From Information Technology to Knowledge Technology – Adaptive Advisory System for Oil and Gas Operations

Master Thesis

by

DI Theresa Baumgartner

Submitted at the

Institute of Economic and Business Management

University of Leoben

Leoben, October 2011

STATEMENT OF TASK

DI Theresa Baumgartner is assigned to address the topic

"Adaptive Advisory System for Oil and Gas Operations"

in a Master Thesis.

The objective of this Master Thesis is to evaluate the current status of research on business intelligence and expert systems applied in the area of oil and gas production operations. The candidate shall evaluate different methodologies of artificial intelligence regarding their implementation for production monitoring and surveillance systems.

It is then expected that the candidate formulate own ideas on how the concept of a knowledge unit can be realized in a fully automated production monitoring system. The knowledge unit shall be a learning layer in the production monitoring system that detects events, captures decisions, evaluates their impact and uses them to improve its future event detection and advisory. The candidate has to come up with the necessary metrics to measure performance and performance improvement and has to document a workflow to deploy the knowledge unit. It is expected that a working prototype version for the Knowledge Unit can be built based on the work of the thesis.

Leoben, October 2011 o.Univ.Prof. Dr. Hubert Biedermann

AFFIDAVIT

I hereby affirm in lieu of an oath that the present diploma thesis entitled

"From Information Technology to Knowledge Technology - Adaptive Advisory System in Oil and Gas Operations"

has been written by myself without the use of any other resources than those indicated and quoted.

Leoben, October 2011

Theresa Baumgartner

ABSTRACT

In today's increasingly digitalized world, companies have to deal with ever enormous data floods, which eventually lead to a data overload undermining decision quality of their employees. The petroleum industry in particular faces the problem of an aging workforce with more experienced people retiring than fresh engineers stepping in. This means that within the next years fewer engineers have to cope with increasing volumes of data and information in a more and more complex operational environment and the permanent loss of accumulated knowledge of the retiring experts.

The work suggests that a solution lies in the focus of knowledge management's efforts on the level of information technology (IT). Data, information and knowledge form a pyramid where the amount of data generally decreases towards the top, whereas the structure and logic increases. Today's IT systems are stuck at the information level of the pyramid. Knowledge can be automatically generated from data and information and universally reapplied to new situations. Pushing information technology towards a knowledge technology (KT) with highly adaptable and learning network structures could bring a long required software innovation to companies. The work emphasizes that with intelligent systems, companies learn how to learn faster, thus secure a sustainable competitive advantage.

As a practical example of the implementation of KT, the Adaptive Advisory System is introduced. It is a decision support system for oil and gas operations that makes use of artificial intelligence to infer knowledge from data and information, store and adapt to new challenges and therefore steadily increases the quality of results along with its usage. It is designed to complement oil company's assets teams for highly complex problems with limited expert availability, but with a great flexibility of application. In addition to user-system-interfaces, it comprises from a Data, an Information and a Knowledge Layer, with the latter being the centerpiece of the Advisor.

The Knowledge Layer consists of the three different sections: the Event Detector identifies real-time data deviations and alerts in case of emergencies. Combining these events, the Problem Classifier assesses the situation and infers the most likely problems. The Decision Supporter calculates the utilities of actions and recommends the best solutions to the user. In a final evaluation, the decision projection is compared to the actual results some time after the action to further improve the results.

Event Detector, Problem Classifier and Decision Supporter are based on Bayesian networks, an instrument of artificial intelligence, that is widely and successfully used for knowledge representation and reasoning under uncertainty. Bayesian networks are defined as acyclic directed graphs with nodes (random variables) that are interconnected by edges (conditional probabilities).

The work describes the principles, structures as well as the operating modes of the Adaptive Advisory System and illustrates these by examples from the area of well production monitoring. It proposes several ideas for a user interface in order to meet the requirements of a transparent and flexible Advisory System. Finally the focus is put on the human factor, since the best system is worthless without being accepted by the user.

KURZFASSUNG

In einer zunehmend digitalisierten Welt stehen Firmen heute oftmals vor dem Problem von regelrechten Datenfluten, die ab einem gewissen Volumen negativ auf die Qualität von Entscheidungen auswirken. Die Erdölindustrie sieht sich zudem mit dem Problem einer kurz alternden Belegschaft konfrontiert, von der mehr Experten aus der Arbeitswelt ausscheiden als nachbesetzt werden können. Kurzum, in den nächsten Jahren müssen immer weniger erfahrene Ingenieure ein Mehr an Daten und Information bewältigen, und das bei zunehmend komplexeren Aufgabestellungen der Erdölgewinnung und einem dauerhaften Verlust von angesammelten Wissen zusammen mit den pensionierten Mitarbeitern.

Diese Arbeit schlägt als Ausweg aus dieser Situation vor, die Bemühungen des Wissensmanagements auf der Ebene der Informationstechnologie zu forcieren. Daten, Information und Wissen bilden eine Hierarchie, bei der die Menge an Zeichen nach oben hin ab und die Strukturen zunehmen. Heutige IT Systeme bleiben auf der Informationsebene stehen. Eine Vielzahl an Daten und Informationen kann zu allgemein einsetzbarem Wissen subsummiert, als solches gespeichert und für neue Situationen angewendet werden. Der Schritt von IT hin zu einer Wissenstechnologie (Knowledge Technology – KT) mit anpassungsfähigen und selbstlernenden Netzwerkstrukturen würde eine lange ausständige Software Innovation mit sich bringen. Die Arbeit macht deutlich, wie Firmen durch intelligente Systeme schneller lernen können und sich damit einen nachhaltigen Wettbewerbsvorteil sichern können.

Das Adaptive Advisory System, als praktische Umsetzung von KT, ist im weiteren Sinne ein System zur Entscheidungsunterstützung für Öl- und Gas Operationen, das mit Hilfe von künstlicher Intelligenz fähig ist, eigenständig zu lernen, sich stetig neuen Herausforderungen anzupassen und damit mit zunehmendem Einsatz immer bessere Resultate liefert. Es ist für die primär für die Anwendung des Well-Production-Monitorings ausgelegt, ist aber grundsätzlich sehr breit einsetzbar. Das System besteht aus neben einer Benutzerschnittstelle aus einer Daten-, Informations- und Wissensebene, wobei die Wissensebene (Knowledge Layer) die Herzstück des Advisors darstellt. Der Knowlede Layer ist wiederum in 3 Abschnitte unterteilt: der Event Detector stellt Abweichungen in den Echtzeit-Signal (Events) fest und alarmiert den Benutzer. Der Problem Classifier analysiert die Events und kann diese bestimmten Problemen zuordnen. Der Decision Supporter berechnet den Nutzen für verschiedene Handlungen und schlägt die bestmögliche Problemlösung vor. Eine abschließende Evaluierung der Entscheidung als Vergleich mit realen Ergebnissen sorgt für eine zusätzliche Verbesserung der Resultate.

Event Detector, Problem Classifier und Decision Supporter basieren auf Bayes’schen Netzen (BN), einem Instrument der künstlichen Intelligenz, denen der Satz von Bayes zugrunde liegt. Ein Bayes'sches Netz ist ein Graph bestehend aus Knoten (Zufallsvariablen) und Kanten (bedingten Wahrscheinlichkeiten); es vermag Wissen aus Daten zu extrahieren, zu speichern und anzuwenden, Rückschlusse unter Unsicherheiten zu ziehen und zu lernen.

Die Arbeit beschreibt die sowohl Prinzipen, Strukturen und Operationsmodi des Adaptive Advisory Systems, demonstriert die Prinzipien anhand eines Beispiels aus dem Well-Production-Monitoring und bringt Vorschläge für eine Benutzeroberfläche. Abschließend werden die Implikationen einer resultierenden erhöhten Transparenz auf eine Organisation diskutiert.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF EQUATIONS

LIST OF ABBREVIATIONS

AI	artificial intelligence
ANN	artificial neural networks
BI	business intelligence
BN	Bayesian networks
BPT	Bayesian probability theory
cf.	confer (lat.) = see also or compare to
D	data
DBMS	data base management system
DDS	decision support system
DGMS	dialog generation and management system
DS	Decision Supporter
e.g.	exempli gratia (lat.) = for example
E&P	exploration and production
EBM	evidence based medicine
ECD	equivalent circulating density
ED	Event Detector
ES	expert system
et al.	et alteri or et alii (lat.) = and others
I	information
i.e.	id est (lat.) = which means, in other words
IO	integrated operations
IT	information technology
K	knowledge
KM	knowledge management
KPI	key performance indicator
KT	knowledge technology
MBMS	model base management system
N.N.	nomen necsio (lat.) = unknown author
NN	neural networks
O&G	oil and gas
PC	Problem Classifier
TVD	total vertical depth
UML	unified modelling language

INTRODUCTION

"We are drowning in information, but starving for knowledge."[1]

Many oil and gas asset teams nowadays face increasing challenges in daily operations. First of all, an increased digitalization of the fields has led to more and more field sensors that deliver huge amounts of high frequency data. Furthermore it requires highly skilled and experienced people on site as well as in the offices, who early enough detect performance problems in the field, analyse them and suggest remedial activities to correct them. These highly-experienced experts and together with them their highly valuable knowledge are not available at all times and for ever. Asset teams have to strive for lean technologies and efficient processes to maintain production and at maximize profits in very competitive environment. Maintaining a high level of production requires complete asset awareness and hence occupies a significant amount of the asset's resources. Hite[2] discussed how 91% of participants in a survey conducted by the SPE Real Time Optimization Technical Interest Group spend more than 50% of their time to look for, access and prepare data, which ultimately leaves less than 25% of their professional time for analysis, evaluation options and decisions. Based on that work, Brulé[3] proclaim in his paper how "faster decisions with precision have tremendous value, and provide much leverage in any industry hindered by a shortage of qualified people".

In contrast to conventional batch processes (e.g. in a factory conveyor belt or in a racing car where processes are repeated in every loop) a reservoir cannot be seen as a repeatable process. Conditions in the reservoir continuously change as production is going on. Constraints may change or even disappear completely while new constraints may come up. Production from a greenfield for example is typically constrained by the number of wells, whereas common brownfields mostly are constrained by the deliverability of the reservoir or the capacities of the facilities.

Under these conditions, asset teams are looking for the right technologies to capture, continuously update and apply knowledge of skilled personnel, in order to streamline not only production processes but also knowledge processes. Knowledge capturing technologies would enable asset teams to not only automate repetitive processes or model executions, but also to support or automate complex decision making processes that typically involves experts from various disciplines. The ever-changing environment in the oil and gas industry requires a system to be adaptive, thus the system is able to learn together with their daily tasks and knowledge gained in previous processes can be easily applied to other processes, which are not necessarily similar.

An intelligent technology, or knowledge technology, would enable asset teams to make better and faster decisions in oil and gas operations, leading to more efficient, sustainable and transparent processes to increase production with minimizing the costs and better meeting the world's energy demand in the future. Therefore an Adaptive Advisory System is proposed in this work, which employs methodologies of artificial intelligence to assess the complex conditions of oil and gas operations, continuously improving results and to learn and adapt to unstable conditions.

---

[1] John Naisbitt, American businessman
[2] Hite et al. (2007), p. 1
[3] Brulé et al. (2008), p. 1

1 BACKGROUND & THEORY

1.1 Knowledge Theory

The Adaptive Advisor proposed in this work is a system that helps engineers to make decisions. Its main characteristic and the innovation that its implementation brings along is the application of artificial intelligence to a general approach of well production management. The system the work proposes is capable of learning and capturing, sharing and reapplying knowledge utilizing methodologies of artificial intelligence.

This chapter tries to explain what learning actually means, in mathematical sense as well as for humans and for organizations. We need to know what learning is to investigate the benefits it brings to companies by applied in the right way, giving them a sustainable competitive advantage. Since there is one-above-all yardstick for intelligence - the human brain - its functional principles will be described and compared to the attempts to artificially model the brain and its learning capabilities.

1.1.1 Learning

According to the Oxford dictionary, *learning* in general is the acquisition of knowledge or skills through study, experience, or being taught. For optimal usage, next to the process of learning, the skills of remembrance (memory) and recalling (application of learned matter) are required. Learning is more than the simple storage of information; it includes the perception and evaluation of the environment, connections with previous knowledge and experiences and the recognition of structures or patterns.

Human beings learn with everything they do. Learning can be done in a fully conscious and intentional way, but it also happens implicitly all the time. Our brain consists of a large network of interconnected neurons, which change its structure when we are learning. Learning has to be understood in a broader sense: not only humans and animals can learn, but machines, systems, and even organizations are capable of acquiring knowledge.

Machine learning is the artificial generation of knowledge from experience. Computers can be more than tools that save data and information. With certain algorithms they are capable of extracting structures from data and recognize patterns from new inputs.

A *learning organization,* according to David Garvin[4], 'is an organization skilled at creating, acquiring, interpreting, transferring and retaining knowledge, and at purposefully modifying its behaviour to reflect new knowledge and insights.' This definition contains three elements: First the generation of new knowledge, second the knowledge transfer and retaining and thirdly the practical application of knowledge. Hence learning is a function of actions with regard to knowledge. [5]

Math of Learning

Steven Flinn[6] uses the terms of *knowledge* and *learning* in a very abstract way and visualizes them with mathematical expressions: From the definition of learning we can conclude that it leads to an increase in knowledge, or in business terms, increase in intellectual capital.

[4] Garvin (2000), p.11
[5] Flinn (2010) p. 17
[6] Flinn (2010) p. 17

Knowledge can be measured in absolute terms at every point in time, therefore it is stock variable. Learning is the difference between the same stocks of knowledge at two different points in time and considered a flow variable that feeds the stock of knowledge with a certain rate.[7]

Following this logic, learning is the derivative of the knowledge function with respect to time, or inversely, cumulating or integrating learning over a period of time yields knowledge.

Competitive Advantage: Innovation

Arie de Geus[8] says, 'Learning to learn faster is the only sustainable competitive advantage.' Applying this principle to the mathematics of learning means that the rate of learning has to be increased, therefore competitive advantage is the acceleration of gain in knowledge:

$$competitive\ advantage = \frac{d\ learning}{dt} = \frac{d^2 knowledge}{dt^2}$$

Equation 1: Competitive advantage[9]

In Newtonian terms the knowledge would be the distance, learning the speed and the competitive advantage the acceleration.

Knowledge in companies should grow exponentially, or even super-exponentially. The growth of world population, world economy, CO_2 concentration in the atmosphere, living standards, prices etc. over the years are examples for such a super-exponential function. For knowledge, a super-exponential growth would imply going one step further and even accelerate learning.

Geoffrey West draws a picture of us on a treadmill that goes faster and faster, but we have to exchange the treadmill by a better one faster and faster.[10] He claims that the only way how economies and businesses could sustain an increased acceleration rate of growth is by *innovation*.

In the last 50 years we already could experience some advances in information technology: Steven D. Flinn[11] in his book "The Learning Layer" identified three major waves in IT development since the beginning of the computer era: speed, connectivity and adaption. Speed was the primary limiting factor from the beginnings of IT in the 1960ies until the 1990ies, but infrastructure changes and advances in hardware development allowed a faster processing of data. From the early 1990ies on, the Internet gained a public face and made it possible for everybody to reach any other person that is connected to the world wide web.

According to Flinn, we are currently experiencing a third wave, *adaptation*, which is based on virtual connection. Adaptation constitutes an advancement of virtual connection, because it enables two things: Firstly it offers a much bigger pool and a broader range of experiences to learn from to everybody. Secondly, the Internet has changed the way we structure things: we are turning our backs from the traditional hierarchy structure, towards a

---

[7] A wrong conclusion would now be that the stock has a limited capacity and can be filled by learning, which is clearly not the case. Quite the contrary, the rate of learning is greatly increases with the "amount of knowledge" already available in the stock. Hence the relationship of learning and knowledge is a non-linear function.

[8] Arie de Geus is former manager of Royal Dutch Shell, business strategist and author of books and articles on the 'Learning Organisation' concept

[9] Modiefied from Flinn (2010), p. 17

[10] West (2011)

[11] Flinn (2010)

network structure. A network stands for an interconnected adaptable way of processing data and information to extract knowledge and forms the basis of a new technology of knowledge.

1.1.2 The Next Innovation: From IT to KT

In the last 50 years, IT has undergone some changes in character, but the basic concept of IT has not substantially changed. Considering the incredible increase in digital data we are currently facing, innovation is absolutely necessary. Innovation means creating smarter systems for organizations, businesses and processes. In Chapter 1.1.4 it will be explained that data, information and knowledge form a hierarchy and it is time to make the step from information to knowledge.

Data Floods

Despite the immense increase in productivity and more efficient and effective decisions we could make with the help of smarter systems, the increased amount of data makes it necessary to change the way we are dealing with it.

The IDC, the international data corporation, estimates in its 2010 Digital Universe Study[12] that the amount of all digital data produced worldwide, or the 'Digital Universe', despite the global recession has grown by 62 % in 2009 to nearly 800,000 petabytes[13]. According to the IDC report, the 'Digital Universe' will grow by a factor of 44 until 2020. In contrary, the money spent on IT staff will only grow by a factor of 1.4. At the same time, the money invested in managing one byte will be dramatically decreased. So we can expect businesses being flooded with even more data in the future.

Jay Bourland, group technology officer at Pitney Bowes says that CIOs consider making sense of all the data they are dealing with as their biggest challenge in the upcoming years:[14] 'It's a mass of structured, unstructured and real-time data, data in storage, operational data, marketing data, data from external sources, and model, predictive and past-result data. To deal with all of that, you need mechanisms in place to manage it and make sure it's fit for use and that it's being used appropriately. We've been calling that data quality and data integration.'

The pyramid of data, information and knowledge, that will be described in Chapter 1.1.4 The DIK(W) Model suggests that when going up in the hierarchy, the logic and structuration of data increases, whereas the number of data points decreases. Therefore the next innovation in data handling will be the transition to KT - Knowledge Technology.

Knowledge Technology or Making IT Systems Smart

Information technology (IT) is the acquisition, processing, storage and dissemination of data and information in various forms by a combination of computing and telecommunications. The term in its modern sense first appeared in a 1958 article published in the Harvard Business Review[15], in which authors Leavitt and Whisler commented: 'The new technology does not yet have a single established name. We shall call it *information technology*.' The authors already proposed progressively that IT would process large amounts of data, apply

---

[12] Digital Universe Study (2010), p. 1
[13] 1 petabyte = 1000 terabytes or 10^{15} bytes
[14] Sperling (2009)
[15] Leavitt, Whisler (1958)

statistical and mathematical methods and simulate higher-order thinking through computer programs.

Imagine the following situation: On a dinner table, Mark tries to take part in a conversation. Whenever Mark hears a set of keywords he recognizes the topic the other guests talk about. He simply digs out previously collected information on the topic and drops it in form of sentences in the middle of the conversation. Mark's comments may add to the debate if he is lucky, but they do not properly connect to what the others said. The other guests would immediately realize that Mark does not actually know what he is talking about.

IT systems in companies today are just like Mark, they contain an enormous amount of data and information, we can look for it with powerful search tools, but the results have to be chosen, adapted and interpreted from intelligent human beings. The IT systems we are using today are capable of processing data and information, but cannot yet make the step to intelligent systems and hence building a stock of knowledge. I would prefer working with a computer that acts like a smart person, rather than Mark.

The term *Knowledge Technology* (KT) stands for a long outstanding revolution of the IT systems that are in place for decades. KT systems are capable of delivering knowledge in form of recommendations and arguments for it. But not only can they offer results, they help the user by selecting the right information for display to the user from any source available. Actually, "source" is the wrong terms, since an important change that KT brings, is to virtually interconnect all data, information and knowledge, such that we do not speak of single sources like isolated databases, other engineer's desktops, books, etc. anymore, but combining them all into one network. KT systems are adaptable to the requirements of the users as well as the data processed. They are evolving in time the faster the more often they interact with human users and learn from them. KT will change the way organizations work into much more transparent structures with lower hierarchies and less interdependence of people. Lowering the walls of power and competition, it will enable the knowledge sharing and the cooperation across various disciplines and locations.

Knowledge Management

Knowledge Management (KM) comprises a range of strategies and practices used in an organization to identify, create, represent, distribute, and enable adoption of insights and experiences. Such insights and experiences comprise knowledge, either embodied in individuals or embedded in organizational processes or practices. It can be seen as a process that includes six core activities: *formulation of knowledge goals* and *knowledge identification* form the starting point and trigger an inner operational knowledge management process circle, which consists of the activities of *knowledge development, storage, distribution* and *application*.[16] Knowledge technology focuses on the can "hardware" aspects of knowledge management, with still having the same goals. But it is obvious that technology alone cannot obtain the best results, because the people can still chose whether to use it or not.

Knowledge can be *explicit* or *implicit*. Michael Polanyi has defined *explicit* knowledge in 1966 as knowledge that has been or can be articulated, codified and stored in certain media. It can be readily communicated to others for example via certain media like encyclopaedias, books, blogs, etc.; or in Polanyi's words: 'What is usually described as knowledge, as set out in written words or maps, or mathematical formulae, is only one kind of knowledge; while unformulated knowledge, such as we have of something we are in the act of doing, is an-

[16] Bodendorf (2006), p.133

other form of knowledge.'[17] Bringing it down to "we know more than we can tell", Polanyi thinks that verbalizing, and therefore also formalizing of tacit knowledge is not possible due to its nature.[18]

Implicit or *tacit* knowledge however is can be seen as experience that is hard to describe. We often call it "having a feeling for something", just like a driller implicitly knows how much weight on bit is required for each situation, but he could not put it in words. Tacit knowledge can also be represented in machine learning: Neuronal networks do the right thing, but we have no insight in their "black box" of internal structures and cohesion of nodes.

In practice, for example in large oil operator companies, knowledge management is perceived by employees as a system of intranet platforms, reporting tools, lessons-learnt-approaches or forums where other employees or experts answer posed questions. An important task of knowledge management is a change of paradigm of how mistakes are only negative if nobody else but the person who made it can learn from it. Up to now the value KM brings to an organization is very much dependent on the commitment of each and every single person voluntarily sharing his or her explicit knowledge and information.

For the explicit parts of knowledge, the *knowledge transfer* can be realized relatively easily with tools of information technology. Whereas implicit knowledge is generally transferred by time-consuming and complicated ways, such as imitation, observations, analogies or learning-by-doing with quite limited technological support.[19]

In chapter 1.1.1 "Math of Learning" however offers a clue to this problem: In order to transfer knowledge, people can look at the *stock of knowledge* itself or look at the *flow of learning*, hence how people did get to the knowledge and get an idea of what is in the stock as well. Also if the knowledge is not visible because it is tacit knowledge, we still can look at the flow and derive it by the integration of the learning function.

KT however would provide the means of directly extracting knowledge at the source – at the moment it is created in our brains in the first place. The system ideally mirrors the flow of data and information that is perceived by employees and learns in the same way as a human being and by these means it is therefore capable of transforming some tacit knowledge to explicit knowledge.

1.1.3 How Our Brain Works

John Haugeland said in 1985 that AI was "the exciting new effort to make computers think … machines with minds, in the full and literal sense". This approach of AI is called the cognitive modelling approach or simply *mind design*, which aims to model the brain.[20] If we want to create machines that think like humans, we first need to understand how the human brain works in principle, although it must be mentioned that the exact process is not yet fully understood by scientists.

The study of the nervous system is called neuroscience. Our brains are made out of many *neurons*, individual simple cells that manage to interact in a way that leads to thought, action and consciousness. Figure 1 shows the parts of these nerve cells or neurons. Each neuron consists of a cell body that contains a cell nucleus. Branching out from the cell body are a number of fibers called dendrites and a single long fiber called axon, which is typically 1 cm

[17] Polanyi (1958), p. 12
[18] Ahlert (2006), p. 45
[19] Ahlert (2006), p. 66
[20] Russel, Norvig (2010), p. 3

long (approx. 100 times the diameter of the cell body) but can reach up to 1 meter.[21] One neuron has connections with 10 to 100,000 other neurons at junctions known as synapses.[22] So about hundred billion (10^{11}) neurons form about hundred trillion (10^{14}) connections. Signals, i.e. neurotransmitters are released and propagate through the brain, from neuron to neuron. At a neuron, the cumulative effect of the incoming signals can change its electrical potential. When a certain threshold is reached, the neurons "fire", they are setting off a pulse that can do the same in another neuron and so on. The connections are weighted, so the effect of signals on the electrical potential differ for different connection paths. The signals control brain activity in the short term and also enable long-term changes in the connectivity of individual neurons.

