time data, a corollary answer is the level at which the measured data is most relevant for the defined event(s).

We can join the diagram of Figure 9, which structures the relationship between the physical reality and our observables, with the initial ID of Figure 5 that represents the operational decision of what to do next, to obtain the ID in Figure 10.
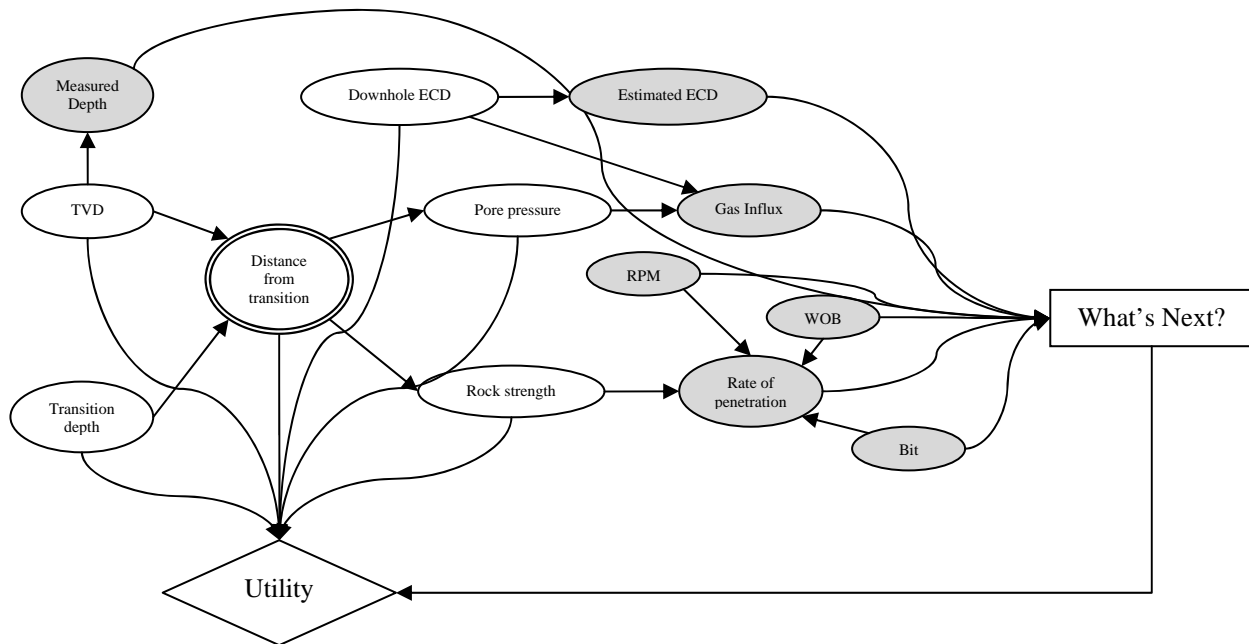

Figure 10 - An ID for gas influx and rock strength including utility and decision nodes

We see that all observable variables have an arrow into the decision node "What's Next", representing the fact that their values are all available when the decision is to be made. All the hidden variables we obtained by decomposing the "Physical Reality" node can influence the utility.

As before, we can now observe that the way the various hidden variables influence the utility is too coarse grained to make the quantification easy or natural. We can gather the various components influencing the utility measure the decision maker(s) has decided upon; e.g., net present value. Some values of interest include:

- The cost of setting one more casing than the drilling program calls for

- The net present value of the finished productive well. This may be reduced; e.g., by reduced and delayed production

  due to reduced pipe diameter if an extra casing is needed

- The cost of waiting and circulating for a given amount of time

- The cost of fracturing the well or the casing structure through too high mud weight