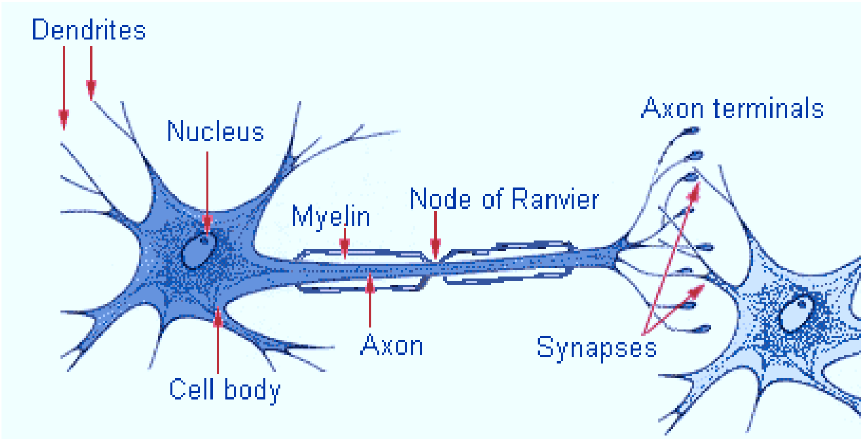


Figure 1: Parts of a neuron[23]

Recently, with the development of imaging techniques of the brain activity, such as functional magnetic resonance imaging (fMRI), studies[24] could prove that learning correlates with adding, deleting, strengthening and weakening the connections between the neurons in the brain. Experiences trigger a modification of the physical structures in the brain.

By 2008 the *Blue Gene* by IBM was the world's most powerful supercomputer. At that moment in time, it was able to process 10^{16} bits per second. This processing speed is approximately comparable to what a human brain can do. The main difference in power between computers and brains its energy consumption. Whereas the supercomputer needs 1.5 mega watts of energy, the brain takes 100,000 times less – approximately 10 watts.[25] Talking about efficiency only, there is still a long way to go until computers come only close to our brains.

1.1.4 Data, Information, Knowledge

The terms *data*, *information* and *knowledge* are integral part our general linguistic usage. Their meanings are obvious to all of us, but when it comes to integrated adaptive systems a clear definition and proper distinction is necessary, since data, information and knowledge is processed in different ways.

---

[21] Neurons and axons are not exclusively found in the brain but connect all body parts with the brain and spinal cord.
[22] Russel, Norvig (2010), p. 11
[23] PATTS (2011)
[24] for example Keller (2009)
[25] Boahen (2007)

The task of finding one valid definition for these concepts seems not quite promising. Philosophers and scientist tried to solve questions like "What is Knowledge?" in a general sense since ancient times – without success. In a study by Chaim Zins, an Israeli information researcher, 130 different definitions of the concepts of data, information and knowledge were collected.[26]

The DIK(W) Model

The *DIK model* or *knowledge pyramid* shows the hierarchy of data information and knowledge. Figure 2 shows the pyramid structure of data, information and knowledge. Whereas the amount of data generally decreases when going up in the pyramid, the structure and logic increases. Further definitions and interrelationships of the concepts will be explained in the next chapter.

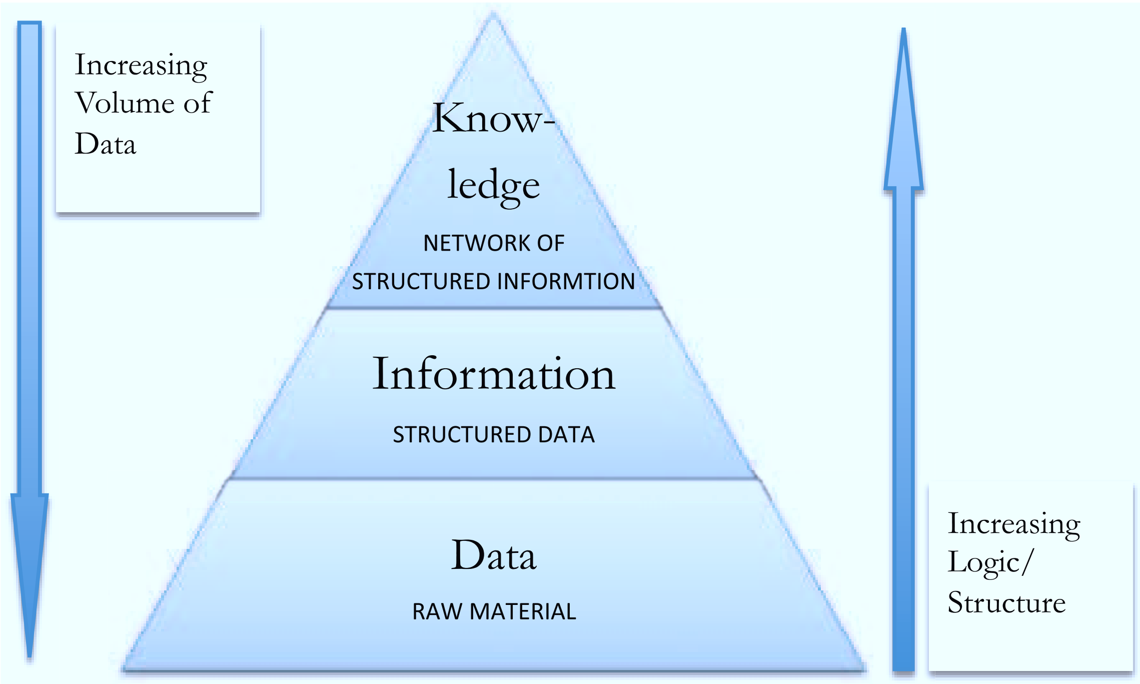


Figure 2: Knowledge pyramid[27]

Some sources extend the model by *wisdom*, to the *DIKW model*.[28] The model is used in knowledge or organization management model of decision-making. The introduction of the DIKW model is attributed to Dr. Russell Ackoff, pioneering the field of operations research, systems thinking, and management science.[29] The concepts can also be seen as *know-nothing, know-what* and *know-how*.

Wisdom comes from the repetition of the DIK cycle and according to Richard Gayle[30] it represents the best and most appropriate action, and requires reflection of knowledge – the *know-why*. It is the key to making the most effective decisions and predicting the future by inference.

---

[26] Zins (2007)
[27] Modified from Voß, Gutenschwager (2001), p. 13
[28] for example see Rowley (2007)
[29] Drogseth (2011)
[30] Gayle (2009)

Computational Information Processing

Did you know that it is not quite right to talk about "knowledge in a book, unless the book has reasoning capabilities? The following chapter should prevent the ambiguous use of the terms data, information and knowledge.

Agnar Aamodt and Mads Nygård[31] defined data, information and knowledge from the perspective of computational information processing, which is in line with conventional theories of knowledge, like the DIK or DIKW model.

Finding fixed definitions for the three terms turns out to be not an easy task. There is no known way to distinguish data, information and knowledge on a representational basis, i.e. items or structures "look" the same on paper or in a machine.

The authors created a definitional framework, discussing the concepts of data, information and knowledge in the simple context of single agent decision-making. The concepts by Aamondt and Nygård are summarized below and illustrated in Figure 3:[32]

- Data are syntactic entities
  - Data are patterns with no meaning: they are input to an interpretation process, i.e. to the initial step of decision making
- Information is interpreted data
  - Information is data with meaning; it is the output from data interpretation as well as the input to, and output form the knowledge – based process of decision making
- Knowledge is learned information
  - Knowledge is information incorporated in an agent's reasoning resources, and made ready for active use within a decision process; it is the output of a learning process.

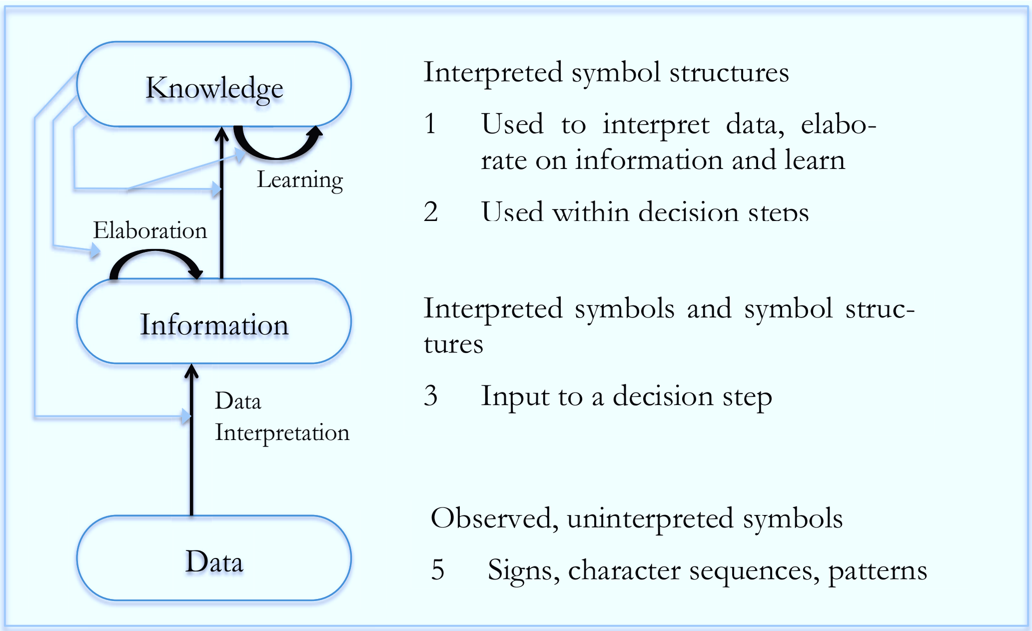


Figure 3: The data – information – knowledge model[33]

[31] Aamodt, Nygard (1995), p. 6-8
[32] Aamodt, Nygard (1995), p. 6-8

Data Interpretation

Data are uninterpreted signals, characters, patters, codes, etc. that have no meaning for the agent, i.e. the human or the machine. The transition from data to information occurs when there is given meaning to the sequence of signs. To be able to do that, the agent or system has to have certain knowledge. The process of creating information out of input data using certain knowledge is also known as *inference*.

"01123581321345589" is a simple sequence of numbers; data with no meaning to most of us. But this sequence actually represents the first 11 numbers of the famous Fibonacci series, that starts with 0 and 1 and always the last two numbers are added. We interpret "0, 1, 1, 2, 3, 5, 8, 13, 21, … " as Fibonacci sequence, which has a certain meaning to us, and therefore is information.

Another example of this is inferred information on a gold ring using input data on temperature and the condition of aggregation with the knowledge of the melting point of gold T_m = 1064.18°C.[34]

Input Data	Inferred Information
The ring is of solid gold. Temperature is 1000°C	The ring won't melt
Temperature is 1000°C. The ring melts.	The ring is not of gold.
Temperature is 1100°C. The ring does not melt.	The ring is not of gold.
Temperature is 1100°C. The ring melts.	The ring is possibly of gold.
The ring is of solid gold. It does not melt.	The temperature is lower than T_m.

Figure 4: Inferred information from input data[35]

In other words, for data to become knowledge, an interpreter is required. A human interpreter may use his social and cultural background, unconscious associations, textbook knowledge, certain memories or previous experiences to determine the contextual meaning of data. Computer systems unfortunately can yet only draw on a limited and pre-defined pool of findings.

Elaboration of Information

The elaboration of information is represented by the arc in Figure 3. Once an initial set of information is gained from data, it is elaborated upon for a better understanding and new information is derived from it. This can be seen as a circle: the elaboration leads to a better understanding of data and therefore new data can be interpreted to extract information and so on.

Both interpretation and elaboration process require knowledge. In the case of elaboration, knowledge serves a different role than information, which functions as an input whereas knowledge can be seen as a inherent resource of reasoning agent. For example: 'If in an

[33] Aamodt, Nygard (1995), p. 8
[34] Pourret (2008), p. 4
[35] Pourret (2008), p. 4

production advisory system the information "flowline pressure has increased from 50 to 100 bar during the last hour" is given, a system may use its knowledge to infer "strongly and rapidly increased pressure", from which "possible plugging of pipe" may be concluded, in turn leading to the action "shut down pump"'.

Amodt and Nygard emphasise that computational methods for inference and elaboration should be non-deterministic processes, considering that strict algorithmic approaches lack degree of flexibility necessary. They argue that AI research had shown cognition-inspired language of knowledge structures and inference methods to be well suitable in this case.

Learning of Knowledge

Figure 3 shows two learning arrows, which represent two different types of gain in knowledge. In the context of a decision-making system we can say that learning is the integration of new information into an existing body of knowledge. But new knowledge also comes from inference processes within the body itself.

One method to distinguish knowledge from data and information is by its flexibility. Once extracted from a specific context like "apples on a tree always fall down on the ground" we can infer the general principle of gravity.

1.1.5 Artificial Intelligence

Artificial Intelligence (AI) is one of the most exciting areas of research nowadays. Sebastian Thrun and Peter Norvig from the Stanford University in California have offered a free online course on Artificial Intelligence to everybody who is interested for Fall 2011 and 135,000 people from all over the world had signed up by September. This interest originates first of all from fundamental reasons: How our brain works is still one of the unsolved mysteries of science, by mirroring brain activities artificially we hope to get further insight into human mind. Secondly, the areas of application are widespread and fascinating: from driving cars automatically, understanding written text, search engines, medical and technical diagnosis and problem solving, modelling the financial world, etc. Peter Norvig[36] thinks that AI is a truly enabling technology and it will lead the way into the next century.

AI is the science and engineering of making intelligent machines, especially intelligent computer programs. Since John McCarthy coined the term in 1954[37], artificial intelligence has had many ups and downs: It has been hyped and then, having failed to live up to the hype, been discredited until being revived again. Attempts to define AI are doomed to failure, since the definition of intelligence in the first place seems to be a never-ending story among scientists.

AI is a truly universal field, its application range from deduction, reasoning and problem solving, knowledge representation, automated planning and scheduling to machine learning and classification, natural language processing, robotics, machine perception and speech recognition, etc. The tools are almost as numerous and diverse as the applications of AI, but what they have in common is that they are designed to work with large and complex data sets in dynamic environments. AI is seen as the key to managing the ever increasing amounts of data and information in businesses, and therefore also as the most promising "hard" approach to knowledge management in businesses. Artificial Intelligence incorporates the system described in this work on a small scale, intelligent methods, as well as on a big scale, how businesses can learn and make use of smart systems.

[36] Norvig (2011)
[37] Skillings (2006)

Soft Computing is a multidisciplinary field that was proposed 1981 by Dr. Lotfi Zadeh[38], professor of computer science at the University of California, Berkeley, who had the goal to construct new generation AI, known as Computational Intelligence.

Soft Computing can be seen as the fusion of the fields of fuzzy logic, neuro-computing, evolutionary and genetic computing, and probabilistic computing into one multidisciplinary system in order to develop intelligent machines and to solve nonlinear and mathematically unmodelled system problems.[39] Therefore, lately a lot of research on combinations of intelligent methods, such as neuro-fuzzy systems, fuzzy Bayesian networks, fuzzy petri nets, is performed.

1.2 Methods

In the previous chapter we got an idea of how our brains achieves amazing skills although it consists of many very simple and stupid cells. Efforts to model the neurobiological structures of the brain are not the only approach of copying the human intelligence by artificial means. The next chapter describes some of the concepts and methods of artificial intelligence as well as the framework they are integrated in to employ them to real problems.

In this chapter intelligent methods that are frequently used for expert systems or decision support systems are described. Bayesian networks are considered to be the best main method for the Adaptive Advisory System and therefore described in further detail to provide a basic understanding for the application in chapter 3. The basics of some alternative methods like artificial neural nets, fuzzy systems and others are described and are compared to Bayesian networks.

1.2.1 Integration Framework

Decision Support Systems

Starting off with single, isolated data systems like an engineer's personal excel sheet, there is a general trend to connect systems like databases, information systems and knowledge base systems to *integrated systems*. Examples for these integrated systems are among others decision support systems.

Decision support systems (DDS) are interactive computer based systems that aid users in judgement and choice activities.[40] DDS consist of a data component, a method and model component and a connection between them, i.e. a "dialog component".[41] The concept of decision support systems DSS is extremely broad, and it is used in a variety of disciplines such as statistics, economics, engineering and operations research development.

Human judgement and decision making has been the subject of numerous studies.[42] It has been shown that we judge situations and base decisions rather on intuitive strategies than on theoretically sound reasoning rules. These intuitive strategies, referred to as *judgmental heuristics* in the context of decision making, help us in reducing the cognitive load, but alas

[38] Zadeh (1981)
[39] Moussa (2003), p. 8
[40] Druzdzel, Flynn (2002), p. 6
[41] Voß, Gutenschwager (2001), p. 337
[42] See Dawes, R.M., *Rational Choice in an Uncertain World.* Hartcourt Brace Jovanovich, Publishers, 1988; or Daniel Kahneman, D., Slovic, P., *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge University Press, Cambridge, 1982.

at the expense of optimal decision making. Therefore, our unaided judgment and choice exhibit systematic violations of probability axioms; we are biased.

The desire to improve human decision-making provided motivation for the development of a variety of modelling tools in disciplines of economics, operations research, decision theory, decision analysis, and statistics.

The most obvious solution to the problem is to build a system that imitates human experts. But while human experts are excellent in structuring a problem, determining the components that are relevant to it and providing local estimates of probabilities and preferences, they are not reliable in combining many simple factors into an optimal decision. The role of a decision support system, especially those that employ decision-analytics, is to support them in their weaknesses using the formal and theoretically sound principles of statistics.

Decision support systems in general have three main components:[43]

- The database management system (DBMS) stores large quantities of relevant data and the logical structures that are required for decision support.
- The model-base management system (MBMS) has the role of data transformation of data into information useful for decision-making.
- The dialog generation and management system (DGMS) is the interface of the system and supports model building as well as the utilization of the software.

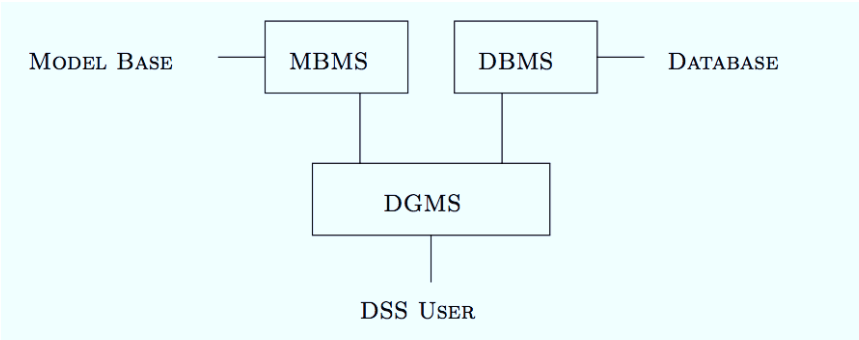


Figure 5: The architecture of a DSS[44]

The components are illustrated in Figure 5. Essentially, the user interacts with the DSS through the DGMS, which communicates over the MBMS with the model base and the DBMS with the database.

The model component integrates various (mathematical) methods of three main categories: exact methods, heuristic methods and meta-heuristic methods. Methods that employ artificial intelligence (such as the later described Bayesian networks) fall into the later class of DDS, also known as *decision-analytic DDSs*, which employ the principles of decision theory, probability theory, and decision analysis to their decision models.

Expert Systems

Expert systems can be seen as a special type of DDS, which aim at building support procedures or systems that imitate human experts. Expert systems use a knowledge base of hu-

---

[43] Druzdzel, Flynn (2002), p. 6
[44] Druzdzel, Flynn (2002), p. 7

man expertise and inference procedures for solving complex problems. They are a traditional application of artificial intelligence and are most commonly used for specific isolated problem domains. Therefore, there exist a wide variety of different systems. The operating principles of expert systems is to a large part the recognition of certain conditions plus following actions combined with some analytic capabilities.[45] Consequently, every expert system consists of two principal parts: the knowledge base and the inference engine.

Expert systems already separate the programming code from the expertise or data structures, such that a change of reasoning rules or knowledge does not require a change of the program code, as it is the case in traditional problem solving programs. Also, new rules can be added to the knowledge base rather easily.

In comparison to decision support systems, expert systems go a step further, applying the knowledge that had been taught to them so that they can provide answers and suggestions for decisions. They are usually using a much wider set of information, equations, rules, etc. than decision support systems, which only support a human user in the process of finding a decision, not executing a solution independently.[46]

An example for an expert system and a first step toward automation are *rule-based systems* with an IF-THEN-ELSE logic to extract knowledge from information. This system is comparable to a decision tree (e.g. *if* the well diameter is lower then 5" *then* take this action, *else* that one) and supports engineers to make fast decisions based on simple operational parameters under information overflow. These rule-based systems are easy to handle, but they face serious limitation when it comes to uncertainties, noisy or missing data and result ambiguities.[47]

1.2.2 History of Bayesian Networks

There exist two views on probability, the "*frequentist*" view and the "*Bayesian*" view. Although the approach of Thomas Bayes to treat probability as a degree of plausibility is the original approach and has many advantages over the "more objective" approach of probability being a frequency of occurrence, it fell into oblivion around the turn of the last century from which it still has not fully recovered yet. The number of Bayesian followers is steadily growing with showing many successes over the "orthodox" statistics. The history of Bayesian theory should show that it is by no means a fancy and isolated mathematical tool. It had the potential of becoming what we call statistics now and it might fully return from oblivion one day.

Thomas Bayes

Thomas Bayes, after whom the Bayes theorem was named, was born most probably born in 1701, as offspring of a prominent nonconformist family of Sheffield in the north of England. He studied logic and theology at the University of Edinburgh and became mathematician as well as a Presbyterian minister. In his lifetime he published only two books, one on mathematics and one on theology. In 1742, Thomas Bayes was elected to be a Fellow of the Royal Society.

Bayes famous findings on the problem of inverse probability were published posthumously by a friend, Richard Price, who gave his paper 'Essay Towards Solving a Problem in the

[45] Michalek (2001)
[46] Voß, Gutenschwager (2001), p. 337
[47] De la Vega et al. (2010), p. 3

Doctrine of Chances' to the Royal Society only two years after Bayes' death in 1761[48]. The essay did not contain the statement of Bayes' Theorem that we think of today, which was introduced later by Laplace.

Bayes Theorem

Scientists today are not sure if Thomas Bayes actually discovered the theorem that carries his name, since he might had not fully understood the concept of what we call Bayesian probability today.[49] Nevertheless, Bayes' work pioneered the one of the French mathematician and astronomer Pierre-Simon Laplace (1749 – 1827), who is famous for the Laplace equation, Laplace transformation or the Laplacian differential operator. In an essay from 1974 he pointed out more clearly than Bayes that the posterior distribution of the hypothesis should be proportional to what we now understand as the likelihood of the data:

$$f(H|D_i) \propto f(D_i|H)$$

Equation 2: Proportionality of posterior distribution of hypothesis and likelihood of data[50]

The posterior probability is proportional to the likelihood of the observed data, multiplied by the prior probability. Thus the simplest full version of Bayes' theorem is

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)}$$

Equation 3: Bayes' theorem

where

- P(H) is the prior probability of the hypothesis H
- P(D|H) is the conditional probability of seeing the data D given that hypothesis H is already true
- P(D) is the marginal probability of data D
- P(H|D) is the posterior probability, the plausible model of hypothesis H given that we have observed data D

Two Views on Probability

The *definition of probability* that is taught in nowadays is understood as the long-run relative frequency of occurrence of an event in a sequence of experiments that are repeated a number of times or in equally prepared systems, the "frequentist" view.

The *Bayesian probability theory* (BPT) takes a much more general approach: In BPT, probability is regarded as a real-number-valued measure of the plausibility of a proposition when incomplete knowledge does not allow us to establish its truth or falsehood with certainty. The measure is taken on a scale where 1 represents certainty of the truth of the proposition and 0 represents certain falsehood. Hence it can be seen as a kind of "quantitative epistemology", a numerical encoding of one's state of knowledge. Already Laplace in 1812 viewed probability theory as "common sense reduced to calculation".[51]

---

[48] Bellhouse (2004), p. 23
[49] Stigler (1999), p. 291 – 300
[50] Fienberg (2006), p. 3
[51] Loredo (1990), p. 85

In the late 19th and early 20th century, the mathematicians of those days declined the Bayesian view on probability, but created their own definition as a relative frequency of occurrence. Firstly, seeing probability as a degree of plausibility seemed to be a too vague concept. Secondly, some problems were associated with finding the prior probabilities, which have to be mutually exclusive and selectively exhaustive for a discrete set of set of propositions.[52]

The Bayesian way often seems to be the more "natural" approach to uncertainties: Typical example given by Loredo of plausible inference are the doctor, who diagnoses an illness by considering the plausibility of each of several diseases given the symptoms of the patient; or a judge in court deciding on the guilt or innocence of the defendant in the light of the evidence presented in the trial.

In these examples under the presence of uncertainty, a variety of hypotheses (cold/flu/bronchitis or guilty/innocent) are assessed in the light of a single set of evidence presented and some prior knowledge creating a tendency towards one of the hypotheses. Or in Bayesian terms: Assessing a variety of hypotheses H_i by calculating the posterior probability of each H_i, which depends on both the prior probability of the hypothesis and on the probability of actually observing the specific set of data.

In Bayesian theory, the set of data is fixed and we are looking for which of the many random hypotheses has the highest probability. The frequentist theory, however, has to assume that the data or the set of evidence are random variables, not the hypotheses; despite the fact that the data that we see is the only proven part of the equation. Also, frequentists ignore any prior knowledge that one might already have regarding the hypotheses.

The Revival of Bayes

One of the first in the 20th century to criticize the orthodox or frequentist statistics and to rediscover the ideas of Bayes, Price, Laplace, etc. was the Cambridge geophysicist Sir Harold Jeffreys. In his book 'Theory of Probability'[53] from 1939 he presented Bayesian solutions to many statistical problems, some of them inaccessible to frequentists. In the 1940's and 1950's names like R.T. Cox or E.T. Jayens are connected with research of Bayesian theory.

Since the 1960's the Bayesian minority has been steadily growing especially in the fields of economics and pattern processing. In the 1990's the problem of speech recognition was assessed using the Bayesian technique of Hidden Markov Models as well as the image reconstruction algorithms using the Maximum Entropy theory.[54] Some say, Thomas Bayes ideas were useless without the processing power of modern computers. Therefore, with the success of IT infrastructure, the amount of data is growing exponentially and with it the success of Bayes theory applied in many different ways.

The Beginnings of Bayesian Networks

Judea Pearl, an Israeli American computer scientist and philosopher first developed Bayesian Networks in 1985. In his paper "Bayesian Networks: A Model of Self-Activated Memory for Evidential Reasoning" he argues that the textbook definition of probabilities (frequentist view) delivers a rather distorted picture of the human way of reasoning.[55] Also,

[52] Loredo (1990), p. 88
[53] Jeffreys (1939)
[54] MacKay (1992), p. 2
[55] Pearl (1985), p. 3

the computer powers required to using joint probability distributions $P(x_1,\dots, x_n)$ is simply too high for simple calculations already.

He observed that whereas a person may showed reluctance to giving a numerical estimate for the conditional probability $P(x_i|x_j)$[56], no hesitation would normally be encountered when that person was asked to state merely whether x_i and x_j were dependent or independent, i.e. whether knowing the truth of x_j will or not will alter the belief in x_i.

Therefore, in order to represent causal knowledge by a computational architecture, Pearl prefers to approach this by the Bayesian view on probability. Moreover, he suggests that the fundamental structure of human judgmental knowledge could be represented by dependency graphs and that mental tracing of links in these graphs were responsible for the basic steps of querying and updating that knowledge.

Judea Pearl defined Bayesian networks as directed acyclic graphs in which the nodes represent propositions or variables, the arcs between the nodes signify the existence of direct causal dependencies between the linked propositions and the strengths of these dependencies are quantified by conditional probabilities in the Bayesian sense.[57]

1.2.3 Theory of Bayesian Networks

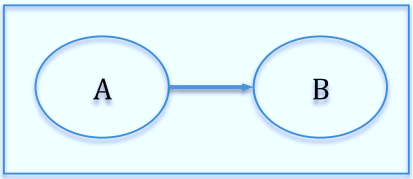
Bayesian Networks (BNs) are widely used for knowledge representation and reasoning under uncertainty in intelligent systems. A BN consists of the following components:[58]

- A set of variables or nodes (A) and a set of directed edges (E) between the variables.
- Each variable either has a finite set of mutually exclusive and collectively exhaustive states, or it is a continuous Gaussian random variable.
- The variables together with the directed edges form a directed acyclic graph (DAG or G), therefore G = (A, E).

To each variable A with parents pa (A), there is a conditional probability table $P(A|pa(A))$ attached. If A has no parents, a table of absolute or prior probabilities is attached.

Figure 6 shows a simple form of a BN, a cause-effect relationship that includes two nodes A and B and an arc pointing from node A to node B. A causal relationship means from A to B means that A, being in some state, influences the state of B. Node A has no parent node, i.e. no edge directing to it. In this case, a table of absolute probabilities, which are considered to be given, represents the variable. Node B is a child of A (i.e. one or more edges pointing at it), therefore it is characterized by conditional probabilities: for each state of A different states of B have to be defined.[59]

[56] Probability of state x_i in case state x_j has already occurred.
[57] Pearl (1988), p. 27 - 50
[58] Jensen (2001), p. 4
[59] Bidyuk, Terent'ev, Gasanov (2005), p. 588



Node	Probability Distribution	State	Conditional Probability
A	$P(a_i)$	a_1	0.5
		a_2	0.5
B	$P(b_i \mid a_1)$ / $P(b_i \mid a_2)$	b_1	1 / 0.6
		b_2	1 / 0.2
		b_3	1 / 0.2

Figure 6: Trivial Bayesian Network[60]

Pearls defined the 'chain rule' theorem in 1988, which represents the basis for the application of BNs:

$$P(U) = \prod_{i=1}^{n} P(A_i | pa(A_i))$$

Equation 4: Chain rule theorem[61]

With U being a set of nodes $U = \{A_1, A_2, A_3, \ldots, A_n\}$ and $pa(A_i)$ being the parents of A_i. The BN specifies a unique joint probability distribution P(U) given by the product of all the conditional probability tables specified in the BN.

A BN is a compact representation of a joint probability distribution of all variables, but it is also a graphical model representing cause-effect relations of a domain and reflecting the inherent uncertainty in a domain.

Not all edges in a Bayesian network necessarily have to represent actual "real-world" cause – effect relations. According to Jensen, probability theorists may accept links not being strictly causal as long as the model represents the proper probability distribution. However, Jensen warns that if the model is used for decisions, it is absolutely critical that the model edges represent causal relationships only.

Simple Example of BN – Wet Grass

The principles of Bayesian Networks (BN) and its most widespread use, probabilistic inference, will be demonstrated in a very simplified way with an everyday example:

[60] Bidyuk, Terent'ev, Gasanov (2005), p. 588
[61] Pearl (1988), p. 114

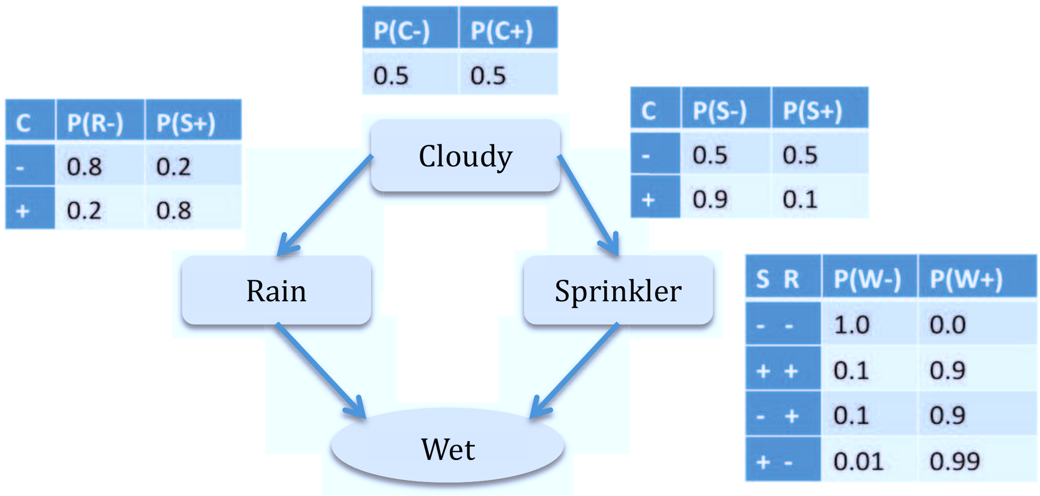


Figure 7: Simple Bayesian network and conditional probability tables[62]

Suppose the grass is wet. There are two possible explanations: The grass can be wet due to rain (R) or due to the sprinkler (S). We use a Bayesian network to find out which of these events has a higher probability.

For simplification, the system is binary; hence all variables can either be true (+ or 1) or false (- or 0) with a certain probability. Both rain and whether the sprinkler is turned on or off is influenced by the clouds. The probability that it is cloudy is 50%. On a cloudy day it is more likely to rain. This is shown in the conditional probability table: if it's cloudy there is an 80 % chance of rain. If there are clouds, I will not turn on the sprinkler in 90 % of the cases. The wet grass depends on two variables: if there is no rain and no sprinkler, the grass won't be wet for sure, but if I have both, it will be wet a 100 %.

Using Bayes rule we obtain:

$$P(S=1|W=1) = \frac{P(S=1,W=1)}{P(W=1)} = \frac{\sum_{c,r} P(C=c,R=r,S=1,W=1)}{P(W=1)}$$

$$P(R=1|W=1) = \frac{P(R=1,W=1)}{P(W=1)} = \frac{\sum_{c,s} P(C=c,R=1,S=s,W=1)}{P(W=1)}$$

with $P(W=1) = \sum_{c,s,r} P(C=c,R=r,S=s,W=1)$,

Equation 5: Probabilities of sprinkler and rain obtained with Bayes' rule.[63]

where the small captions c, s, r represents the different states of cloudy, sprinkler and rain.

Types of Bayesian Networks

In *discrete* BNs (like in the example given above) the nodes represent events that are described by discrete variables. This means that random variables can have a set of stages. Nodes with parents are defined by a table or a function of conditional probabilities; nodes without parents include unconditional (absolute) probabilities. In practice, the parameters are often binary (true / false) or have only a few states (e.g. low / moderate / high). Ar-

[62] Bidyuk, Terent'ev, Gasanov (2005), p. 591
[63] Bidyuk, Terent'ev, Gasanov (2005), p. 588

ranging values in a limited number of states results in a loss of accuracy, but it may simplifies the data input procedure.

In *continuous* BNs, the nodes represent events that are described by continuous variables. Events can take any state from an infinite point set. Continuous random variables are described by probability distribution functions or probability distribution densities.

Hybrid BNs contain both nodes with discrete and nodes with continuous probability distributions. In this case, restrictions apply for such a type of BN, since discrete nodes cannot have continuous parents. For large networks, this form of BN will be the most common one, since they will include a wide range of possible input parameters.

Dynamic Bayesian networks (*DBN*s), in contrast to *static* BNs, are networks with time – varying nodes. The structure of the model does not vary, but the nodes at time t can influence nodes at time t+n. Figure 8 shows a dynamic Bayesian network in its unrolled form. Variables at the same node may differ from time to time and influence different nodes (hence the blue arcs change), whereas the static network structure (nodes and grey arcs) remains the same. The network can be 'unrolled' further and further, for each time step time 0, time 1, time 2,…, time n new variables can be defined.

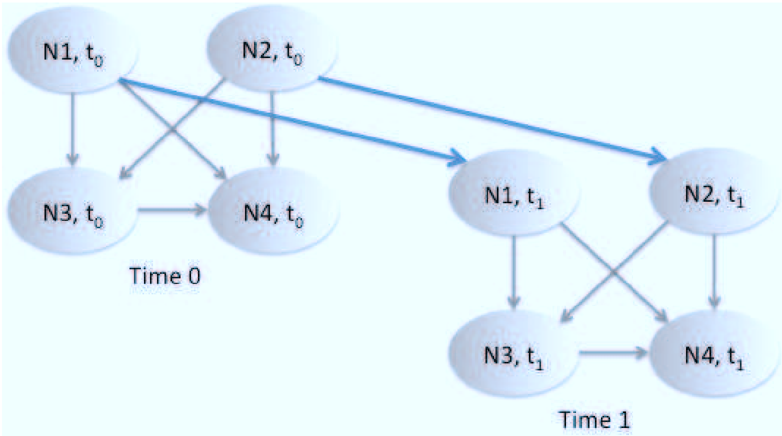


Figure 8: Dynamic Bayesian network

An example of a DBN is the *Hidden Markov model,* a tool for modelling time series data. Hidden Markov models are used in speech recognition systems and other areas of artificial intelligence and pattern recognition. The name descends from two parts. First, the basic assumption is that the observation at time t was caused by an event whose state was *hidden* from the observer. The second part of the name comes from the application of the *Markov property* that states that given value of a state at time t-1, the current state at time t is independent from all states prior to time t-1. Therefore, all you need to know for the "history" of states is only the previous one.[64]

Dynamic Bayesian networks allow a prediction of the future events and probability distributions. Algorithms, such as Kalman filtering have been developed for that purpose. This can be a very helpful tool in operation systems, for instance to predict future failures of tools or materials and adapting a preventive action approach.

In *naïve* BNs, all input parameters are parented by a single output parameter, called the target node. Therefore, given the parent node, all child nodes are independent from each other. In order to overcome the limitations in cohesion with the simple network structure, the

[64] Ghahramani (2001), p. 2

naïve BN can be extended to an *augmented* naïve BN, by creating dependencies between the child nodes. Figure 9 schematically shows a naïve and an augmented naïve BN.[65]

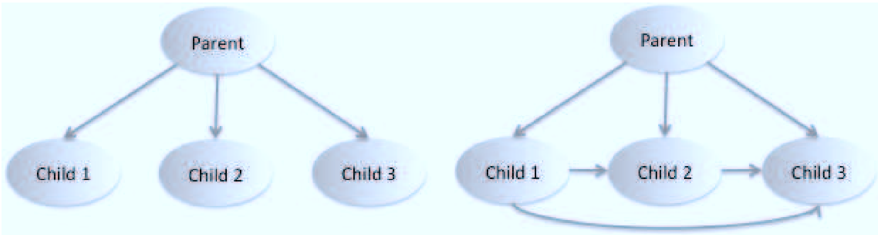


Figure 9: Examples for naïve and naïve augmented Bayesian networks.[66]

BN Extended by Influence Diagrams

Decision trees as a graphical representation of uncertainties (probabilities), decisions and payoffs (equivalent utilities) are used in a wide range of applications. On a small scale, decision trees are intuitive and easy to handle, but for larger, more complex systems, IDs can provide a better level of insight and transparency.[67]

The framework for (multi-criteria) decision-making based on BN includes the extension of BNs by influence diagrams IDs. IDs extend BNs by adding utility nodes and decision nodes, which allows to deploy Bayesian networks in an even greater variety of tasks, including the computation of the expected utility, given observations and decision choices, thus finding the optimal decision.

In an ID, let A = $\{a_1, a_2, \ldots, a_n\}$ be a set of mutually excusive actions, and *H* the set of determining variables. A utility table U(A, H) shows the effects on H for each of the actions A taken. The problem is solved by maximizing the expected utility for action a EU(a) by multiplying the utility of H for action a U(a, H) with the conditional probability of H in case action a is taken:

$$EU(a) = \prod_{H} U(a, H) P(H|a)$$

Equation 6: Expected utility[68]

1.2.4 Learning of a Bayesian Network

Learning is seen as an increase of knowledge; knowledge in the context of Bayesian networks is the structure of the network, representing knowledge about the causal relationships of variables as well as the conditional probabilities of each variable from data, expert knowledge or inference. In the context of Bayesian theory there are three types of learning: parameter learning, inferential sensing if parameters are unobservable and learning of structure.[69]

[65] Friedman, Geiger, Goldszmidt (1997), p. 10
[66] This is a figure created by the author; if not indicated otherwise, all other figures without a foodnote are created by the author.
[67] Giese, Bratvold (2010), p. 3
[68] Adapted from Jensen, Nielsen (2001), p. 3
[69] Woolf (2001)

Parameter Learning

The primary application of Bayesian Networks is parameter learning or belief updating.[70] Setting up a Bayesian, we start with some *a priori* knowledge about the model structure and the probability distributions of the nodes. From this first model that is then updated with data we get the *posterior* probability distribution of node variables.

Equation 7 shows the Bayesian theorem augmented by adding background information I available that indicates how hypothesis H and data D are related.

$$P(H|DI) = \frac{P(D|HI)P(H|I)}{P(D|I)}$$

Equation 7: Bayesian theorem augmented by background information[71]

Again, the "after data" or posterior probability of H can be calculated by multiplying the data we already had or prior probability P(H|I) by the probability of the data assuming the hypothesis is true P(D|HI) and normalizing it by P(D|I), which is the probability that we get the data regardless of the hypothesis. Therefore belief updating means adding evidence to the actual BN model and calculating the posterior probability distribution for the new set of variables. This means that our state of knowledge changes through the acquisition of data; the model learns.

Inferring Unobserved Variables

Bayesian networks can be used to gain information on a subset of variables when other variables (the evidence variables) are observed. Probabilistic inference is the process of computing the posterior distribution of variables, given an evidence.

Exact and approximate interference algorithms have been developed. Exact inference methods include the variable elimination, clique tree propagation or recursive conditioning. The complexity of all these methods is exponential to the width of the network tree. The most common approximate inference algorithms are stochastic MCMC (Markov Chain Monte Carlo) simulation or the mini - bucket elimination.[72]

Structure Learning

The task of updating the network structure or deriving a correct network structure from data is more complex than deriving node parameters only.[73] But in special cases it might be beyond human capacity to define the network structure. In this case, algorithms like the structural maximization of the expectation (SME) or the bound and collapse algorithm are used. This approach is rather complicated in most cases, since the size of the learning dataset required is unrealistically large for most practical learning domains.[74]

---

[70] Jensen (2007), 109 - 157
[71] Modified from Lored (1990), p.86
[72] Vanek (2011), p. 5 - 6
[73] Bidyuk, Terent'ev, Gasanov (2005), p. 596
[74] Mitchell (2010), p. 3

1.2.5 Alternative Methods of AI

Artificial Neural Networks

(Artificial) neural networks (ANN) are information processing mathematical or computational models that are inspired by the structure and functional aspects of biological neural networks. They consist of numerous simple simultaneously working, interconnected groups called neurons.

By basing neural networks on biological systems, engineers hope to make use of the following advantages:[75]

- Biological information processing is robust and fault-tolerant; it includes a lot of redundancy that allows loosing some routes every day or signals to be wrong without problems.
- Biological information processors are flexible; learning helps the brain adapting to new environments without a total change of structure (i.e. reprogramming).
- Brains can handle fuzzy, noisy and inconsistent data, which in conventional programs requires a lot of programming and complex algorithms.
- Brains work in a highly parallel way; they are small and compact and need little power.

ANNs typically include three different layers. The first layer consists of a set of input neurons, were input data enters the network and propagates through the network across the weighted connection until the activation reaches the output layer. The middle layer is known as hidden layer as it is invisible from outside and we in most cases not affect its activation directly. The goal of an AAN is to discover some association between input and output patterns.

In Figure 10 a typical ANN is shown, in this case it consists of 3 layers of neurons and 2 connecting layers of weights.

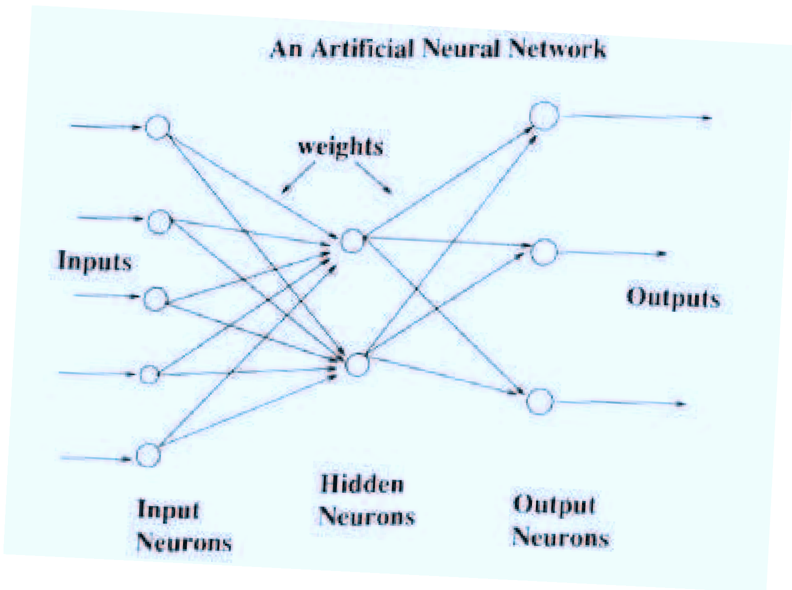


Figure 10: A typical artificial neural network[76]

[75] Fyfe (2005), p. 57,58

Each neuron in an ANN, just as real neurons, only fires (i.e. allows propagation of a signal) if a certain threshold value is reached. Firing (or activation) rules are programmed that determine whether a neuron should fire for any input pattern. These firing rules are the key to the flexibility, since they are adjustable by the network itself.

Also each connection between the neurons is *weighted*. This means that input values can have more or less influence on whether an output value is activated at a neuron or not.

When an ANN is set up, it is in the learning mode, which can be supervised or unsupervised. In the *supervised learning mode*, input and output data is provided and the network is corrected if the results differ from the required ones. In the *unsupervised learning mode*, only the input parameters are provided and the network has to be self-organized (i.e. to teach itself) depending on parameters from the input set only.

Comparison ANN and BN

Correa[77] and his colleges compared the use of artificial neural networks with Bayesian networks for quality detection in a machining process and found several advantages in favor of the latter.

First of all, BNs achieve the best results in terms of classifier goodness applied to the problem of quality prediction in high-speed milling processes. The results have been confirmed by several hypothesis tests. BNs were significantly faster than ANNs both in time it took to enter the data and construct the models and for the computers to calculate them. ANNs have the problem of non-convergence to a global minimum, when being trapped in a local minimum.

There are no general construction principles for the network structure (e.g. number of nodes, hidden layers, activation functions), whereas a BN is based on actual causal relationships, a feature that again makes it easier for the constructors and users to understand the Bayesian network. As a result the understanding of its working principles allows everybody to optimize its structure and increases the confidence in the correctness of the model. In contrary, ANNs work like a black box. They do not allow to perform inference, which can be helpful to determine the effect of different variables. As an advantage of ANNs over BNs it can be mentioned that ANNs take less memory space.

Fuzzy Systems

Fuzzy systems are suitable for uncertain or approximate reasoning, especially for a system with a mathematical model that is difficult to derive. This allows solutions despite estimated values under incomplete or uncertain information.

In 1965 Lotfi A. Zadeh, who also created the term "soft computing", developed *fuzzy sets* to be able to mathematically represent uncertainties and vagueness in data and information and for dealing with imprecision intrinsic to many problems. [78]

However, the story of fuzzy logic started much earlier: the Greek philosopher *Aristotle* already posited in one of his so called "Laws of Thought" the "Law of the Excluded Middle", which states that every position must be either *true* or *false*. Around 1920, Lukasievicz proposed a systematic alternative to the bi-valued logic of Aristotle, a three-valued logic: The third value could be seen as "possible" and it had a numeric value between true and

[76] Fyfe (2005)
[77] Correa, Bielza, Pamies-Teixeira (2007)
[78] Zadeh (1965)

false. The concept was further developed including more and more values, until Zadeh described the mathematics of his fuzzy set theory. This theory proposed that the values of true and false should operate over a full infinite range from 0 to 1.[79]

A fuzzy set therefore can be defined by its membership function:

$$\mu_A : X \rightarrow [0,1]$$

Equation 8: Fuzzy membership function[80]

with X being a nonempty set, A a fuzzy set in X and μ_A the degree of membership of element x in A for each $x \in X$.

Some of the essential characteristics of fuzzy logic relate to the following:[81]

- In fuzzy logic, exact reasoning is viewed as a limiting case of approximate reasoning.
- In fuzzy logic, everything is a matter of degree.
- In fuzzy logic, knowledge is interpreted a collection of elastic or, equivalently, fuzzy constraint on a collection of variables.
- Inference is viewed as a process of propagation of elastic constraints.
- Any logical system can be fuzzified.

Frequentists have a bi-valued view on events, which are only either true or false. But fuzzy probability expresses a state of knowledge: under given evidence and current state of knowledge an event has a percentage of certainty of occurrence.

In fuzzy logic, for example, "a person whose height is 2.5 meters is tall", is a true statement. Another statement. "A person whose height is only 2 meters is tall" is also a true statement, but it is not as true as the first one.[82]

The typical process flow of a fuzzy system is shown in Figure 11.