- The cost of plugging back the well in order to set a casing

These are again uncertain values, and domain experts need to quantify these uncertainties by giving the probability that any of these costs or values takes on a given value. We can split the utility node into several new utility nodes representing the different utility components. These can then be connected to the decision node and the relevant hidden variables. In some cases the connection between the hidden variables and the utility may be more complex and new variables may be needed. For instance, the cost of potentially fracturing the formation by increasing the mud weight is dependent on the fracture pressure, which is dependent on the depth in a similar way as the pore pressure, as shown in Figure 11.
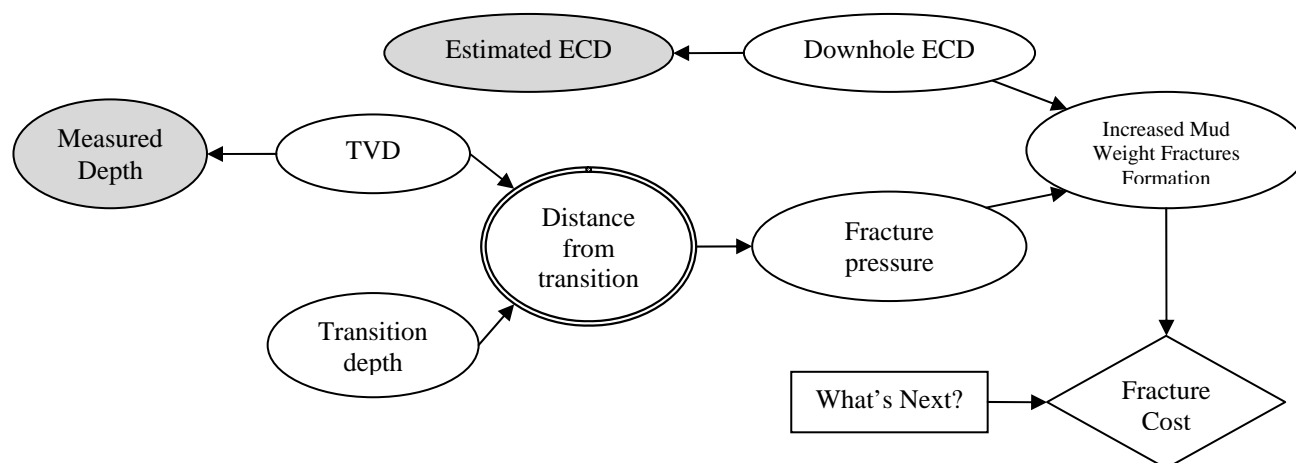
Figure 11 - Structuring the cost of fracturing

In addition to modeling the fracture pressure, we have added a hidden variable to express the fact that the increase in mud weight from action A3, "Increase mud weight and circulate for a set amount of time" would make the downhole pressure exceed the formation's tolerance. The utility node is setup such that if this is the case and action A3 is taken, then the cost for fracturing the well is incurred. In a similar way, we can model the relevant factors and cause-effect chains for the other utility components.

In the CODIO project, we performed this analysis for the complete scenario of our case study. We encountered a number of additional challenges, some of which will be discussed in the remainder of this paper. The intention of this section was to illustrate how the decision analysis process provides a structure, transparency, and the necessary understanding to make good decisions. This includes clearly defining the relevant observable and hidden variables involved, their interdependencies, the utilities to maximize, and how they interact with the decision to be made. Given quantitative assessments of the interdependencies between the variables, we can use the resulting model for computation and thereby deduce which action leads to the highest expected utility.

We should point out that the structure of the ID we constructed is generic: the same structure can be used whenever operational decisions are to be taken when drilling close to a high pressure formation. For a different drilling operation, the structure of the ID can be kept, only the numbers in the probability tables need to be adjusted, using the relevant numbers for the drilling program, including geology data, risk assessment, etc.

Testing the Influence Diagram

The ID constructed for the CODIO case study was implemented in a commercial software application and tested in a simulation environment.

The ID consisted of a fixed structure, constructed as described in the previous section, with most of the conditional probability tables filled in by a software program that converted geological data taken from the drilling program into appropriate tables. Using the ID for a different operation would have required regenerating those tables from the relevant configuration data. Some of the utilities, expressing the cost of various actions under different circumstances, would also have to be adapted to fit the operation at hand.

The system used for testing, called CODIO Pilot 1, uses the Hugin tool (Madsen et al., 2005) to load the ID and perform computations on it. An infrastructure was created to read real time log data from a WITSML server several times per minute. After some preprocessing of the data, for unit conversion and such, Hugin is used to derive the expected utility of the available operational actions. The pilot provides a Web interface that displays a ranking of the possible actions, where the action with the highest expected utility is the one the system recommends.

To test the pilot system, we used the ELAD drilling simulator (Cayeux, 2007, Hulsund et al., 2009) at IRIS in Stavanger. The simulator was configured to provide a similar environment as in the original drilling operation defined by the case study. That included providing it with pore pressures, rock properties, mud flow, etc. Obviously, this data was *not* taken directly from the drilling program, since the actual geology encountered in the case study was different from that predicted by the drilling program. Rather, we configured the simulator with data that corresponds to what is known about the well *in hindsight*. This made our experiment realistic in the sense that the CODIO decision support tool only had information available before the operation, whereas the simulation tool behaved like the actual well did.