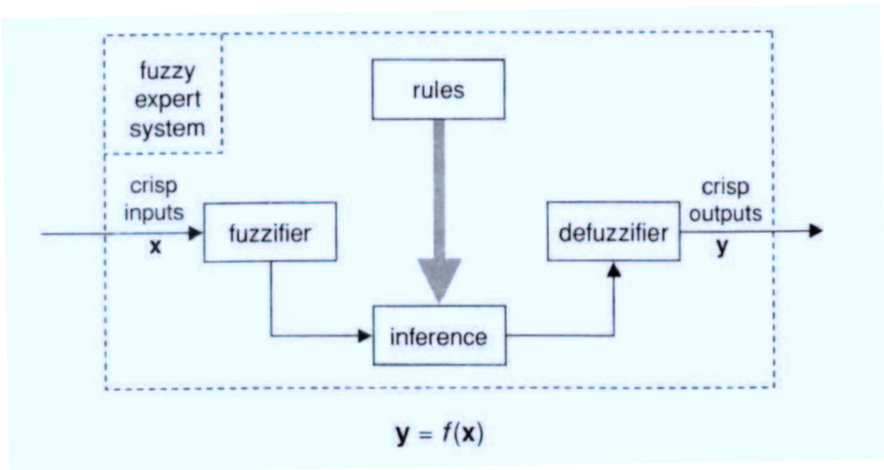


Figure 11: Generic architecture of a fuzzy expert system[83]

[79] Zadeh (1965), p. 339
[80] Fuller (2005), p. 11
[81] Fuller (2005), p. 11
[82] MarkCC (2011)
[83] Garibaldi (2005)

Although fuzzy networks for certain applications show more accurate results than BBNs, they are comparatively easy to construct and can deal with even less accurate data, the biggest downturn of fuzzy networks is, despite its name, the lack of flexibility and adaptability when it comes to new data and results. Whereas BNs learn and update their structure automatically, the rules of fuzzy systems have to be redefined by the system expert.

Other Intelligent Methods

In a survey[84] by the SPE Communities - Digital Energy Technical Section Artificial Intelligence and Petroleum Analytics (AIPA), released in October 2011, several methods that are generally connected with AI in practical work, were listed by the SPE research group. Besides the in detail described ANNs, fuzzy logic and expert systems they list: data mining, rule-based reasoning, genetic algorithms, machine learning, intelligent agents, automatic process control, proxy models, workflow automation and virtual environments. A more detailed description by the AIPA can be found in the Appendix.

Workflow automation: is a set of methodologies and technologies which aims to integrate data and applications into automated workflows, which reflects the business processes developed in a company over a well structured information management platform. Workflow automation has been one of the main areas of interest of the Oil & Gas Industry in the last 5 years, since is one of the key elements of the Digital Oil Field and Integrated Production Operations trends.

Virtual environments: is the combination of simulation, computation and visualization technologies in order to reach partial or total immersive environments for the analysis of production and reservoir data.

Some other methods that are in use in connection with decision support systems are described below.

Petri Nets (PN) are a graphical tool for the formal description of the flow of activities in complex systems. In particular, petri nets represent the logical interaction among parts and activities in a system. A PN is also called place/transition net and consists of place nodes transition nodes and directed arcs. A typical application example would be the modelling of a concurrent behaviour of a distributed system like the description of chemical processes, where molecules interact if they meet at the same place and at the same time.[85]

Clustering or *cluster analysis* is the unsupervised classification of patterns (observations, data items, or feature vectors) into groups (clusters). Clustering or cluster analysis one of the most important unsupervised learning techniques. Clustering is useful in several exploratory pattern-analysis, grouping, decision-making, and machine-learning situations, including data mining, document retrieval, image segmentation, and pattern classification.[86]

Self-organizing maps (SOMs) are a type of artificial neural networks, capable of unsupervised learning, that display multidimensional datasets in a low dimensional (typical two-dimensional) map. SOMs use the Euclidian distance between two points in the multi-

[84] AIPA (2011)

[85] In fact, some sources say that the developer of Petri Nets, Carl Adam Petri, invented them in 1939, at the age of only 13, for exactly that purpose of describing chemical processes.

[86] Jain, Murty, Flynn (1999), p. 1

dimensional space as measure of their similarity.[87] This makes SOMs useful for visualizing low-dimensional views of high-dimensional data.

Conclusion

In the last chapter we looked at a selection of intelligent methods that are commonly in use for expert systems or decision support systems. Comparing these methods we can conclude that Bayesian networks are the best suitable methods for classifying problems and giving decision support in oil and gas operations.

Petroleum engineering has always had to deal with great uncertainties. These uncertainties are eliminated by calculating the node state probabilities based on information about the values of the other nodes of the network. For the build-up, the combination of expert knowledge and automated structure and probability inference promises an optimal design. Their transparent structures allow inferring roots of problems and influences on utilities and they promise the biggest acceptance by the user. The results are relatively easy and fast to calculate and more importantly, they are better in terms of accuracy than for other methods. Also, Bayesian networks have the ability to learn in multiple ways, which is crucial for a smart and adaptive system.

1.2.6 Applications of Bayesian Networks

Decision support and expert systems are used in various applications and industries. In particular, BNs and other forms of AI are used in systems, where modelling of knowledge under uncertainties is required, such as computational biology, bioinformatics (gene regulatory networks, protein structure), modelling of eco- or climate-systems, document classification, information retrieval, image processing, data fusion, various decision support systems, engineering application, gambling or law.

A very prominent example for one of the first applications for decision support is medical diagnoses. Other popular application examples are in engineering areas, in the financial sector including risk management and insurance, monitoring and alerting, sensor fusion, etc.

Medical Diagnosis

A recent development in medical diagnosis is *evidence-based medicine* (EBM). EBM requires the integration of individual clinical expertise, available external scientific evidence, where preferences, desires and expectation of the patient should be central to the decision making process. The hopes are put on medical informatics in general and artificial intelligence in particular to help improve the healthcare quality.

Figure 12 shows an influence diagram for the treatment of patients. The disease is the hypothesis variable, the shaded circles are evidence variable or measurable findings and the decision nodes are represented by rectangles. The final utility "U" may show the satisfaction of the patient. Based on a symptom, the doctor may choose whether or not to make certain tests and finally decide on a treatment.

[87] De la Vega (2010), p. 4

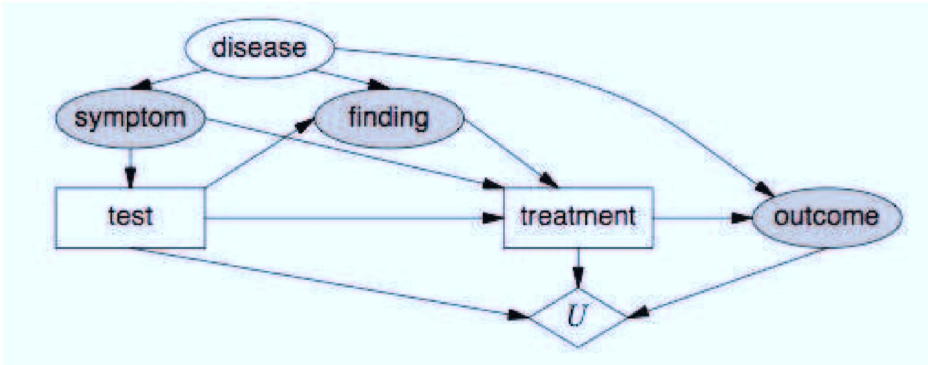


Figure 12: Influence diagram for medical treatment[88]

Financial Applications

Financial risk management, portfolio allocation, banking, insurance, etc. traditionally include a lot of uncertainties. Many causal relationships at capital markets are yet unknown, leading to unexpected developments and over-reactions at stock markets. Figure 13 shows a schematic BN for a stock market application, including conditional probability tables. A real life example would be much more complex.

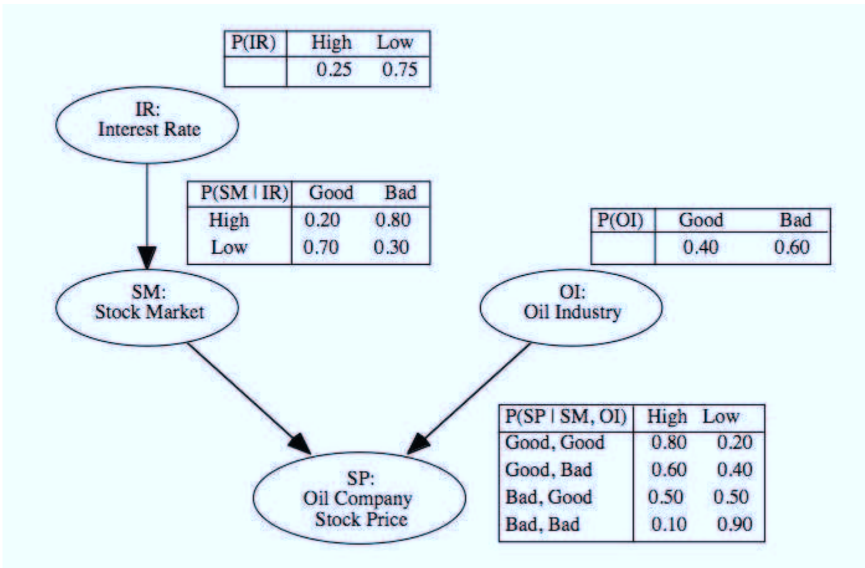


Figure 13: Bayesian network for stock price with conditional probability tables[89]

Sensor Fusion

If various sources must be integrated to come to interpretation of a situation, this class of problem refers to sensor fusion. Examples for this are the integration of several cameras from different angles or resolutions into one picture; or in industrial applications, sensors often redundant and signals are noisy, thus an integration of all signals is required to receive the most exact reproduction of the situation. Whereas each sensor has only a limited chance of giving a correct interpretation, the combination of all the sensors' chances, typi-

[88] Van Gerven (2007), p. 21
[89] Shenoy (1999), p. 4

cally increases the likelihood of a valid interpretation. BNs are can very well deal with missing data and combine information accurately.[90]

Monitoring and Alerting

Eric Horvitz developed the *Vista system*[91], a computational technique that uses a Bayesian network to reduce the cognitive load of human operators responsible for complex monitoring systems in the NASA space shuttles. It has been successfully used by the NASA Mission Control center in Houston for several years.

Figure 14 shows the key components of the Vista system. A belief network is used to consider the diagnostic relevance of shuttle telemetry. A model of time-dependent utility is employed to prioritize alternative possible faults, given the likelihood of the faults, and to control the size and amount of detail dedicated to a panel describing the system in which the possible faults may be occurring.

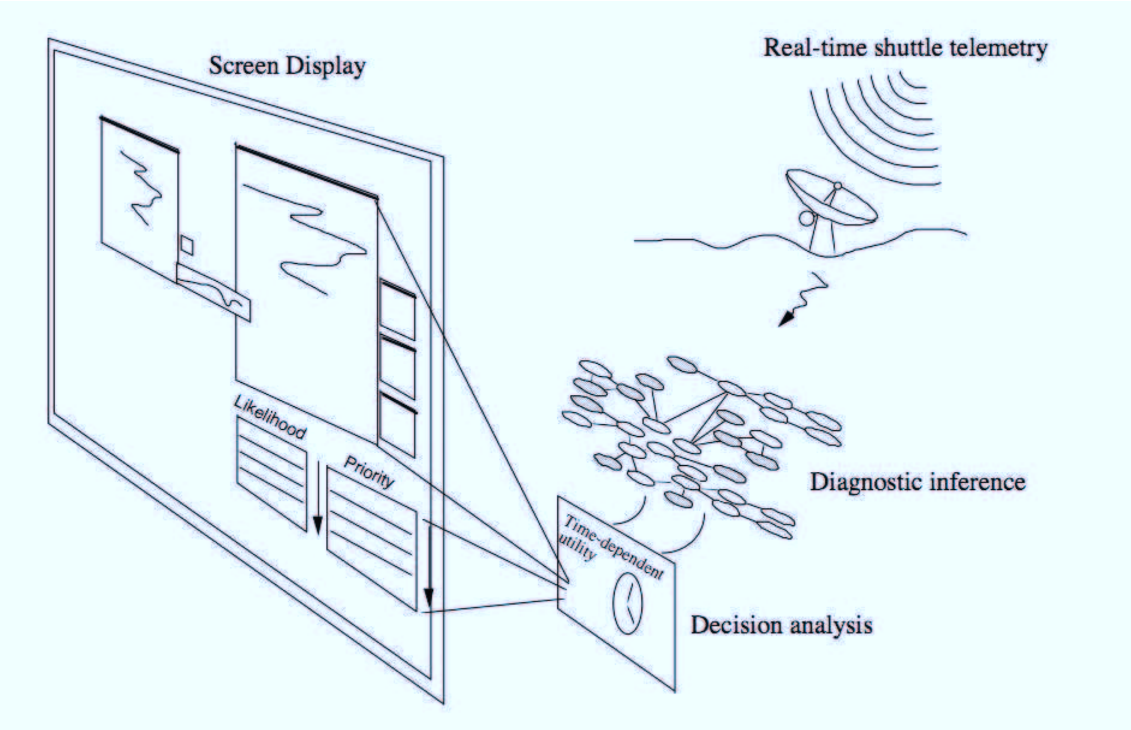


Figure 14: Vista system overview[92]

The display includes a problem summery, and a list of possible disorders shown in a column listed by likelihood of the cause. In another column, the disorders are listed in terms of priority, which is the expected time-criticality of the problems. The expected time-criticality is computed by weighting a relative cost of delay for each possible disorder, assessed from the ground controller, by the likelihood computed for that disorder.[93]

---

[90] Norsys (2011)
[91] Horvitz (1992)
[92] Horvitz (1992), p.10
[93] Horvitz (1992), p. 8

1.3 Oil & Gas Operations

The oil and gas industry in general and particularly the E&P business is traditionally characterized by huge cost and revenue figures with high risks, and high margins. Provocatively speaking, one may find evidence that lean management, target costing, etc. are as unfamiliar in the oil industry as an attitude of watching every penny considering the millions and billions spent on drilling wells or building production plants. Time is money: Offshore platforms can cost several hundred thousand dollars every single day. So decisions have to be made fast in the first place, irrespective of the cost of the outcome.

Oil and gas operations are also characterized by their unrepeatability. No reservoir in the world shows the same properties, therefore, none of the numerous wells drilled so far was the same or no production history of a field is exactly comparable. In Formula 1, for example, the car can around the racetrack as often as desired. This offers the possibility to test for every single variable, by changing it and keeping all others constant. Results are shown within minutes.

These issues lead to difficulties for the application of decision support systems in oil and gas operations because data samples and testing data is not easily produced on demand. On the other hand, it is even more important, to intelligently capture processes, extract knowledge in form of causal relationships and probabilities and apply this knowledge to other assets where there is a lack of relevant data.

Another consequence of the large uncertainties is that accuracy and amount of measurements are naturally lower than in repetitive and exact industries like automotive or other consumer goods. With the drop in oil price in 2008 as a consequence of the global economic financial crisis, oil companies suddenly faced troubles financing their operations. Exxon Mobil, a company that was described as very risk averse, cautious and with high time and budget discipline, was better off than its competitors with shares falling just by 15 % and 22 % respectively.[94] This could hint to a relation of accuracy or transparency and success.

Catastrophes like the oil spill in the Golf of Mexico in 2010 cast a damp over the reputations of oil companies as well as their balance sheets. Although risks and uncertainties are omni-present in oil and gas operations, a smart handling of data, improvements in data quality, transparency, and hence better knowledge management could bring the right data, to the right people at the right time and help to prevent mistakes and make fast and right decisions.

1.3.1 Well Performance Monitoring

The primary design for the Adaptive Advisor described in the Chapter 3, is for well performance monitoring, which generally includes well surveillance, troubleshooting and production optimization. The following issues add to the complication of the task with regards to an improvement of the system and at the same time call upon the necessity of smarter systems:

- Many different disparate well and process data sources exist within an asset. The various databases are operated with different software that is seldom compatible.
    - There are high frequency *real – time data sources*, gathered directly at the control system that operates the production facilities. Production volumes of

[94] LeVine (2009)

oil, gas or water, temperature or pressure data are examples for this. They consist of a data pair of measurement and its associated timestamp.
   - *Historical* or *tactical* data sources consist of a wide range of different databases, they include manually entered laboratory measurements, time aggregates like daily production, reports on maintenance, etc. Also this includes non-time based data such as well planning data (e.g. well trajectory), or reservoir properties. The data exists in form tables, figures, documents, texts, orally communicated, etc.
- In many companies, well monitoring activities are *not centrally organized,* but differ between assets, which reduces the comparability, makes central support and inter-asset communication difficult and introduces a large learning curve when staff moves between assets.
- Particularly real time data leads to *data overload* and data *quality problems*. Because many different sources of data exist, data overlap and data duplications between systems often occur, leading to confusions and maybe wrongly reported data as a result.
- Well Performance Monitoring often is not a *specific job position*, but is carried out by personnel as one of many other tasks. Therefore not more time than necessary is spent on this task and the creation of a feeling of responsibility for a proactive improvement of the system is difficult.

1.3.2 The Big Crew Change

Another issue that particularly concerns the oil industry is its aging personnel: The late-2000s, with an oil price as low as 12 $ per barrel, the IOCs had to decrease recruitment quite drastically and the oil industry was not quite an attractive field for students. Not enough fresh engineers could replace the retiring experts. As a result, the oil industry faces the challenge of a very old average workforce. Figure 15 shows the percentage of petrotechnical professionals in a certain age group in 2009 and the prospection in 2015. Especially the age group of 50-54 will experience high retirement rates in the next years.

Schlumberger Business Consulting[95] investigated the role of petrotechnical professionals for success and growth in the oil business. A study displays that technical talent plays a strategically important role in the oil and gas business. Also it confirms the large demographical shift, which will materially reduce the number of petrotechnical professionals within the coming years.

[95] Rostand (2011)

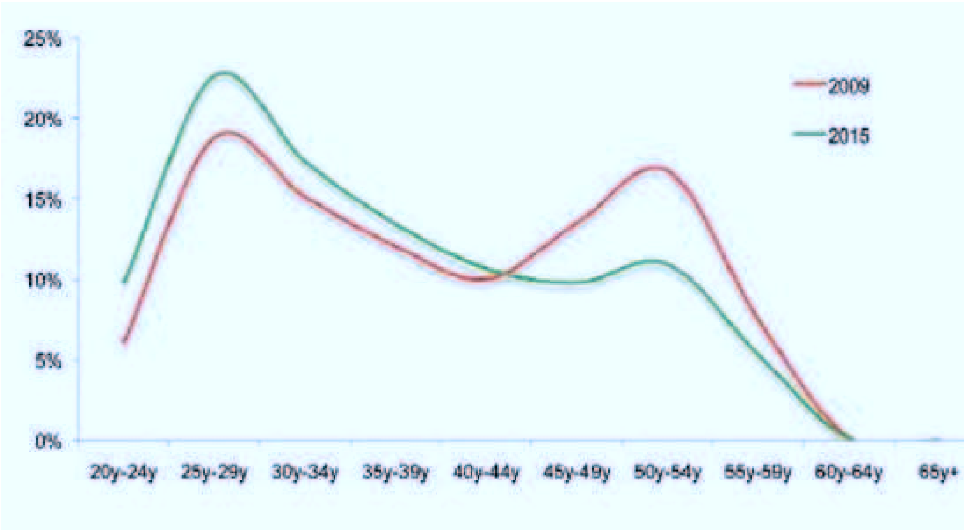


Figure 15: The big crew change[96]

What is the impact of "the big crew change"? Every day, experienced oil and gas professionals leave companies and irretrievably take with them their enormous knowledge and decades of experience, without leaving neither sufficiently trained replacements nor explicit records of what they know and used to do.

Although the number of young engineers increased within the last years, the gap in the middle-aged workforce is challenging. Oil companies need to react fast, if they want to capture as much knowledge as possible. An Advisory System requires an expert to set it up, who has an in-depth understanding of causal relationships and a feeling for probabilities of occurrence.

Once a system is in place and has been used successfully for a while, it can assist young engineers to fulfil tasks that would normally be beyond their capabilities. As the times of "cheap oil" are over, projects become less profitable but much more complex at the same time. The workload and responsibility of a single engineer therefore increases; a circumstance that calls for the help of intelligent KT systems.

1.3.3 Digital Oilfield

Efforts to employ smarter systems or knowledge technology are connected with the expression of *digital oilfield*. A digital oilfield (some may call it integrated operations (IO), smart field, i-field or intelligent energy (iE)) is about specially adapting and designing IT systems for the petroleum business to meet the objective of improved production and increase efficiency of people, technology and knowledge in order to make better decisions faster. The Cambridge Energy Research Association (CERA) provides another commonly used definition:

'The vision for the Digital Oil Field is one where operators, partners, and service companies seek to take advantage of improved data and knowledge management, enhanced analytical tools, real-time systems, and more efficient business processes.'[97]

Business Intelligence (BI) has found its way into the E&P industry. Digital oilfield can be seen as the application of business intelligence or operational business intelligence. The pioneers

[96] Rostand (2011)
[97] CERA (2003), p. 2

of BI were data-intensive industries that digitalized tools and processes to analyze large amounts of transactional or financial information; digging into what customers are buying or analyzing spend in a particular business unit.[98]

BI is the combination of practices, capabilities and technologies that companies use to gather and integrate data, apply business rules and deliver visibility to information in order to better understand the business and ultimately improve performance.

Operational BI is the application of business intelligence capabilities within operational areas of the business that typically involve information and data that changes frequently during the business day.[99]

Mike Brulé formulated 5 stages of a company's Operational BI development:[100] Stage 1 is the reporting what has happened – standard in the E&P industry. Stage 2 is analyzing why things happened, including ad hoc queries and KPI. In stage 3 companies are predicting what will happen in the future using analytical modeling, full – physics models or integrated asset modeling. Stage 4 involves the operationalization of the status quo: continuous update, time sensitive queries and in-database analytics on huge amounts of multi-subject data. In stage 5, real-time decision-making is happening.

According to Brulé, the E&P industry is somewhere between the first and the second stage, but manages to reach stage 5 for certain isolated production areas like gas – lift optimization.

Too Much Data

For oil and gas operations, the number of collected data points has vastly increased within the last years. In drilling, for instance, operators have increased the installation of downhole and surface sensors to collect data, included high-speed data links between downhole and surface instruments and increased telecommunication capabilities to share data at unprecedented speeds and quantities with service companies.[101]

Also, petrotechnical professionals often mistake a higher rate of data and information flow with automatically better decisions. Research across various disciplines has shown that more information only helps to the extent that it can be used intelligently.[102]

Figure 16 displays the concept of information overload. There exists a positive correlation with the amount of information received on an issue prior to taking a decision and the quality of the final decision. When the amount of information provided reaches a certain point we speak of informational overload. Beyond this point, information undermines decision quality.[103]

---

[98] Kuzmich (2009)
[99] Hatch (2009), p.5
[100] Brulé (2009), p. 24
[101] Rajaieyamchee (2010), SPE 135416, p. 1
[102] Meyer (1997)
[103] Rajaieyamchee (2010), SPE 135416, p. 2

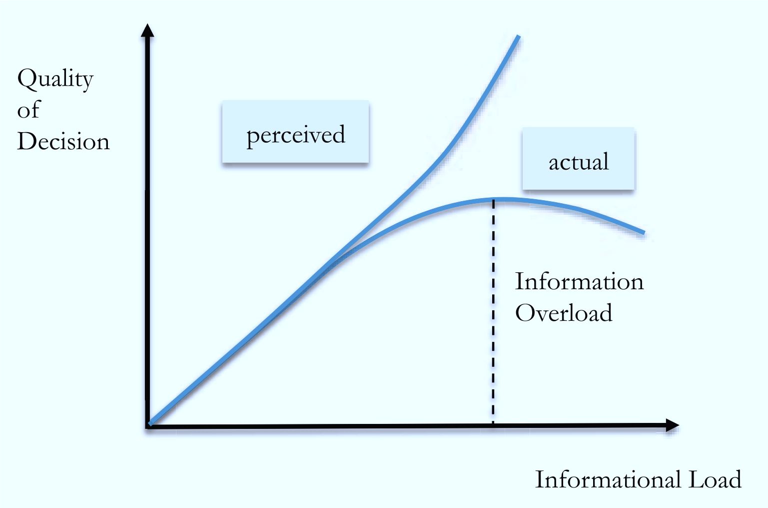


Figure 16: Data / information overload[104]

Therefore it is important to emphasize that digital oilfield does not mean to primarily digitalize operations and measure what is going on. It is all about integrating operations such that collaboration across disciplines, companies & organizational and geographical boundaries, made possible by real time data and new work processes in order to reach better decisions - faster.[105]

Benefits of KT in Oil and Gas Operations

Operational business intelligence for oilfields seems inevitable. Loss of knowledge with more and more experts retiring and difficulties finding fresh engineers call for smarter systems that are capable of capturing the knowledge in a company. Some social psychologists[106] regard intuition as a consequence of unconscious information processing. Methods employing artificial intelligence could to some extent imitate the "gut feeling" when experienced engineers make decision under limited amount of time. The subconscious knowledge is also a result of previous experiences and extraction of information and data.

The implementation of business intelligence in oil companies that are generally known for their antique, cowboy-style, change-resistive way of working would lead to better decisions in less amount of time – a tremendous competitive advantage. High profits in oil and gas operations can be accomplished by high revenues and low cost. Early identification of under-performing wells or need for artificial lift or workover, spending less time for problem solving and fast actions and pro-action instead of re-action, help increasing the production at every point in time. Processes are recorded at every step of a project, which provides a basis for defining and calculating KPIs for strategic and operational goals.

Avoiding wrong decisions and mitigating risks in the oil business that has a long history of major accidents with large international impact, offers a huge cost savings potential. Business intelligence improves transparency and helps companies to understand what they are doing, but also provides more accurate information to auditors. Just like every alarm requires an action, each business process that is not concluded with a decision is a waste of a business opportunity.

---

[104] Rajaieyamchee (2010), SPE 135416, p. 2
[105] Strasunskas (2010), p. 1
[106] for example see Dijksterhuis, Nordgren (2006)

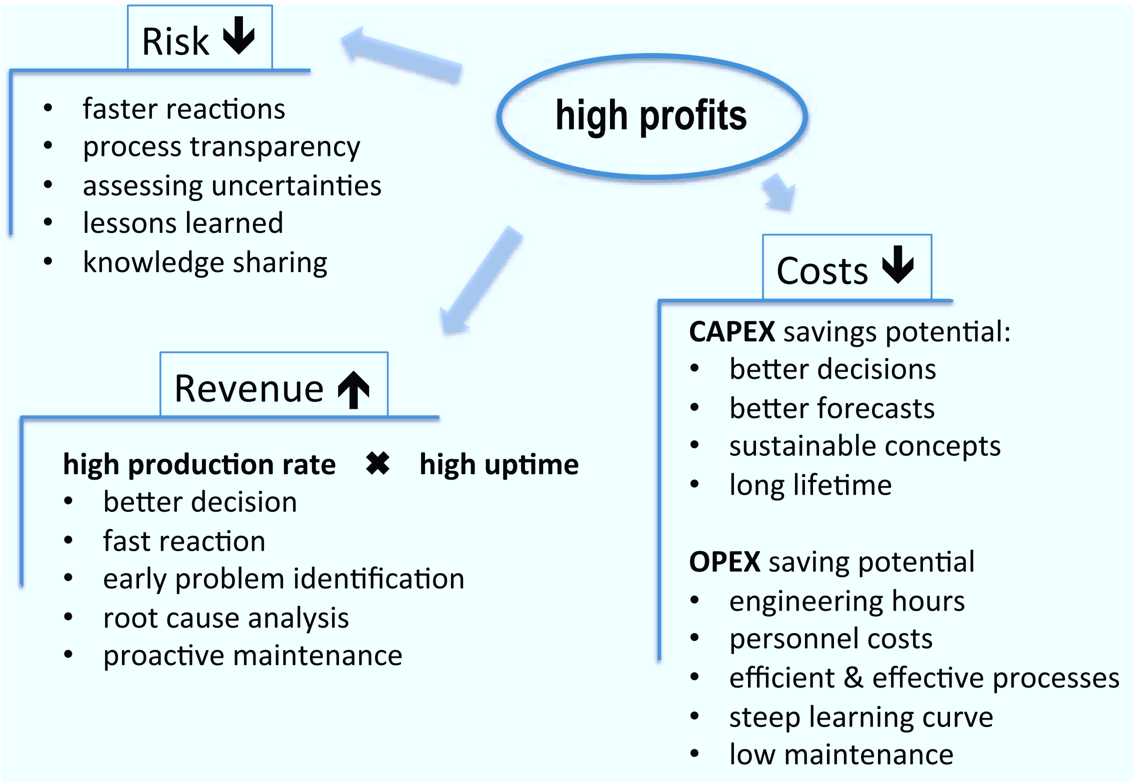


Figure 17: Benefits of an Advisory System in O&G operations

1.3.4 Intelligent Methods in the Petroleum Industry

There are a few examples of the usage of Bayesian networks in petroleum related applications. Lately an increase of the number of SPE paper publications can be noticed, which may indicate an increased effort of research on Bayesian networks for the petroleum industry.[107] The papers so far mostly deal with rather isolated specifically technical issues, no high level approach for multiple operational petroleum related problems (e.g. for field management) can be found in professional publications (SPE) yet. Below some examples are described more detailed.

Predicting Performance of Gel Treated Wells

Ghoraishy[108] and his colleges from the University of Kansas used Bayesian networks to predict the performance of Gel-Treated Wells in the Arbuckle formation in 2008. They used a data set of pre- and post-treatment data of 45 wells to train both naïve and augmented naïve BNs to predict the performance of 14 wells where pre-treatment data was available. Later the performance of the BNs was evaluated in terms of discounted cash flow and net present value (NPV).

Attempts of predicting the performance of wells treated with polymer gels with multivariate statistical analysis[109] or identifying candidate wells with neural networks[110] were unsuccessful, due to the complexity of the gel-treatment process and its physical nature.

[107] Approximately 12 SPE papers that explicitly deal with Bayesian networks have been published, 4 of them in 2011.
[108] Ghoraishy (2008)
[109] Alhajeri (2006)
[110] Saeedi (2006)

The training data set was used to train the supervised naïve as well as augmented naïve BN. The accuracy of prediction was measured by the error fraction, which is the absolute difference of the actual and the predicted output, divided by the actual output. The naïve and the augmented naïve BN reached an accuracy of 77% and 84 % respectively, showing the positive effects of an increase in complexity of the network structure.

Expert System for Optimum Completion

Al-Yami[111] et al. were among the first to propose a software system that depends mainly on previous knowledge and experience, which is used as a guide for optimal completion design using Bayesian Artificial Intelligence. Being aware of a lack of knowledge, knowledge transfer, communication and coordination between the engineer, service companies and rig personnel, they collected best practices for well completions from data, models and expert's knowledge to combine it with a Bayesian decision node (BDN) model with 18 uncertainty nodes and 6 decision nodes using GeNIe (Graphical Network Interface). The variables include multilateral junctions, treatment fluids, lateral completion considerations, perforations, etc. Uncertainties include swelling probabilities with respect to fluid types, fluid properties, temperatures, costs, etc. In a case study the authors proved the feasibility of the system for completion decisions.

Most of the probabilities are assigned by experts; many of them are either "recommended" or "not recommended", therefore set to 0 or 1. They can be updated if more data or different information is available. Although the system is limited to usage for completion optimization only, it can serve as prove of concept for a small-scale application. In this case, the acceptance of the system by the engineers and the integration in their daily work will be of special interest.

Bayes' Model for Oil/Gas Processing Networks Failure

Edeko and Mbamalua[112] from the University of Benin developed a Bayesian Model for the diagnosis of six categories (or hypotheses) of pipeline failures: stress rupture, lack of quality control environment, sand erosion, corrosion and natural cause.

The authors approached the problem of uncertainties by using the Bayesian inference theory that allows for the update of subjective probabilities if new objective data is found. The Bayesian inference requires the connection of observable evidence e_i with the hypotheses H_i mentioned above by the Bayes theorem. The evidence e_i can take on different values 0, 1, 2,..., n depending on various parameters.[113] Then these evidences were combined to evidence vectors or patters[114], which again were related to a set of hypotheses. If a certain pattern can be found in the field application, there is a certain probability of the truth of a hypothesis.

The patterns of evidence displace the Bayesian network structure in this case, with the simplification that the evidences were assigned limited discrete values, whereas in a Bayesian network continuous probability values that better display uncertainties in most cases are assigned to nodes.

---

[111] Al-Yami (2011)

[112] Edeko et al. (2004)

[113] For example: for pipeline corrosion, e_1 can take the value 0 if the failed section is located in a straight run, or 1 if the failed section is located in a bend.

[114] The number of possible combinations is rather large, it can be computed by n^k, where n is the number of values for e_i and k is the number of evidences

Optimal Placement of Horizontal Wells

With drilling costs taking over a large part of the upfront expenditures in oil and gas operations, the placement of the well including real-time downhole monitoring is crucial for the success of the operations. Well placement decisions are made under uncertainties and have to meet multiple objectives, like costs, production rate and well design that are often conflicting and non-commensurate. These conflicting objectives cannot be resolved by purely physical models, which followed by a cost/benefit analysis. Rajaieyamchee[115] and his college from the University of Stavanger present a multiple-objective decision making approach based an influence diagram framework with two different extensions, namely MAUID (multi-attribute utility influence diagram), which allows to calculate an overall utility value from diverse monetary or non-monetary utilities, and MOID (multiple objective influence diagram), where the inferior decision rules are eliminated in different steps, finally applying trade-off values for the non-inferior decisions. The authors found that the MOUID was a feasible and suitable decision-analysis tool for that purpose, whereas the MOID still required research in order to support large decisions.

Fuzzy Reservoir Ranking

Another method for dealing with high uncertainties and multiple objectives has been demonstrated by Agbon and his colleges.[116] They used a fuzzy model to rank 45 reservoirs for gas opportunities. The model had 8 input variables, 20 fuzzy rules and an output variable, the gas opportunity index. The results were compared to those obtained by conventional ranking methods, such as Net Present Value, Internal Rate of Return, etc. and Decision Tree - Monte Carlo method. The authors concluded that the Fuzzy Model Method could better deal with the vague, non-linear and complex relationships between the input variables than the Monte Carlo method.

Conclusion Background & Theory

The chapter has provided an overview over concepts the of data, information and knowledge from a perspective of computational information processing, which will be mirrored in the basic structure as well as in the purpose of the Adaptive Advisory System described in the next chapter. Background information on the primary field of application, the production monitoring in the oil industry, has been given. The description of the petroleum industry shows the peculiarities that lead to difficulties in the application of an intelligent system, but at the same time illustrates the necessity for new approaches to data floods and highly complex problems. Bayesian networks, which are described in detail here, seem to be the optimum methodology of AI to make the Advisory System learn and adapt to new problems and tasks.

[115] Rajaieyamchee, SPE 129984 (2010)
[116] Agbon (2003)

2 APPLICATION

2.1 System Overview

The *Adaptive Advisory System*[117] is a framework for an intelligent technology that captures knowledge from experts and experiences, and reapplies it when it processing real time field data to identify problems and to help engineers to take the best actions. The Advisor integrates a whole operational process from data collection to an action into one system. The system uses methods of artificial intelligence (i.e. Bayesian networks) to constantly learn with new tasks adapt to ever changing environments.

The system is displayed in Figure 18 and Figure 19 (simplified version); it includes three layers, the Data Layer, the Information Layer and the Knowledge Layer. The structure is based on the knowledge pyramid in described in the "knowledge theory" chapter. The system is capable of adapting and leaning in two ways: It follows the general theory and logic of learning from extracting knowledge from data and information as well as the Bayesian network learning approach within the Knowledge Layer.

The Knowledge Layer forms the heart of the Advisor and consists of three main elements, the Event Detector, the Problem Classifier and the Decision Supporter. The framework that embeds the Knowledge Layer includes a Data Layer and an Information Layer.

In the *Event Detector,* unexpected signals are detected, such as values above or below threshold, missing signals, most likely values for noisy or redundant signals, deviations from trends, etc. Alarms can be triggered in serious cases.

In the *Problem Classifier*, the identified events from the Event Detector are considered to be symptoms for problems. Bayesian networks provide a tool to calculate the most likely problems under the evidence of events. For each problem a traceable ticket is opened.

In the *Decision Supporter*, a set of actions for each problem is evaluated by their utilities. The user can then execute the recommended action, what leads to a closing of the ticket.

We differentiate three different phases:

- The *(initial) set up phase*, where expertise on operations, causal structures, probabilities, etc. together with set up and testing data is used to build the Data, Information and Knowledge Layers.
- The *online phase* is the operating mode of the Advisor and includes four cases.
- The *adaption phase*, where experts can update the Knowledge Layer after evaluation of results

Next to the standard case, we differentiate three cases in the online mode:

In the case of an *immediate action* the Event Detector identifies an event, which is classified either as alarm or as standard event. For an emergency (alarm), the user is alarmed and actions that influence the assets (e.g. shut down) might be triggered automatically. For a standard event the resulting problems and decisions are well known and approved, therefore actions can also be executed automatically.

[117] In the following also referred to as Advisor or Advisory System.

A *guided workflow* is initiated either at any stage of the Knowledge Layer either by the user or by the system. In this case the user can investigate all steps that led to the symptoms, problems or actions and try to adjust, update or correct the system to find a suitable solution.

If the *results are ambiguous*, i.e. the system does not identify a clear utility for an action, a decision committee or and expert are required to evaluate the priorities and nonetheless find a good solution.

Once a problem is identified, a ticket is opened for this specific problem and resulting processes are subordinated. Each ticket requires an evaluation of predicted and actual results that help the system to improve its recommendations.

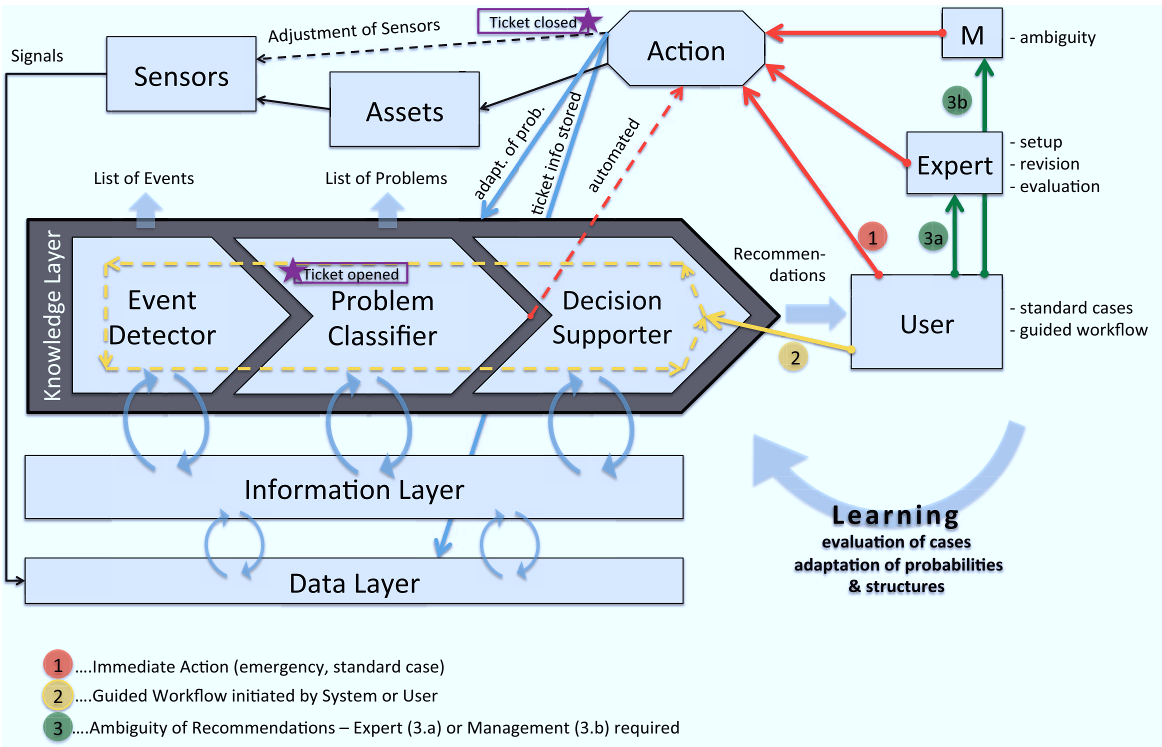


Figure 18: System Overview

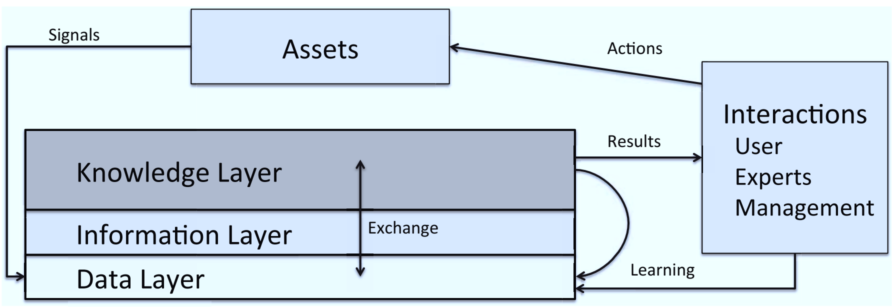


Figure 19: Simplified illustration of the Adaptive Advisory System.

The real time operations process, displayed in Figure 20, starts with the acquisition, preparation and storage of data. The engineers use newly acquired data as well as historical data, previous experience, guidelines, etc. to carry out a technical analysis, which already includes one or more concepts of actions. For the business analysis, costs of different options have to be estimated using cost databases, invoices, contracts, service and equipment bidding documents, etc. Based on the results of the business analysis (highest profits, lowest costs & risks), recommendations are given and a decision is made by the engineer or the management that leads to an action being taken. The Advisor is capable of integrating the complete process described above into one single system.

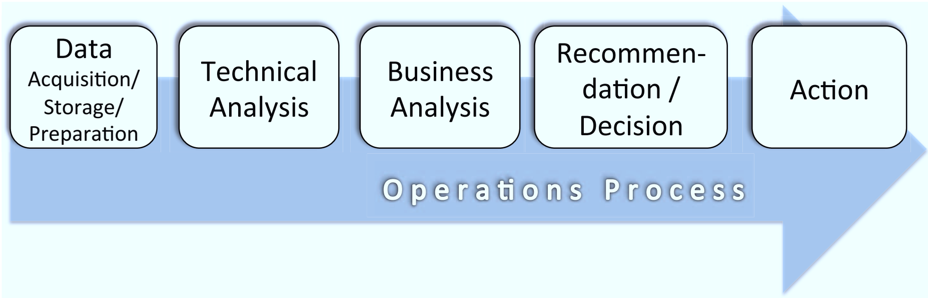


Figure 20: Real time operations process[118]

2.2 System Elements

2.2.1 Data Layer

The *Data Layer* is a database that integrates data from various sources, in various times sequences and in various formats. Next to historical data, newly generated real time data as well as activities in the Advisor, such as opened tickets, events, problems, recommendations and final actions are stored and accessible any time later.

Garbage In Garbage Out

A very crucial step in the whole system is to provide sufficient quantity and the quality of data. Any kind of expert, either human or artificial, can only deliver results in a quality that is in line with the quality the data it bases its decisions on. Or in other words, if we have a perfect decision support system, but do not feed the right data, then there is no way the decisions are right.

As already mentioned, the data available for oil and gas operations diverse in time sequence and nature and type of measurement. Often, different software applications have different means of collecting and storing data. A very difficult but crucial task is to integrate all that data into one system in a way to reduce uncertainties rather than to increase them by divagating evidences.

Another benefit of using Bayesian networks is that it does not fully rely on data, as for instance a neural network would do, since experts still set up the structures and probabilities and therefor can compensate for missing or not correct data.

[118] Hite, Crawley, Deaton, Farid, Sterneskey (2007), p. 2

2.2.2 Information Layer

The *Information Layer* fulfills two purposes: Firstly generates information and secondly it provides a link to applications.

The *generation of information* from data has been described in further detail in Chapter 1.1.4 Data, Information and Knowledge; it includes the interpretation of data, hence inferring information from data with the prerequisite of certain knowledge as well as the elaboration of information. For example in the Information Layer, models are generated out of data points with the knowledge that workover times have a Rayleigh distribution. Static and dynamic thresholds or KPIs are defined and elaborated with new data.

Information Layer provides an interface with the system to external *applications* or specific software and therefore is sometimes referred to as application layer. Examples for such applications are reservoir or production modeling software programs that deliver important data input as well as may profit from taking data output from the Advisor and complement areas of complex modeling that are beyond the Advisor's capabilities. In this context the Adaptive Advisor can be seen as a very general data processor, enabling a linkage between people and technology across various disciplines.

Value of Information

If the Information Layer cannot provide enough input into the Knowledge Layer, the costs of determination of additional information can be calculated by the system. In some operations, it is possible to make additional measurements outside the normal range, like running wireline instruments in a well or analysing the chemical properties of reservoir fluids. What influence on calculated utilities does this new information have and how does it change our decisions? What is the value of the additional information?

Bayesian networks in the Problem Classifier and the Decision Supporter include the basic function of recommending additional data, if the costs for this are lower than the expected value resulting from less uncertainty when making decisions. Rajaieyamchee[119] included the option of taking an azimuthal test in his dynamic BN. The hexagon is a multi-utility node. The optimal decision "What to do?" with the highest utility results could be to conduct an azimuthal test.

[119] Rajaieyamchee, SPE 129984 (2010), p. 8

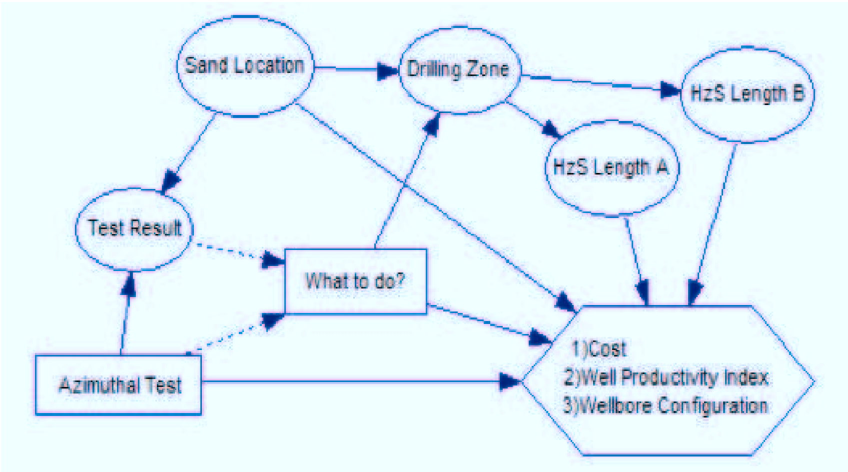


Figure 21: DBN for calculating the value of azimuthal test information[120]

2.2.3 Event Detector

The first element of the Knowledge Layer is the *Event Detector* (ED). Events can be expected or unexpected conditions in the production system that signifies a change of the operating state of the system. The methodology for the Event Detector consists of a conventional event and alarm management system, complimented by Bayesian event detection methods. The Event Detector is strongly linked to the next element, the *Problem Classifier* (PC), such that clear boundaries cannot easily be drawn. The opportunity arises to construct one Bayesian network that includes both, ED and PC, or to merge two separate networks in an early stage of the system installation.

Challenges in Event Detection[121]

Situational dependence refers to the problem that no two alarms are exactly the same. Complex systems require similarly complex event detection mechanisms. Certain parameters and variables that show one problem in a specific domain are not necessarily applicable to other very similar problems. Addressing an event detection problem in one domain may require alternative or never-before-attempted methods within other domains.

Criticality of application means that event detection often requires a high degree of precision, i.e. providing a high true positive (detecting an event if it occurs) rate and a low false positive (prevention of incorrect detection) rate. Also, a high degree of timeliness is required for many systems, e.g. a blow out in drilling operations. Therefore, especially for complex system event detection, the software processing time should be minimized.

Large and diverse data sources: The bandwidth of communication, cheap storage and developments in sensor technologies lead to an explosion of the data sources and the amount of data readily available. Sources include unstructured data, text documents, images, audio, spatio-temporal, etc. Any single event detection problem may consider a variety of these diverse data sources consisting of different data types and formats. *Large data volumes* require immense digital storage space and either high computational power or high computing time. Too much data can lead to a "paralysis of analysis", which means that an over-

[120] Rajaieyamchee, SPE 129984 (2010), p. 8

[121] cf. Kerman, Jiang, Blumberg, & Buttrey (2009), p. 5 - 9

analysis with too many indifferent options might prevent an actual action being taken by the decision maker. Too little data, however, might lead to missed detections and high uncertainties. *Raw sensor data* is often plagued by incompleteness and inaccuracies, they are noisy and redundant sensors deliver divagating results.

Network topology refers to the pattern layout of interconnections of the various elements of the event detection network, including sensors, transmitting and receiving stations and nodes, connected through cables, wires or other wireless communication. Important factors for efficiently working event detection include care and maintenance of the network or the network throughput and capacity. Problems can occur at different stages, such as the measurement (sensor failure, wrong calibration, etc.), during signal transfer (fiber-optic cables, mud pulses while drilling, satellite, etc.) and at the receiving station.

Bayesian Event Detection

An event is an occurrence or activity that is unusual relative to normal patterns of behavior. Typical event detection may be classified in four rather broad categories: statistical, probabilistic, artificial intelligence and composite methods.[122] Statistical methods include static threshold methods, different types of regression (linear, polynomial, etc.), or a time series analysis, when successive data points occur after a uniform time interval. Probabilistic methods in contrast to deterministic methods with exact values allow accounting for uncertainties. Various models, like the time-varying Poisson process model, the distributed Gaussian method are implemented for event detection.[123] Methods employing artificial intelligence usually are both computationally and informationally intensive; stand-alone they require advanced fusion algorithms to integrate data from multiple sources.[124] Fuzzy methods, neural networks, petri nets etc. can be used according to the requirements of the system.