The operation was simulated with drilling engineers controlling the system. ELAD published simulated log data to a WITSML server, which the CODIO pilot was connected to. The drilling personnel were shown the interface of the CODIO pilot with its operational recommendations.

The system showed sensible behavior in the test: as soon as the gas levels rise over a certain amount, the system advises to pause drilling and to increase the mud weight. It continues to do so until the expected cost of fracturing the formation or the casing structure is asserted to outweigh the potential benefit of containing the gas influx or at least being able to set a casing without plugging back. Given the configuration in the case study, the system then decides that the down-hole pressure is already too high to be contained by the drilling fluid, and advises to set a casing immediately, i.e. earlier than envisaged by the drilling program.

We also tested the model on data corresponding to other circumstances, e.g. gas from a gas pocket, or no gas at all and the transition at the depth where it was expected. The influence diagram gave sensible recommendations also for these cases.

While this overall behavior is to be expected from a system that takes all the major aspects of the problem into account, the subtlety lies in the thresholds: at which distance from an expected formation boundary is a gas pocket more likely than that the formation boundary is already reached? When does the risk of destroying the casing structure outweigh the benefit of controlling the gas influx by mud weight? A system based on decision analysis computes these thresholds based on a systematic assessment of the structure and magnitude of the uncertainties involved.

Challenges

Although our approach turned out to be feasible, and led to a successful test, there were a number of practical challenges to overcome in the construction of the influence diagram. We will discuss some of them in this section.

One challenge comes from the multitude of parameters in formulae provided by specialists in petroleum engineering. For instance, we initially tried to quantify the relationship between the pressure difference down-hole and the amount of gas measured. That would have made it possible to use the amount of gas observed to infer an estimate of the current pore pressure, which in turn would lead to an improved estimate of the depth of the transition boundary. Formulae that quantify a gas influx have indeed been published (e.g. Dake, 1978) but they require parameters describing rock porosity, gas reservoir size, and shape. While an up-front estimate of porosity might be available, it is practically impossible to obtain usable estimates of the geometry of the gas enclosed in a formation, making it impossible to use such formulae for our purposes. The matter is further complicated by the difficulty of describing the way gas percolates to the surface while being transported by the mud in a precise way. The consequence of this, as practitioners confirm, is that it is very hard to say anything about the pressure difference down-hole from the amount of gas measured. The best we can do is to say that if there is gas, then the pore pressure must have been higher than the ECD.

We made a similar observation with respect to the computation of ECD: there is extensive engineering knowledge about the behavior of drilling fluids in the annulus, and these could be used to obtain a precise estimate of the down-hole ECD given mud flow, viscosity, annulus geometry, and several other parameters. However, these parameters were hard to elicit for the drilling operation in our case study. On the other hand, we asked practitioners at the oil company about this, and they were easily able to give an empirical quantification of the impact of mud flow on down-hole ECD when switching off pumps to make a connection for the operation we were considering (ca. 0,5ppg). There was surely a higher uncertainty connected to this estimate than to one obtained with a correctly parameterized hydrodynamic model. But the inaccuracy of this practitioners' estimate was still small compared to the other uncertainties described by the ID.

So in general, a practitioner's rule of thumb was more useful for this particular kind of probabilistic model than a precise formula. This could convey the impression that the recommendations derived from the influence diagram cannot be better than the intuitive recommendations of an expert, but that is not the case. Any rules of thumb used in the ID are explicitly expressed, i.e. they are part of the explicit assumptions made when modeling. The recommendations derived from the diagram are optimal given the assumptions and probabilistic input. A human decision maker can usually not name all the assumptions made to arrive at a decision, and the decision is often not optimal with respect to the knowledge and assumptions it is based on. So the formalization has its value even if it is not built on a precise theory. It is however very beneficial if the consulted expert is trained in the principles and methods of decision analysis, and therefore has an understanding for the necessity of quantifying uncertainties.