Integrating event detection in the Advisor has the advantage of using data that is already prepared as input for the Bayesian networks and therefore does not require much additional data processing. Still, for some event signals, static thresholds, trend comparisons, etc. contribute to optimal event detection.

A Bayesian network methodology for event detection can deal with uncertainties and noisy or redundant signals. Since the network models the physical reality, it is beneficial for closed systems containing many non-measurable variables like many reservoir parameters. Interferential sensing helps to investigate what is actually going on in the underlying system. To setup the event detection system, the Bayesian network can either be trained using historical event data, or for the detection of rare events, where no sufficient historical data is available, the network and according probabilities can be determined by experts.

Event Classification

The Event Detector offers a list of events to the user that is updated in real time. The events can be classified by their impact (i.e. monetary value), urgency (i.e. change of impact with time) and frequency, as shown in Figure 22. The *impact* of the event is measured in terms of monetary or other value to the company, such as loss of production, costs of maintenance, etc. The calculation of the impact value will be an important input for the evaluation of utilities of recommended actions in the Decision Supporter. The *urgency* of an event is determined by the required response time. Urgency means that the costs or other

[122] Kerman, Jiang, Blumberg, & Buttrey (2009), p. 8
[123] Kerman, Jiang, Blumberg, & Buttrey (2009), p. 11
[124] Kerman, Jiang, Blumberg, & Buttrey (2009), p. 11

risks of an event strongly increase with the time it takes to react, or that another event can only be prevented when reacting within a certain timeframe. For example, a sudden increase in sand production can erode the bearing of a compressor very quickly and immediate action is required. *Frequency* also is an important classification of events. Some events occur only once in a lifetime of a well, whereas other events, like pressure or volume fluctuations, etc. occur hourly or daily.

The users have the possibility to change the position of the dots on the interface, hence changing impact or urgency. The Bayesian network that created the classification is able to learn from these changes and better position similar events the next time they occur.

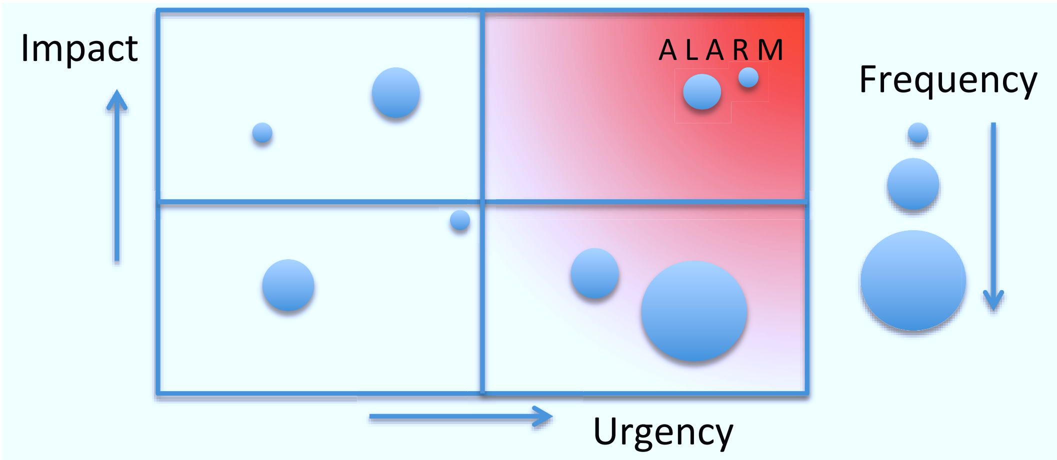


Figure 22: Display of events categorized by impact, urgency and frequency.

Alarm Management

Some events may require immediate action. A threshold urgency can be set above which the events are classified as alarms (red area in Figure 22). In exceptional cases, a callout alarm is generated by the system. Predefined personnel are immediately informed via phone, smart phone application, pager, email, etc. according to the severity of the event. Meanwhile the Advisor identifies the problems that caused the alarming events in order for the engineers to optimally and timely deal with dangerous situations.

Setting the thresholds for alarm generation is crucial for the success of the alarm management system. Thresholds can be set for single signals or events, combinations of signals or events, and at problem analysis stage at the PC. On the one hand, unnecessary alarms floods greatly increase the costs of the system, since alarms require more attention than normal events; alarm floods may overcharge the capabilities of the personnel and critical situations may be missed. On the other hand, a system not reacting to critical situations is even more dangerous than not having an alarm system at all.

2.2.4 Problem Classifier

A *problem* is an occurrence that currently or in the future acts against the objectives of the company hence the elimination of the problem results in a business opportunity of the company. Problems in most cases are indicated by a combination of single point measurement deviations and point out complex issues that cannot be measured directly. Therefore problems usually include a probability of occurrence.

The Bayesian network of the Problem Classifier includes nodes for symptoms and nodes for problems. Symptoms can either be data, information from the Information Layer or the Data Layer respectively, real time measurements from the field, or the events that have been identified by the Event Detector.

Naïve Bayes and Logic Regression

There exist two types of models, the *generative* and the *discriminative* model as shown in Figure 23. Dealing with the interactions of problems (Pr) and Symptoms (S), discriminative models directly estimate the parameters of P(Pr|S), whereas generative models directly estimate parameters for P(Pr) and P(S|Pr). The former is often referred to as logic regression, and the latter as naïve Bayes.

With the *generative model*, we can use Bayes rule as the basis for designing learning algorithms in the following way: We wish to learn the probability of a certain problem given the symptoms, P(Pr|S), we use the training data set to learn estimates of P(S|Pr) and P(Pr). New symptom cases can then be classified using these estimated probability distributions, and the Bayes rule P(Pr|S) = P(S|Pr)•P(Pr)/P(S). Learning Bayes classifiers typically requires an unrealistic number of training examples some form of prior assumption is made about the form of P(S|Pr). The Naïve Bayes classifier (generative model) assumes all attributes describing symptoms are conditionally independent given a problem. This assumption dramatically reduces the number of parameters that must be estimated to learn the classifier. Naive Bayes is a widely used learning algorithm, for both discrete and continuous nodes.[125]

Logistic regression is a function approximation algorithm that uses training data to directly estimate P(Pr|S), in contrast to Naive Bayes. In this sense, Logistic Regression is often referred to as a discriminative classifier because we can view the distribution P(Pr|S) as directly discriminating the value of the problem (target value) for any given symptom.[126]

It can be concluded that in most cases the constructor of a Bayesian network is advised to use the structure of the generative model rather than the discriminative model. In case of the discriminative model, a problem that can be detected by 5 symptoms with 3 states each, results in 5^3 or 125 different combinations. In order to solve this, either 125 different probability estimates or unrealistically large data sets are required. With the generative model it is easier to address complex problems and integrate separate expert knowledge domains into the network, since experts can be interviewed individually on their estimations of isolated symptom problem interactions.

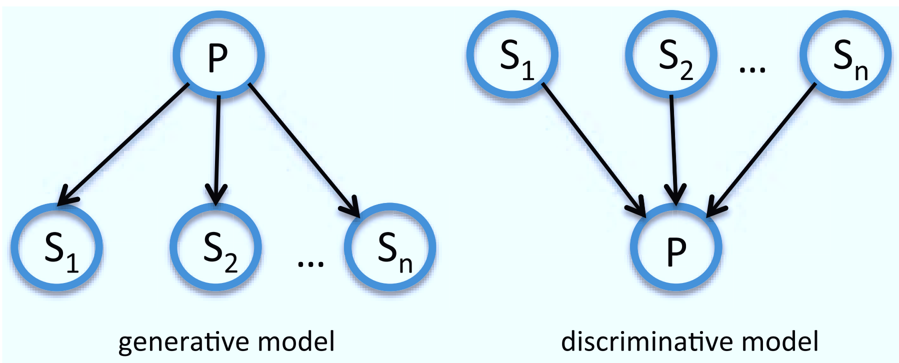


Figure 23: Generative model and discriminative model

[125] Mitchell (2010), p.15
[126] Mitchell (2010), p.15

Temporal Bayesian Problem Solving

Whereas naïve Bayes and logic regression refer to non-temporal problem solving, *dynamic Bayesian networks* (DBN) offer the possibility to include time aspects, as already mentioned in the theory part. The temporal nature of causal relationships is often essential to problem solving and decision-making in oil and gas operations. DBNs allow predictions of future parameters like reservoir pressure or production rates and can to a certain extent integrate functionalities of external models into the Bayesian network.

Many cause-effect-relations include a temporal aspect. For instance a down-hole pump failure at a time t=0 will lead to a reduced flow rate at surface at a time t>0. Surveillance of trends will offer important hints and insights. Generally, time plays an important role in the evaluation of probabilities of certain problems, i.e. older installations and facilities will usually require more frequent maintenance than new ones.

Modelling The Physical Reality

When constructing a Bayesian network, we have to differentiate between actually observable variables and the physical reality that is unknown. In order to better model the physical reality, *non-observable* or *hidden variables* are included in the network. An example by Giese and Bratvold[127] is illustrated in Figure 24. Gas influx depends on the equivalent circulation density (ECD) and the pore pressure. Since it is not possible to measure the ECD, it can only be estimated from mud weight and flow rate. The pore pressure is a function of the depth and the geological properties around the wellbore. The wellbore is not perfectly vertical, so the total vertical depth (TVD) has to be derived from the measured depth, i.e. the length of the wellbore. The more the hidden variables are considered in the model, the better the physical reality is mirrored by the BNs, which leads to more accurate results and a better understanding of cause-effect-relations of problems.

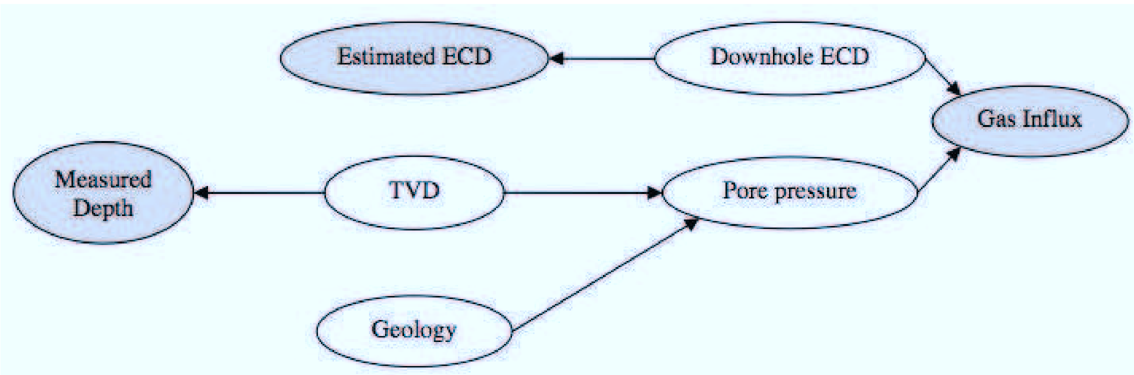


Figure 24: A network for gas influx including observable and hidden variables.[128]

2.2.5 Decision Supporter

An important step in a real time operations process (see Figure 20) is the *business analysis*. The business analysis is usually conducted by the engineer, who designs the concepts, but is often regarded as an annoying task as being not within the core responsibilities of the engineer. Accordingly, some engineers may pay not enough attention to this part of the analysis although it is decisive for the finally decision. The decision supporting part of the Advisor enables the evaluation of the utility of several different options with regards to the business

[127] Giese, Bratvold (2010), p.7
[128] Giese, Bratvold (2010), p.7

opportunity. This step is automatically performed for each option selected and includes an evaluation of the monetary value of an option, as well as risks and other factors like environment, availability of equipment or the compliance with the company strategy.

The *Decision Supporter* is a Bayesian network extended by influence diagrams with input in form of real time data, data and information from the Data Layer and the Information Layer, identified events and classified problems and an output in form of utilities for different recommended actions. Just like for the Problem Classifier, hidden or intermediate nodes complement observable variables and the differentiation between discriminative and generative model is valid.

Multiple Criteria Decision Making

In most complex real world problems, decision criteria are conflicting in nature and their interrelations are often interdependent in complex and uncertain ways. Moreover, additional external factors, like the oil price or costs of services, play an important role and have to be included.

In practice, Bayesian networks for decision making in oil and gas operations will be dynamic, include multiple criteria and are extended by influence diagrams (IDs). They include chance nodes, decision nodes and utility nodes. For each problem identified in the Problem Classifier, the best action over all available alternative actions has to be chosen. The best alternative is the one that meets the objective in the best possible way, hence maximizes the overall utility. Figure 25 shows a schematic Bayesian network including a set of possible actions (alternatives), a set of criteria and sub-criteria, factors influencing them and a set of objectives subsumed in one utility node.

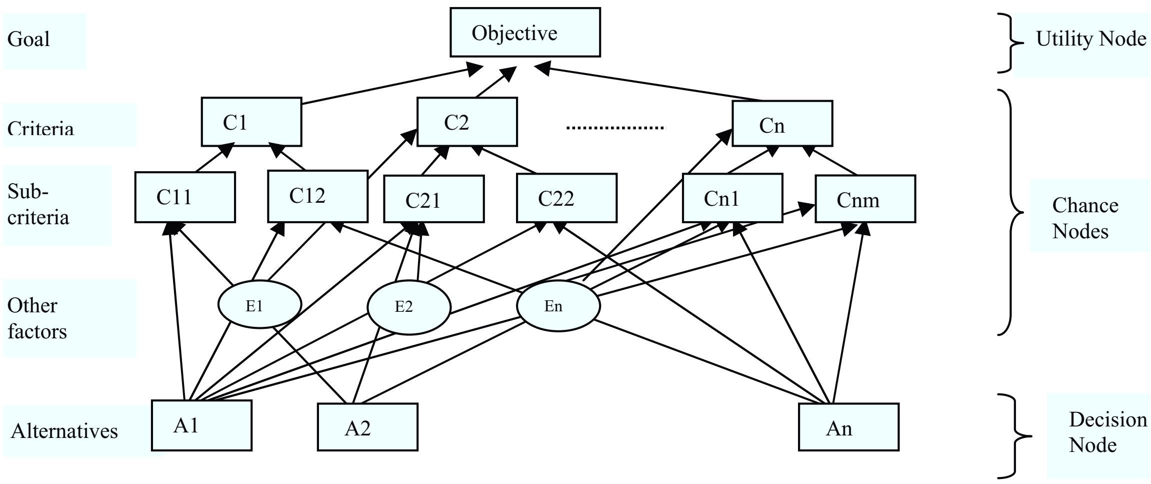


Figure 25: Multi criteria decision making[129]

Objectives

The objectives and the weighting of different criteria in oil and gas operations strongly vary with varying policies and goals in companies. The objectives for oil and gas operations can be summarized in four categories: economic, strategic, operational targets and HSEQ & Risk[130].

[129] Watthayu, Peng (2004), p. 5
[130] HSEQ = health, safety, environmental and quality

The *economic* target can be assessed with the calculation of the net present value (NPV). This compares capital and operational expenditures with additional revenues generated (oil price, production rate) under temporal aspects (discount rate). The calculations may include simpler profit calculation, or use other criteria (e.g. internal rate of return) according to the company's accounting policies.

HSEQ & risk includes the probabilities of accidents like oil spills, blow outs, injuries, etc. For example, many injuries on an offshore rig occur while lifting operations, if a lifting operation is included in one of the options, its utility is decreased accordingly.

Operational targets are short-term to medium-term targets and address efficiencies of processes (e.g. reduction of non-productive time), logistics (e.g. high utilization of equipment), etc.

Strategic targets address the company's preferences and long-term targets. An example for a strategic target would be a companies plan to increase its overall barrels produced per day, then additional production would be weighted more than costs to reach this goal.

The main utility value is a combination of various sub-utilities, therefore a unified evaluating system has to be established that allows to compare monetary and non-monetary targets.

Dynamic Decision Making

Like in problem classification, DBNs account for the temporal aspects of decision making. The history of an asset and right time for taking actions such as maintenance, work over operations, employments of artificial lift methods, etc. play an important role in maximizing the overall profit. Figure 26 shows a BDN for decision making including a super-value utility node that subsumes sub-utilities at different points in time.

The dynamic character of the network allows accounting for temporal *interdependencies*. An example for this can be found in the case study in the next chapter. There is one option to increase the initiation pressure of a gas lift valve. Assuming the Problem Classifier has detected problems with the surface gas flowlines and the compressor, it is only recommendable to increase the pressure if the problems are solve in the first place, because otherwise the gas consumption and pressure losses are too high for the rather small resulting gain in oil production.

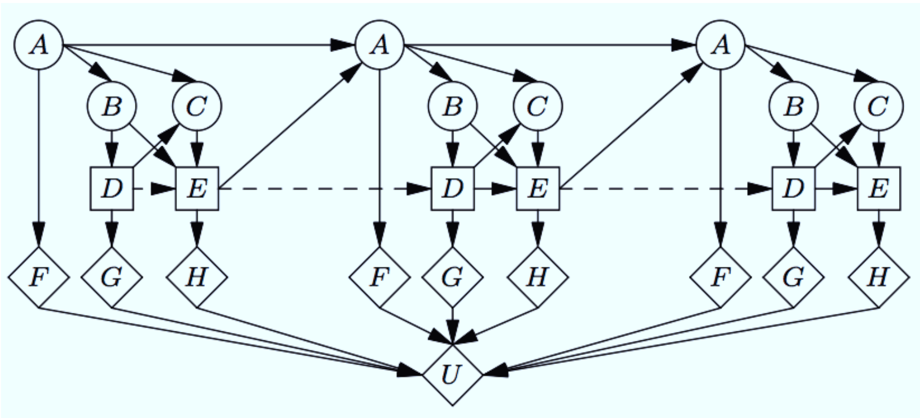


Figure 26: Dynamic influence diagram with local utility functions and a super-value node.[131]

[131] Van Gerven (2007), p. 91

2.2.6 Other Elements

The Layers of the Advisor are just one part of the system. The structure shows a close link to the human actors, the users (engineers, expert or the management), to the element of an action, which has to be carried out to close a ticket and the asset that closes the loop.

Users

Users are an integral part of the Adaptive Advisory System. It should not be seen as an intelligent machine replacing engineers and experts, but as a knowledge sharing enabler: The system learns from its users and users learn from the system.

The term "users" refers to all personnel interacting with the system. Users can be engineers, experts or managers. The standard user could be a young *engineer*, who does not have a lot of practical experience but understands the nature of the operations the system deals with, hence he/she benefits optimally from the knowledge provided by the advisor.

Experts are engineers or scientists with a more profound knowledge on either specific areas of oil and gas operations (geologists, drilling engineers, etc.) or generalists, who have aggregated a lot of experience with an asset, including error-proneness of tools and other peculiarities. The person responsible for the initial design of the Advisory System will be called an expert as well. The in-depth knowledge of different experts is required to initially set up the system for the determination of causal relationships, for verifying the plausibility of probabilities learned from data and for estimating probability values and utilities.

Managers interact with the system if it provides several alternatives with a similar utility value. The manager or any other decision committee takes over the responsibility if the results are ambiguous. Also, managers may use the system to retrieve higher level KPIs and other data they require for certain analyses.

Action

One of the most important principles in alarm and event management is that each event requires an action to be taken. The process of event detection, problem classification and decision support is only completed with an appropriate action that causes a change in the asset and yet in the signals reaching the Advisor.

Assets

The term asset is used in a broad sense and can refer from a single machine, a well, a group of wells, an offshore platform, a field, reservoir, an underground gas storage plant, a production plant, etc. to the whole production system of a company. Signals are measured by instruments directly at the asset and whenever an action is performed, they influence the newly recorded signals and thus the Advisor.

2.3 System Workflows

2.3.1 Initial Setup

The construction of the Bayesian network can be conducted in four steps: variable definition, structure specification, factor association and parameter estimation.[132]

[132] Van Gerven (2007), p. 42

Variables are the nodes of the network; they can either be chance, decision or utility variables. Along with the category, the type and state of the variable have to be defined. The type of a variable refers to its nature being either discrete with mutually exclusive states, or continuous. In practice it is easier to start with a simple initial model, carefully selecting variables according to its importance and then in a next step adding other variables, until the model is accurate enough.

The *structure* of the model ideally mirrors the physical reality of problem causalities. Experts can use their own knowledge and previous experience, literature knowledge and in some cases even training data to set up the structure. Again, starting off with a simple model and gradually increasing the complexity and adding small domain fragments will ensure the functionality of the model. In case of the Advisor, the Event Detector, Problem Classifier and Decision Supporter should be set up as individual systems and then be integrated in a subsequent step.

Factor association refers to the relationship parent variables and child variables. A factor defines the functional form of how the outcome of the random variable of the child depends in the state of its parents.

Parameter estimation can be carried out in three different ways. The easiest and most accurate way is to make the network learn the probabilities from sufficiently large and good quality training datasets. Datasets with these features are often not available. In this case experts can estimate the probabilities based on their previous experiences within a certain technical aspect of the Advisor. The third option is to derive probabilities from literature. The initial probabilities are adapted automatically in the online mode of the Advisor, when it deals with new data, problems and solutions.

2.3.2 Online Mode

The online mode is the normal operational mode of the Adaptive Advisor System. Signals are measured at assets in real time and complement already available data and information at every point in time. Users, experts and managers have access to the system and interact with it for their daily tasks. The system alarms, lists of events and problems and recommends actions for the users. It captures all human interactions with the system, evaluates decisions and automatically adapts its calculations to changing environments and new datasets.

In the online mode besides the standard mode we differentiate three cases, the *immediate action*, the *guided workflow* and the *ambiguity of results*. The modes are displayed in form of sequence diagrams and can be found in the appendix. These three cases are described verbally and are displayed in form of sequence diagrams that can be found in the Appendix. The sequence diagrams shall support the software engineer when programming the Adaptive Advisory System.

Sequence Diagrams

Sequence diagrams are dynamic diagrams in UML (unified modelling language), which standardizes the visualization, specification, construction and documentation of a variety of object-oriented systems. Sequence diagrams emphasize the temporal aspects of dynamic diagrams and show how processes operate with one another and in what order. The sequence diagram has two dimensions: On the horizontal line, objects are listed and the vertical line represents the time. Objects possess parallel dashed lines, so-called lifelines, showing the existence of the objects in a certain period of time.

Interactions between objects are represented by horizontal arrows with full heads, which point from the emitter to the recipient (activation of method) or stick heads (return messages); the name of the activation method is written on top of the arrow.

Additionally, the activation boxes, or method-call boxes, are opaque rectangles drawn along a certain time sequence in the lifelines to represent that processes are being performed in response to a message. Sequence diagrams for immediate action, guided workflow and ambiguity of results can be found in Appendix h, i, j.

Standard Case

In the normal operating mode, signals that are measured at the assets arrive at the Data Layer of the system, where they are stored. The Information Layer processes the newly included data and integrates it with already available information. In the Knowledge Layer, the *Event Detector* receives data and information to be able to identify events as unusual situations. The event detector calculates urgency and impact of events and displays them for the user.

The events, together with other data and information, function as symptoms (input) for the *Problem Classifier*, which ranks different problems or sets of problems according to its probabilities, given the symptoms. At this stage a *ticket* it opened for each most likely problem. A ticket stores all information available to a certain problem; it requests an action, since it is only closed after a decision has been made and an action has been taken. Problem tickets are the main classification system of problems; they establish interrelations with and referrals to previous and similar problems.

For each ticket (problem) the Decision Supporter lists recommended actions according to the calculated utility of an action and provides it to the user. Based on the recommendations of the systems the user can make a final decision and initiate actions. The user will be asked to adapt calculations if necessary and provide additional information on the action taken (e.g. time to evaluation of action, establish connections with other comparable actions, etc.). Once the action has been performed the ticket status is changed to *“evaluation required”*.

Immediate Action

There are two cases, where the Advisor takes immediate action, bypassing the decision process with users, experts or management: emergencies or standard situations. The Event Detector is designed to identify *alarms*, which are events with a high urgency requiring immediate action (e.g. if certain predefined thresholds are exceeded). In this case, the user is alarmed and actions that influence the assets (e.g. shut down) are triggered automatically to keep the damage as low as possible. A *standard situation* is a category of problems or situations that have occurred frequently in the past. The resulting problems and decisions are well known and approved; to save time and reduce the workload of engineers, the actions can be executed automatically.

Guided Workflow

A guided workflow can be initiated at any stage of the Knowledge Layer. The initiator can either be the user, if the recommendations, utilities, problems etc. seem unrealistic, or by the system, if it encounters obstacles in its tasks, such as missing information, completely new situations, errors related to the networks algorithms, etc. In this case the user or expert can investigate all steps that led to the symptoms, problems or actions in detail and try to adjust, update or correct the system to find a suitable solution. Figure 19 shows an example for the guided workflow investigation process: The system displays the obstacles and further allows the user to click deeper into inputs, calculations and results.

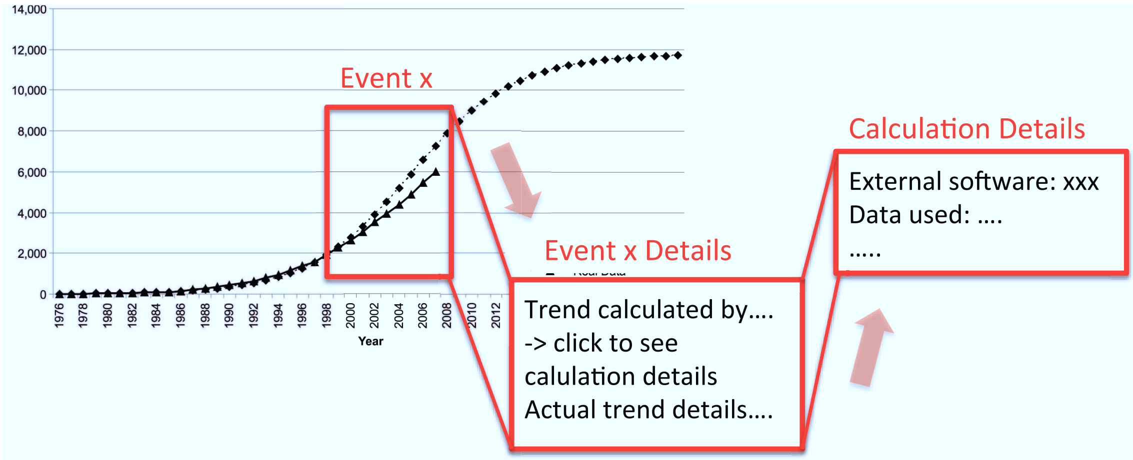


Figure 27: Guided workflow

Ambiguous Results

In case the Advisory System cannot offer a distinct satisfactory solution (e.g. several problems with the same likelihood or similar utilities for recommendations, human interaction is requested). Probabilities, priorities and weights of intermediate utilities can be adapted by the user or experts to identify clear recommendations. In case the results are still ambiguous, the alternatives are handed over to the management or another decision committee that makes the final decision. Capturing the decision maker's reasons for or against various alternatives is an important step for the further improvement of the Advisory System.

2.3.3 Learning

The Advisor captures all interactions of users, experts or the management and automatically adapts the structures of the Bayesian network(s) and the probabilities of the nodes. With each set of events or problems, the Advisor can increase its database and knowledge space; each data input, decision or action further increases the accuracy of recommendations, just like an engineer, it learns with every task. It is important for the engineers and other users to understand that each step they do not perform together with the system will result in a knowledge gap and thus lead to a lower performance.

Routine Evaluation

The evaluation is an important step to sustainably improve the system and learn from bad as well as good experiences. For each ticket (problem) an evaluation is requested by the system. Before the ticket status is changed from "active" to "evaluation required", the user is asked to provide additional information necessary to optimally perform the evaluation, such as the time frame for an evaluation that changes according to the problem or action. For example, if the system recommends to install artificial lift system, it can take up to months to several years until the action has been performed and even longer until the expected profit can be realized. After some time, the system can suggest time frames automatically. Also the KPIs that determine the success or failure of an action should be provided or reviewed by the user.

Adaptation

Once the time limit for the evaluation is reached, the user is informed and provided with details on KPI's, calculations, etc. He/she compares expected and actual outcome and in a semi-automated process determines the reasons for divagating values.

In order to also take good performance into account, a weighting system can be implemented: for satisfactory performance, the user can define a weighting factor for the whole set of data, so the relative influence of the specific set of data on the network is increased.

For bad performances, further investigation similar to the guided workflow is required and the systems structures and probabilities are updated. It is essential to perform an evaluation of all steps and calculations, since correct final results are not a guarantee of correct intermediate results.

2.4 Case Study

The following case study demonstrates the principles of using Bayesian networks for event detection and symptom and problem classification for a gas lift oil well. For the demonstration case the software 'Netica' by Norsys Software Corp. has been used. Although the example presents a real problem set in petroleum production operations, the numbers are not based on actual data, evidences, symptoms and problems do not represent the full scale of possibilities. The examples are over-simplified and their purpose is pure demonstration of the principles of Bayesian networks.

Overview

The example deals with problems in an oil field with several wells that are produced with continuous gas lift. When the reservoir energy is too low for the well to flow, or the production rate desired is greater than the reservoir energy can deliver, it becomes necessary to put the well on some form of artificial lift to provide the energy to bring the fluid to the surface. Gas lift is one of these artificial lift methods, whereby gas is injected in the production tubing to reduce the hydrostatic column of the fluid column. The principles of gas lift method are shown in Figure 28.

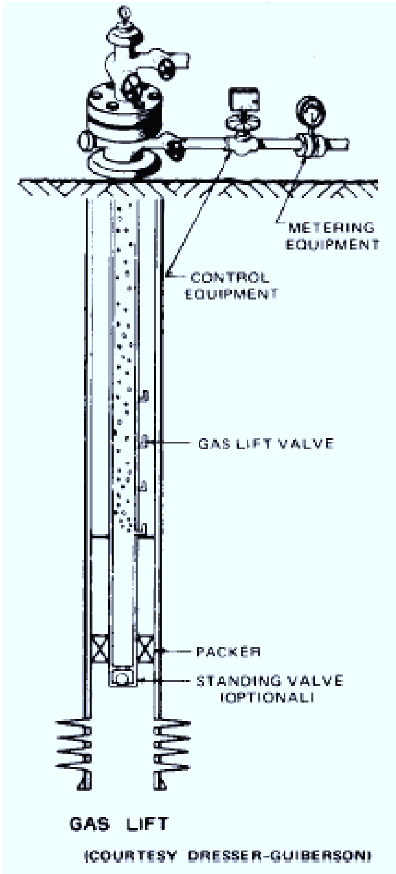


Figure 28: Gas lift method[133]

2.4.1 Event Detection

Bayesian networks will only be one method amongst many for event detection in well problem analysis. Figure 29 shows a schematic Bayesian network that evaluates the actual production rate in relation to the projected forecast. The network includes several measurement nodes, like various pressure measurements, fluid volumes, temperature, etc. The resulting real time production rate can be displayed as an absolute number (i.e. in bbl/day) or trigger an alarm if it is above or below a certain threshold. Classified as "high", "normal" or "low" it functions as input for the problem classification or decision support examples in the next chapter.

[133] Modified from Schlumberger (2000), p. 3

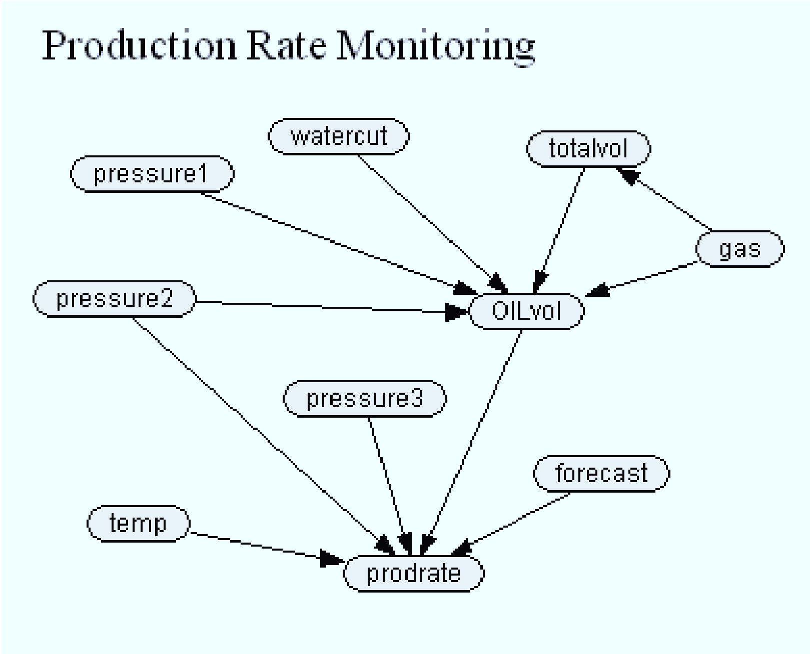


Figure 29: Bayesian network for production rate monitoring

2.4.2 Problem Classification

The well problem analysis example deals with an oil well being produced with the method of continuous gas lift. 5 Parameters were selected for provision of evidence for the classification of problems:[134]

- Well Flow
  - Flowing
  - Not flowing
- Fluid Production
  - Above anticipated production rate
  - Equivalent to the anticipated production rate (normal)
  - Below anticipated production rate
  - No production rate
- Gas Injection
  - Gas intake
  - No gas intake
  - Irregular gas intake
- Fluid Pressure: The actual pressure curve from the bottom of the well to surface shows a change in slope at a certain depth
  - Slope 1: One distinct slope change at depth of operating valve
  - Slope 2: One distinct slope change at depth different than operating valve depth
  - Slope 3: More than one slope change
- Temperature:
  - One distinct temperature peak and small temperature difference at depth above operating GLV

[134] The basics for the selection, a table on the assessment of well problem analysis, can be found in the Appendix.

- One distinct temperature peak at depth different than operating GLV depth
- Multiple temperature peaks
- No temperature peaks

These 5 parameters function as evidence for the following problems:

- Problem 1: Surface gas flow failure
- Problem 2: Operational gas lift valve is plugged
- Problem 3: Gas lift valve port size too small
- Problem 4: Communication between casing and tubing

The model in Figure 30 displays a Bayesian network with 6 chance nodes. The 5 grey nodes incorporate the evidences of the problem. The user can select a certain state of each parameter, e.g. the "well flow" to either yes or no by clicking on it. The selected state then has a probability of 100, which corresponds to a certain occurrence of that state. The yellow "Problem" node shows the probability of occurrence of problems 1, 2, 3, 4 or other problem according to the selected parameters.

In the example, the most likely problem is problem 1, a surface gas flow failure. The Bayesian network assigns a probability of 83%.

The model has been trained with a set of historic parameters that are shown in Figure 31. The model could also be used the other way around: assigning a probability of 100% to a certain problem state, the probability of occurrence of the states of parameters is displayed in the corresponding nodes that can then be compared to actual findings.

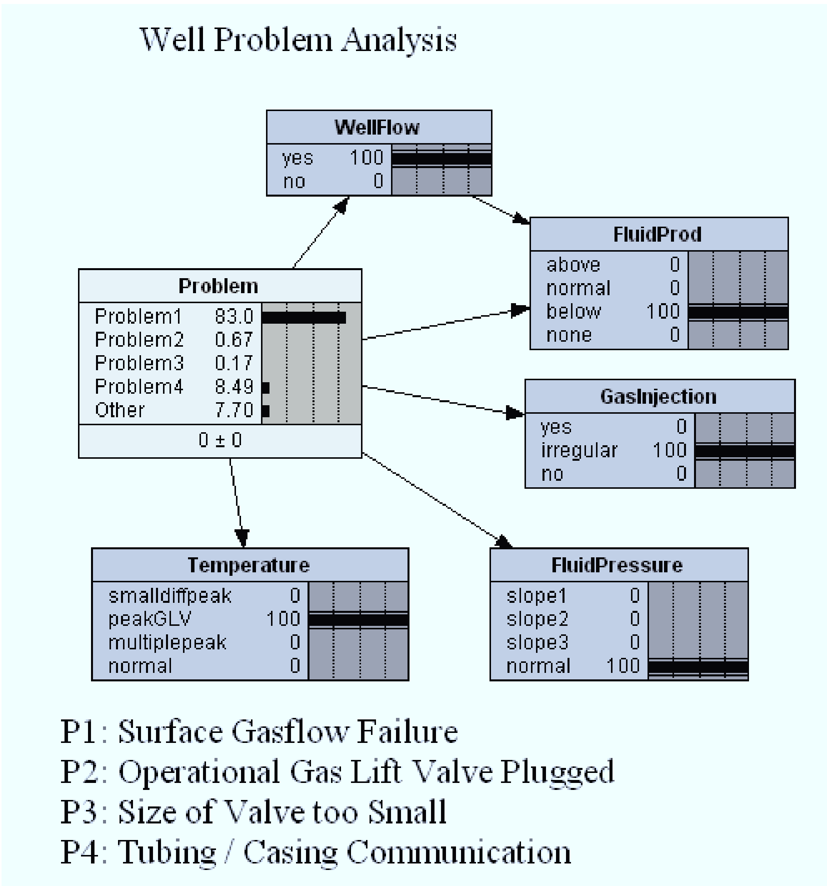


Figure 30: Bayesian network for well problem analysis, modelled with Netica.

//~->[CASE-1]->~

IDnum	FluidProd	GasInjection	WellFlow	FluidPressure	Temperature	Problem
1	above	irregular	yes	normal	normal	Problem1
2	normal	irregular	yes	normal	normal	Problem1
3	below	irregular	yes	normal	normal	Problem1
4	none	yes	no	normal	normal	Problem2
5	none	no	no	normal	normal	Problem2
6	none	irregular	no	normal	normal	Problem2
7	below	yes	yes	slope1	smalldiffpeak	Problem3
8	below	yes	yes	slope1	smalldiffpeak	Problem3
9	below	yes	yes	slope1	smalldiffpeak	Problem3
10	below	yes	yes	slope2	peakGLV	Problem4
11	none	yes	no	slope2	peakGLV	Problem4
13	below	yes	yes	slope2	peakGLV	Problem4
14	none	yes	no	slope2	peakGLV	Problem4
15	none	*	no	*	*	*

Figure 31: Part of training data table for the well problem analysis Bayesian network.

2.4.3 Decision Support

The Bayesian network for problem classification hinted to failures of the gas flow at the surface. This helps to further investigate different parameters, to come to the best recommendation for this problem set.

Technical Background[135]

Continual gas lift method has been applied on oilfield "K" since 1988. In the early period of its operation, in conditions of constant injection pressure, low water content in production (inactive water inflow), and relatively high value of reservoir pressure, an economical and stable performance has been achieved.

The initial specific consumption of gas has ranged between 200-300 m^3/m^3, and has maintained in normal conditions.

However, unforeseeable instability of injection pressure (often compressors' failures and lack of spare parts, as well as injection pressure's dependence on main pipeline pressure), have caused the specific gas consumption volume increase up to 800-1200 m^3/m^3.

Such conditions have led to an average production rate decrease in all the wells (8 in total), and to production fluctuations (1.5-12 m^3/d).

Figure 32 shows a two – pen charts for one of the wells in the K field. The erratic condition of flow is indicated by wide variations in the pressure on both the casing and the tubing. The Event Detector would indicate such pressure variations and forward the information to the Problem Classifier or Decision Supporter.

[135] cf. Solesa, Case Study K Field (2004)

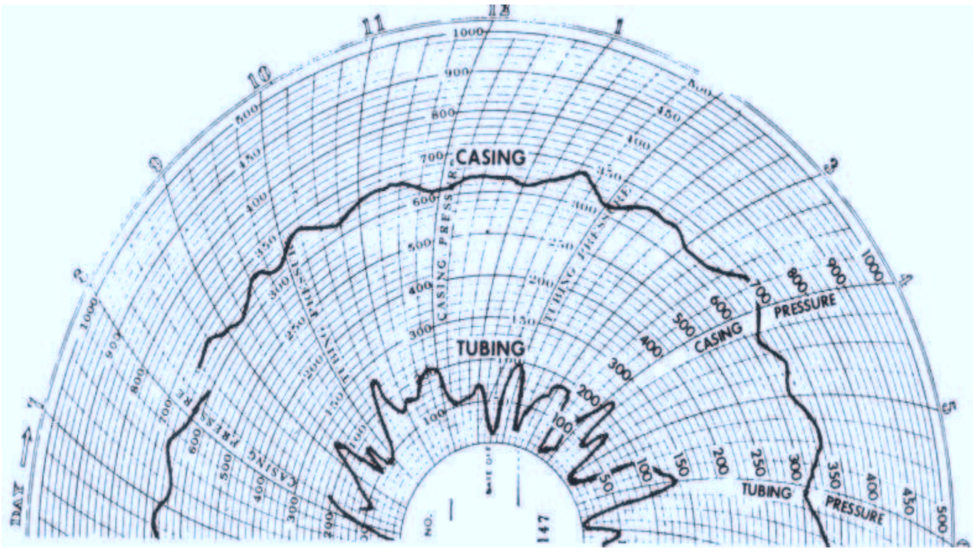


Figure 32: Two - pen chart for continuous gas lift[136]

In Figure 33, the oil production rate is displayed as a function of the injection gas rate and the injection gas pressure. The graph is based on a simulation; hence there is a possibility to increase the gas pressure from currently 28 bar to 32 bar or 40 bar, which probably will result in an increase in oil production. The model is produced by an external application, but the initiation is derived from the Advisor, because it identified a positive "value of information".

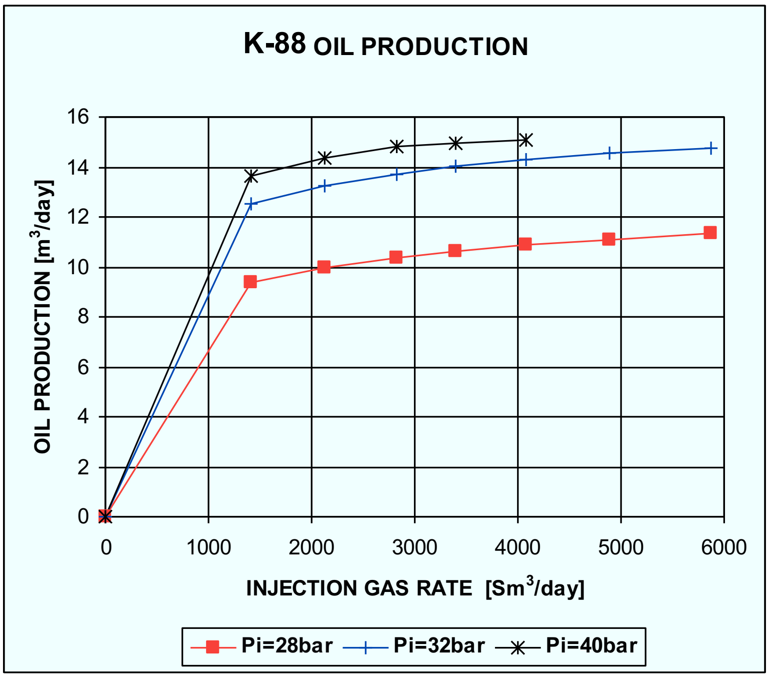


Figure 33: Oil Production as a function of injection gas rate and pressure[137]

[136] Solesa, Case Study K Field (2004)
[137] Solesa, Case Study K Field (2004)

The Bayesian Network

The Bayesian Network for decision support suggests an optimal action facing a certain problem. It consists of 4 grey chance nodes, two blue decision nodes and two red utility nodes (hexagons).

In summery, the parameters and selectable states are the following:

- Conditions regarding the surface pipeline
  - OK
  - Pressure and rate fluctuations observable
  - Wrong diameter, which leads to high friction pressure losses (from previous analysis)
- Gas consumption for the gas lift operations
  - Low
  - Medium
  - High
- Initiation Pressure (Pini)
  - 28 bars
  - 32 bars
  - 40 bars
- Production rate
  - Low
  - Medium
  - High

With decision node "TestPini", the user can decide if he/she wants to confirm the model in Figure 33 with an actual well test. The utility calculation for this node incorporates the initiation pressure and the production rate, with the utility of testing being higher if the production rate is low, since it involves more uncertainty. If the production rate is at its estimated limit already, testing does not add much more value for the decrease of uncertainty.

The test involves costs, but its "value of information" (see Chapter 2.2.2 Information Layer) is a reduction in uncertainty of predicted production rate increase. Therefore it is capable of affecting the utilities of the recommended actions. Comparing Figure 34 and Figure 35, the Decision Supporter network with selection of yes and no at the chance node "TestPini", the utilities of different actions are higher when certainty about increases in oil production rate is increased. A workover operation is only justified (positive utility) if a test approves a resulting increased production rate. Just like in a decision tree, the expected monetary value depends on the probability of occurrence.

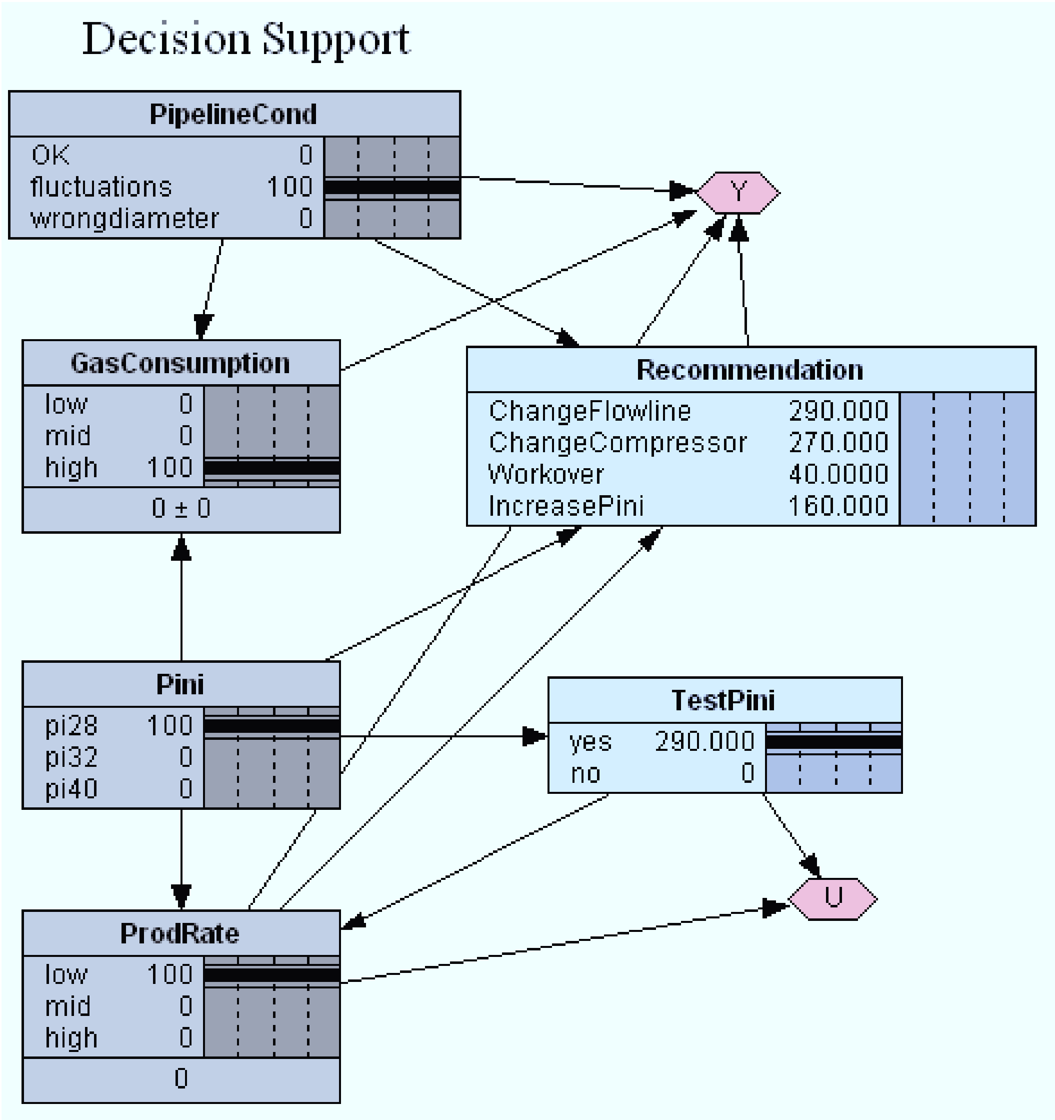


Figure 34: Bayesian network for decision support with decision for the production test