Another difficulty arises with some of the utilities important in operational decisions. For instance, setting an early casing can lead to the need for one more casing than originally planned, and therefore a smaller diameter of the pipe in the finished well. This leads not only to reduced total production, but to *delayed* production, since the petroleum will flow more slowly. A delayed production in turn reduces the net present value of the well; the sooner profit is made, the better. But putting a precise number on the reduction in net present value is very difficult. The solution here is again to ask a practitioner for help. Since the loss in net present value is an aspect that has to be considered every time this kind of derivation from the drilling program is considered, decision makers are already, possibly subconsciously, weighing it against other costs. Making these trade-offs explicit helps in providing more insight as well as to making the decision more transparent and consistent with the stated preferences (utilities).

A more technical kind of challenge comes from the temporal aspects of our case study. To begin with, the sensor values for RPM, ROP, and weight on bit give immediate information of the happenings down-hole, while gas readings have a lag time

attached to them. This means that sensor data needs to be stored in a data buffer and provided to the ID together with data that stems from effects happening down-hole at the same point in time.

A more subtle time-related aspect comes from the fact that drilling is interrupted to make new connections. This means that no assessment of rock solidity is immediately available when a connection is made until drilling is resumed. We solve this by remembering solidity-related data from the last time drilling did progress, and feeding that to the ID.

Finally, if a gas peak is encountered, e.g. from connection gas, and the gas eventually fades, the ID still needs to take this into account for further action. Essentially, connection gas indicates that the pore pressure is only barely contained by the circulating density, which is important knowledge for operational decisions. We solve this by remembering data from the last time a significant amount of gas was observed, and also providing that as input to the ID.

Taking the lag time effect into account, every computation with our ID uses sensor data from up to 5 different points in time. The control software that feeds the log data collected from the WITSML server into the ID managed by the Hugin tool takes care of the management of this data, as well as the computation of the times for which data is required.

Lessons Learnt

Our endeavor shows that it is possible to apply systematic decision analysis to operational decisions in oil and gas drilling operations. We do not pretend that it is trivial to do so: it requires both skills in decision modeling and the availability of domain experts, preferable also having a basic knowledge of decision theory and probabilistic assessments. In the previous section, we have discussed some of the difficulties we encountered, and how we solved them. In this section, we want to make some general remarks about lessons learnt while carrying out our case study, which we feel will be valuable for future endeavors of this kind.

First, given a theoretical model for some physical phenomenon, one has to check whether it is possible to obtain a-priori estimates for the parameters it contains. It is often easiest and entirely good enough for most purposes to obtain a rule-of-thumb estimate from someone with practical experience.

Second, it is important to get full information about the scenario as early as possible. This includes all relevant geological data in numerical, machine processable form. In case the decision support tool is to be tested in a simulation environment, one needs to remember that logs produced by a simulator will not entirely match those from an actual drilling rig. Log data from the simulator should be obtained as early as possible in order to be able to take those differences into account when configuring the ID.

Finally, during the decomposition of the utility function, it turned out that a number of intermediate nodes were required to express aspects that concerned the future development of the operation. E.g. one node was required to express whether the pressures involved are such that it is at all possible to contain the gas influx using mud weight. Another node expressed that the situation was such that a premature casing would eventually have to be set. The necessity for such nodes is a direct consequence of the fact that we are using an ID with only one decision node to model a continuous process in which decisions can occur at any point in time. Finding these nodes requires an understanding of the operation as a whole, and it is advisable to start thinking at an early stage about the possible ways in which an operation might evolve and which (hidden) variables can be used to describe these evolutions.

Desiderata

We have shown that decision analysis can be applied to operational decisions in drilling operations. We have discussed some of the challenges we met, and how they could be overcome. But to truly integrate a rational decision making process into today's drilling operations, a number of more basic challenges need to be addressed. They have to do with the facts that (i) drilling programs today do not include prior uncertainty assessments of the key events and (ii) the domain experts are ill equipped to assess these uncertainties as this is not a central part of their training and operational experience. In addition, we believe the following challenges need to be addressed to improve operational decision making during the drilling process:

- The drilling program needs to identify and describe the key decisions.
- The utility scale needs to be clarified. All involved parties need to agree on which values their decisions are intended to maximize.
- All of the relevant uncertain events should be defined, described, and have a prior probability assessments (this includes uncertainty quantification for cost and value parameters)
- The drilling program should be based on expected values. Today any unexpected event is by definition negative.
- The relevance of measured real time data to the key uncertainties must be identified and quantified.

Today, the operational drilling activity is not described as or thought of as a decision making activity and instead relies on the use of intuitive decision making by more or less seasoned experts. This may work well enough in many cases, but it will not give good, consistent decisions in general, and it is bad for clarity, transparency, repeatability, and learning.

Future Work

A number of additional techniques can be applied to decision making in drilling operations. One example is sensitivity analysis, which can be used to determine how strongly the result of an ID computation depends on the values of the observable variables. Sensor readings are always subject to some degree of uncertainty, and sensitivity analysis can help to find out how

heavily a recommended action depends on the given data. This is particularly useful in combination with human background knowledge, like e.g. that some sensor currently has a particular problem.

A related technique is the *value of information (VOI) analysis*. In many cases, additional data could be obtained about a case, which might help to make a better decision. On the other hand, there is usually a cost to data gathering. For instance, in a drilling operation, an expert might be asked to analyze cuttings. This would possibly give more information about the geology, but it also takes some time. VOI analysis is a method to determine the expected increase in utility from the additional information, and thus the maximum amount that should be spent on obtaining the information.

While these techniques concern real time data, the assessment of configuration data has been a problem in our case study. In some cases, it will not be possible to assess the probabilities required for the ID model, or at least only poorly. We are currently devising a method based on automated logical reasoning that allows deriving recommendations in some cases even when some of the configuration information is missing. Ultimately, with our logic-based approach, we hope to integrate aspects like temporal reasoning, or reasoning about the different state of knowledge of the involved actors into the decision analysis process.

Conclusion

We have applied the decision analytic method to a particular kind of operational decision that occurs in petroleum drilling operations. As a case study, we took a real situation that occurred on a drilling platform on the Norwegian continental shelf in 2002. Through a systematic process, we obtained an influence diagram that expresses the structure of the decision. This structure can be reused for similar decisions in other operations. Given data from the drilling program and further assessments from domain experts, the ID can be configured to recommend operational decisions based on the available sensor data. We tested the approach with a well simulator, and it behaved satisfactorily. We have compiled a number of recommendations for future decision analysis endeavors for this kind of operational decisions.

References

Bratvold, R.B. and Begg, S.H. 2010. *Making Good Decisions*, SPE, Dallas.

Cayeux, E. 2007. eLAD: Well simulator specifications. Report Nr. 2007/182, IRIS, Stavanger, Norway.

Dake, L.P. 1978. Fundamentals of Reservoir Engineering, Amsterdam: Elsevier.

Good, I.J. 1961a. A Causal Calculus I. *British Journal of Philosphy of Science* **11**, 305 - 318.

Good, I.J. 1961b. A Causal Calculus II. *British Journal of Philosphy of Science* **12**, 43 - 51.

Howard, R.A. 1990. *From Influence to Relevance to Knowledge* Influences Diagrams, Belief Nets and Decision Analysis, ed. R.M. Oliver and J.Q. Smith. New York: John Wiley & Sons.

Howard, R.A. and Matheson, J.E. 1981. Influence Diagrams. In *The Principles and Foundations of Decision Analysis*, ed. R.A. Howard and J.E. Matheson 719-762. Menlo Park, CA. Strategic Decisions Group.

Hulsund, J.E., Nilsen, S., Nystad, E., Rø, M., Strand, S., and Bisio, R. 2009. Requirements for experiments and laboratory set-up in eLAD. Report, Institutt for Energiteknikk (IFE), Halden, Norway.

Kim, J.H. and Pearl, J. 1983. A computational Model for Combined Causal and Diagnostic Reasoning in Inference Systems. Paper SPE (IJCAI-83) presented at the Eighth International Joint Conference on Artificial Intelligence

Lauritzen, S. and Spiegelhalter, D.J. 1988. Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems. *Journal of the Royal Statistical Society series B* **50**(2), 157-224.

Madsen, A.L., Jensen, F., Kjærulff, U.B., and Lang, M. 2005. HUGIN – The Tool for Bayesian Networks and Influence Diagrams. *Intl. J. of Artificial Intelligence Tools* **14** (3): 507–543.

Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA: Morgan Kaufmann.

Rajaieyamchee, M.A. and Bratvold, R.B. 2009. Real-time Decision Support in Drilling Operations using Bayesian Decision Networks. Paper SPE 124247 presented at the SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, 4-7 October.

Wright, S. 1921. Correlation and Causation. *Journal of Agricultural Research* **20**(2), 365 - 379.

Wright, S. 1934. The Method of Path Coefficients. *Annals of Mathematical Statistics* **5**, 161 - 215.