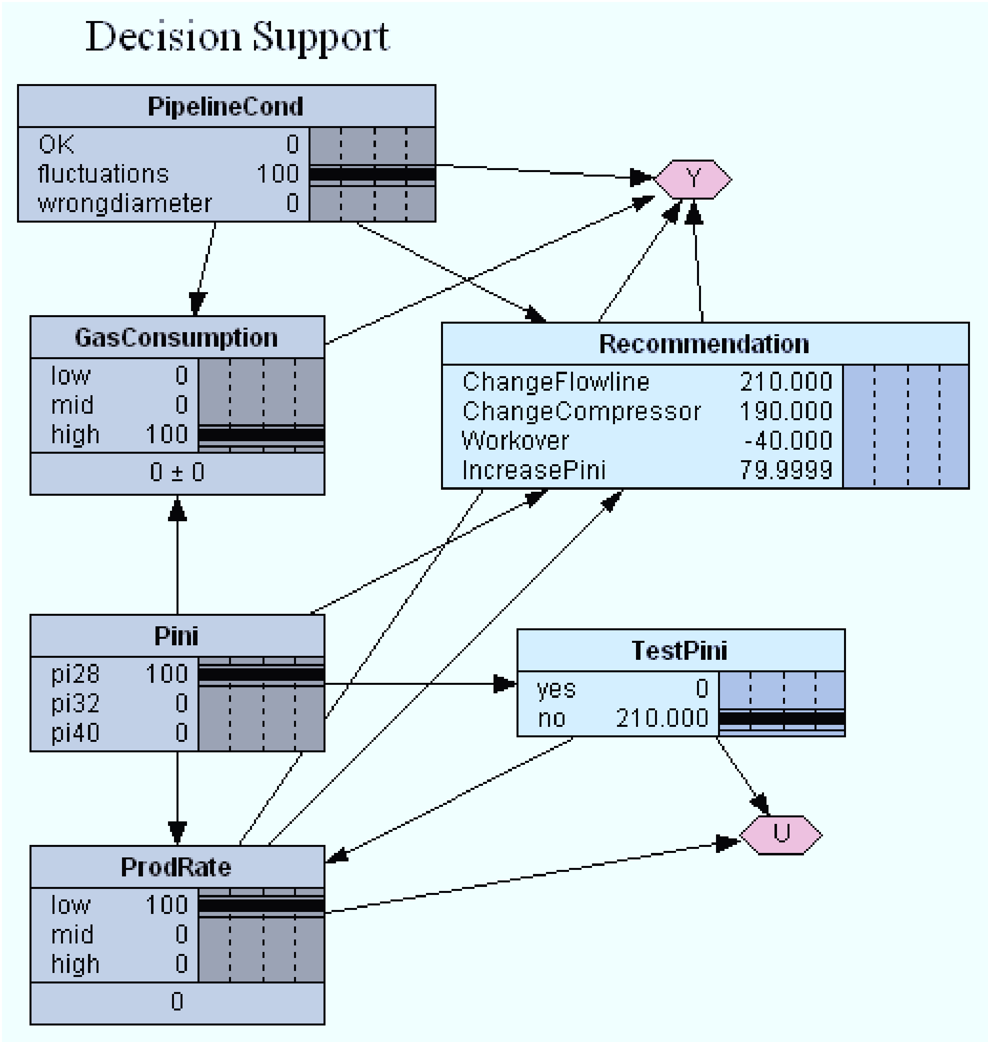
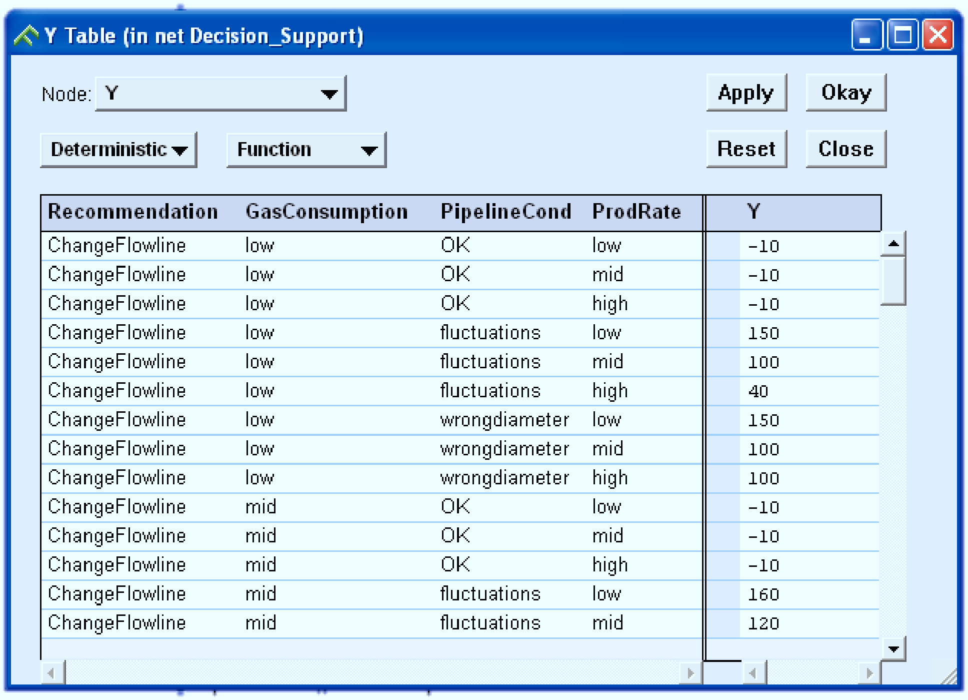


Figure 35: Bayesian network for decision support with decision against the production test

In this example, the utility values are indexes that are manually entered in a table provided by Netica. Figure 36 shows the utility table for the node "recommendations". It involves the type of recommendation, gas consumption, condition of the flowline and the production rate. If the flowline is OK and the gas consumption low, changing the flowline includes higher costs then benefits (negative utility). Whereas if a high gas consumption with fluctuating pressures led to a low production rate, that can be increased by exchanging the equipment the gain may be large.



Recommendation	GasConsumption	PipelineCond	ProdRate	Y
ChangeFlowline	low	OK	low	-10
ChangeFlowline	low	OK	mid	-10
ChangeFlowline	low	OK	high	-10
ChangeFlowline	low	fluctuations	low	150
ChangeFlowline	low	fluctuations	mid	100
ChangeFlowline	low	fluctuations	high	40
ChangeFlowline	low	wrongdiameter	low	150
ChangeFlowline	low	wrongdiameter	mid	100
ChangeFlowline	low	wrongdiameter	high	100
ChangeFlowline	mid	OK	low	-10
ChangeFlowline	mid	OK	mid	-10
ChangeFlowline	mid	OK	high	-10
ChangeFlowline	mid	fluctuations	low	160
ChangeFlowline	mid	fluctuations	mid	120

Figure 36: Utility table for the utility node Y.

The Bayesian network calculates a rather high utility for the action "exchange of flowline" as well as for the "change of compressor". Increasing the initiation pressure is also recommended since its utility is greater than 0 in both cases. A work over operation is not recommended if the pressure test is not conducted, since the costs of such an operations would exceed the unlikely benefits of increased production.

2.4.4 Advisor in Practice

For the real engineering application, all three parts of the Knowledge Layer, the Event Detector, Problem Classifier and Decision Supporter will be *interconnected* systems that do not require separate approaches for state selection, etc. The nodes can also be continuous to incorporate the actual values of the parameter, e.g. gas consumption 300 m^3/m^3, instead of high – medium - low states. This reduces the engineer's subjectivity when defining different states.

The Event Detector enables an *automatic recognition* of different states, e.g. a slope change of the pressure curve in the height of the gas lift valve, such that the workload of the user can be reduced.

Utilities can be calculated more accurately by using functions that include costs of operations, oil price, expected changes in production volumes, etc. The unit can either be a monetary value or another index that includes non-monetary values, like environmental issues, company strategy, etc.

For an actual engineering application the Bayesian networks are much more complex than the shown examples. A *dynamic* component of the network allows a prediction of the future, hence includes a timeline and an "urgency" component. The recommendation could then include conditions, like change initiation pressure only if flowline and compressor are exchanged before, to ensure that enough gas is available.

2.5 Implementation

2.5.1 User Interface

The Adaptive Advisory System must be designed so simply and intuitively that a person of every age and no computer skills can use it without difficulties or training.

The nature of the user interface proposed here is based on the latest development of social media platform and is revolutionary in context of technical applications, since it represents a shift from rigid purely technical hierarchical structures to a network structure that encompasses not only all hardware of oil and gas operations but also closely links people and technology. It may seems unprofessional to include "playing around tools for the bored youth" into this work, but these tools have shown a very efficient way to manage the complex world of social relationships and they will sustainably change the way we are using new web based developments in areas of our lives that we would have never connected with these technologies a view years ago. Slowly the web 2.0 generation will grow into the working age and revolutionize the way we are making use of computers and artificial social and technical networks within companies.

The smartest new software developments and technological often fail when it comes to user acceptance. Often, the highly sophisticated software simply is too complicated for the user, hence the barrier of getting familiar with using the software and the working principles are too high, which leads to substantial investments being made in vain.

Highly sophisticated software does not necessarily require highly sophisticated software users. Quite the opposite, software developers have to pay special attention to the user friendliness of the interface, which includes an intuitive understanding, clarity of the functionality and customizability.

The aging workforce of the petroleum industry has a special implication on the user interfaces, since employees in their 50ies have a different attitude to computers than fresh recruits that grew up with new technologies. Software for technical applications mostly focuses on the correctness of results; the user perception is given a back seat.

The Adaptive Advisory System learns from interactions with the user. Important here are the time of usage, the variety of problems it is used for and the interconnection with tasks next to the pure engineering, like accounting, budgeting or purchasing. The more often the Advisor is consulted for a wide variety of problems, the larger the data base and the more accurate recommended solutions and results get. Next to the engineers, managers or other personnel like accountants might use the Advisor for certain tasks like to access single data points or make an analysis on performance. The size of the network of people that use the system is crucial to its success. The user interface must be designed also in a way that these people do not require special training to use the Advisor, to ensure the biggest possible interaction between employees and between employees and the Advisor.

2.5.2 Developments of Web 2.0

Within a few years, the usage of the Internet has become an integral part of our daily lives. Some years ago, we have checked our mails every other day; now we are online 24 hours using our home computers, smart phones or tablet computers. The Internet coverage has reached 30% of the total population worldwide and almost 80 % in countries like the US, Germany or Austria.[138] Not only the amount of people using the Internet has increased

[138] The World Bank Group (2011)

tremendously over the last years, but also time we spent using it per day has gone up.[139] Web 2.0 or the 'participative web' including *Facebook*, *MySpace*, *Twitter*, *LinkedIn* and other so called social media platforms have strongly participated to as well as benefited from this development.

Facebook, a social networking service, which has reached more than 750 million users from its launch in 2004 to July 2011 after showing exponential growth rates[140], is a good example for the success factors of social media: The dynamics of the social network attract new people on with an even bigger rate, the bigger the network community has grown. Also the time that users spend on Facebook increases with their personal network. It allows the naturally curious users to retrieve information, or gossip, anonymously, which is a truly addictive feature.[141]

Facebook has users among all age groups. Once entering the page it only takes a few clicks to feel familiar with the functionality. Hence another key factor of Facebook's success is its intuitive and easy user interface. The structure of the website is less hierarchical but very interconnected. All names, pictures, photo albums, and external applications are highlighted and include a link that automatically leads to the topic we'd like to investigate. The user can customize his / her personal news wall, by applying filters, following (like, add friend) friends, companies, news providers and other topics of interest.

2.5.3 Instruments

Looking at the overwhelming success of Facebook, Twitter and other social media, there are some lessons companies can learn from them when designing new tools for technical problem solving and knowledge management:

Private Profiles and Participation

If every user of the Adaptive Advisory System is given his/her personal profile the user acceptance, transparency and knowledge management in general will greatly benefit from it. Actions taken are connected to names and faces, hence others can identify experts in certain areas and immediately know whom to address for certain issues. Users can leave comments or pose questions.

There have been knowledge management efforts in creating tools for company wide portals for lessons learned and knowledge sharing among knowledge groups in petroleum companies.[142] Integrating these tools into a system that is in daily use of the engineers and experts would help to solve the problem of user acceptance and participation. The possibility to "comment" on literally everything in the Advisor's interface allows to make individual talks public and to record discussions and solution finding processes.

A side effect of the personal profile can be obtained by the utilization of the addictive nature of social media: The more "personal content" the users share with each other with this tool (e.g. pictures of rigs, family, status updates like "I am on holiday for the next two weeks", and a company internal network of friendship), the more time they spend with the software. One might argue that they might spend more time with the personal data than with the technical data, but as long they are logged into the Advisor and receive alarms and

[139] Beifuss (2008)
[140] Facebook (2011)
[141] Cohen (2009)
[142] For Example: Halliburton, Smith (2007)

follow activities in the assets, the overall benefits are largely increased under the aspects of knowledge management.

Figure 37 shows schematically how the GUI of the Advisor could look like. In the "News" page it lists a selection of tags concerning assets within the engineer's responsibilities and from a certain kind (here: Alarms). Several high level pages can be selected (News, Map, BN, Actions, Analysis, Users) and accordingly subsequent lower level pages (e.g. different stages of the Knowledge Layer BNs). The privacy setting and displayed information of the user are solely in his/her personal responsibility.

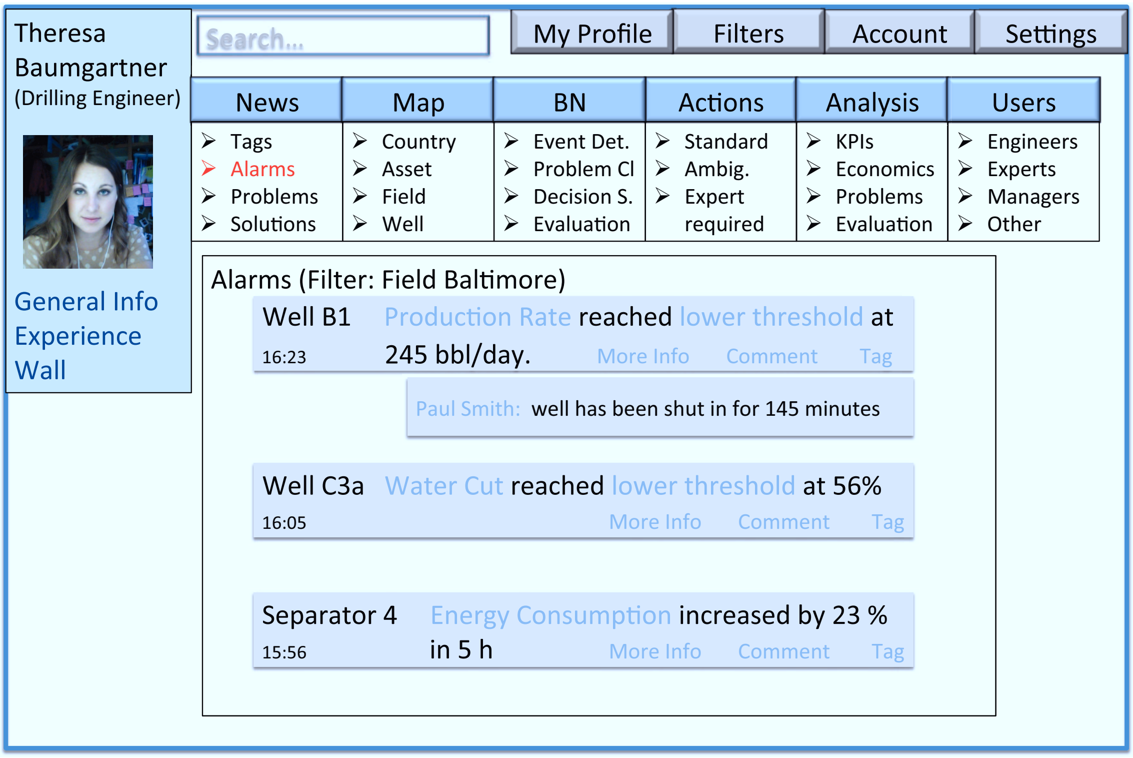


Figure 37: Schematics of the GUI

Hardware Profiles

Just like every user has a profile with important information, the same is valid for assets, fields, wells, machines, etc. according to the level of detail of the operator. The individual profiles contain detailed technical information as well as links to events, problems or actions taken always connected to the person that has taken the action. This shows clearly and publicly related responsibilities and who to address for certain issues.

For example the profile of Well B1 contains an information page with a summery and details on the drilling process and links to the drilling report, on the completion and all relevant information to events and problems that have been identified in correlation with Well B1. The profile page of Well B1 contains actions taken for the lifetime of the well, such as shut-ins, work-overs, etc. in chronological order. Dynamic filters and the search function enable quicker and more precise results.

A *tag cloud* is visual representation of text data, which shows the weighted connection of tags. Important (often used) connections are highlighted by colour or front size. In Figure 38 a tag cloud on the profile shows the connections of the Well B1 within the network.

One glance allows the user to associate the well with a certain set of the most probable issues, responsible people and other similar wells.

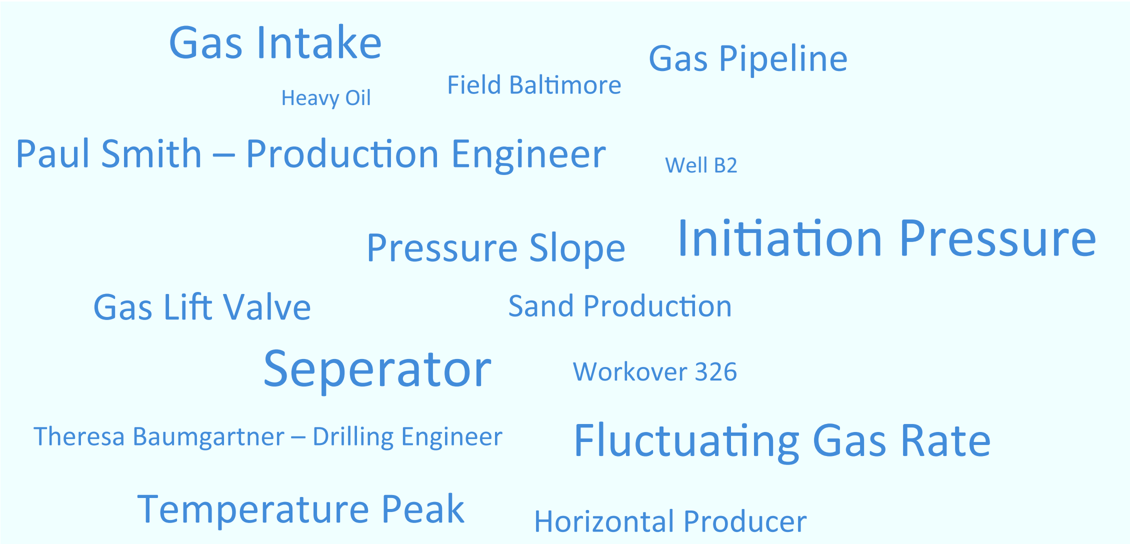


Figure 38: Exemplary tag cloud for well B1

Personalization

"*Create your own Advisor*" might be a motto for the graphical user interface of the Adaptive Advisory System. No single user would see the same things when logging into the system. The interface consists of different element that users display or hide ad libitum depending on their responsibilities and personal preferences.

Personalization of the Advisor is one method to overcome the data flood. Users *subscribe*[143] to the single or groups of hardware profiles according to their personal responsibilities, i.e. countries, assets, wells, fields, problem areas, etc. and according to the level of detail they require information, i.e. country – wide for managers vs. specific tools for engineers. Once subscribed, the users get the latest information on their personal news wall and alarmed by phone, pager, etc. only in their personal area of responsibility.

The Cloud

Cloud computing leads the way towards a computing service rather than seeing computers and network connections as a product. It enables to share computational power, storage capacity, software applications, files and other data via the Internet or Intranet connection, enabling a real time access from any place in the world by any number of people. Cloud computing is currently facing a huge hype in the IT business. IDC estimates that by 2010, 15 % of all digital information in the 'Digital Universe' will be part of a cloud service.[144]

Computational power storage capacities of a whole company could be combined to deal with peak loads, hence cloud computing enables us to reduce the calculation times of complex and huge networks.

Ideally, the Adaptive Advisory System is a web application that allows access from all locations worldwide by computer, laptop, tablet computer or smart phone with only the prerequisite of an Internet connection. Employees can enter their profile and track the opera-

[143] compare to "follow" – Twitter or "like" - Facebook
[144] Digital Universe Study (2010), p. 4

tions on business trips or comment in real time when visiting the rig site and they can use any device for their routine work without being dependent on a specific workstation.

2.5.4 Transparency

When it comes to the implementation of the software, one very important factor for the success of the Adaptive Advisor and a competitive advantage for the company is transparency of the operations. As shown in Figure 39 three main areas of transparency: software, models & calculations, actions & evaluation.

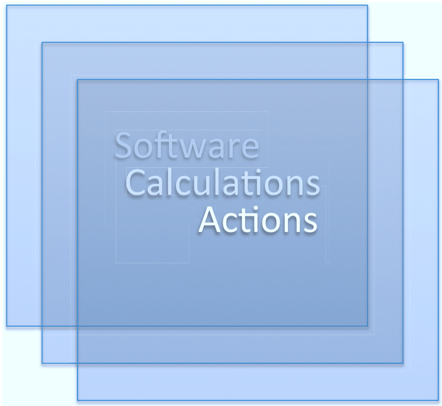


Figure 39: Areas of transparency

Transparency of Software

Most software that is used today is protected and all rights are reserved; hence the code is hidden and can neither be understood nor adapted by IT proficient users. In contrast *open source software* allows its users to study, change and improve[145] the software.

The implementation of the Adaptive Advisory System in different companies sets the tasks of creating interfaces with any different applications and other company specific requirements that may change with time. Allowing the employees of a company to re-program parts of the software of the Advisory System is the optimum way to ensure the truly adaptive nature of the software. People that actually use the software on a daily basis have the best ideas for improvements and how to implement them. However, the changes a software user makes can in turn also not be hidden from the original software developers, who then supervise and direct the evolution of the software. Hence the basics of the software can be improved in a much more innovative and faster pace since a lot of different experiences and opinions can be included.

Creating artificial walls by protecting a product with patents and other regulations creates a short time advantage, since there is a gap in knowledge between the developers and other companies. But this advantage is not sustainable, since it causes the developers to rest on their achievements. Opening up and facing competition creates a constant pressure to be at the cutting edge of technological development and to constantly improve and innovate. In Chapter 1.1.1 Learning, we heard that the only sustainable competitive advantage is to learn how to learn faster. Open source software ensures exactly this principle and hence is the optimum strategy for a truly Adaptive Advisory System in an ever-changing environment.

[145] At times also the distribution of the software is allowed, which is considered an option here.

Transparency of Calculations & Models

The Advisor can be classified as a *decision support* or *expert system*. The name expert system suggests that there is a computer that acts like a human expert with all the good and bad features connected with experts; we automatically assume that they are people with in-depth knowledge on a particular subject and rarely are wrong with their analyses and recommendations. A recent neurobiological study even showed that brain activity was reduced for financial decision-making if people got advice by financial experts.[146]

But seeing the Advisor like an expert is a wrong perception of the system. It is developed, programmed and used by humans; the data it bases its knowledge on was finally created by humans; and humans make mistakes. The system is designed to complement not to replace the human engineer and it lives from the collaboration of the people and technology. Hence we want the users of the system to keep thinking, understanding the processes on the field as well as within the Advisor to constantly update and improve the calculations.

Hence, like described in the Guided Workflow, the users have easy access to all underlying calculations, models, the Bayesian network structures, data and probabilities used in the Adaptive Advisory System. Sharing experiences and knowledge among assets, companies or countries can be a solution where data is typically rare.

Transparency of Actions & Evaluations

The interface of the Advisor allows to directly connect actions, recommendations, calculation, comments, etc. with the user that was responsible for it. There may occur many problems related to aspects of a company's culture. If all actions are public (displayed to other employees) than also mistakes are public and clearly assignable. Since all other employees have the right to comment on others actions, other employees will question expert's actions and opinions much more. The gain in transparency helps the whole organization to learn, improve and increase the efficiency of their daily processes.

To fully benefit from transparency it is crucial to change the organization's attitude to mistakes. Using the Adaptive Advisory System, a decision can either be in line or not in line with the recommendations of the system. If the decision is wrong but it was in line, then it is the system and system supervisors that have to take the responsibility of the decision, not the engineer alone. If it was not in line with the system, there should be a policy to discuss actions with the management or other experts, obtain approval and adapt the system if necessary.

The transparency of engineer's actions greatly adds value to the company. In the short run, a mistake in most cases means lost money for the company. But in the long run, a mistake has a great savings potential, if and only if the company manages to learn from it and prevent others to make the same mistake. Therefore, establishing a "no-blame" culture in an organization together with the implementation of transparent systems is absolutely essential.

Transparency is also critical for the management. Bill Hostmann[147], analyst at Gartner Inc., says that managers don't want other people to look at their numbers. Transparency would mean that everybody else could generate the same KPIs or other sets of data and therefore reduce the management's raison d'être. But Hostmann argues that transparency correlates with better business performance.

[146] Engelmann, Capra, Noussair, Berns (2009)
[147] Tucci (2009), p. 10

3 CONSTRAINTS & CONCLUSION

3.1 Constraints

This work shows how an automated expert system can be used to improve decision-making and reduce the times for it in a petroleum asset team. The system is described and the technical outline for the system is given. Although the technology and tools for decision support systems are already in place since decades, they are not as widely implemented, as one would expect. Considering this fact it becomes clear that the question of implementing a decision support system is not just a technical one. The human factor turns out to be crucial when it comes to the implementation of a system like that. In general there are technical as well as human challenges in the implementation of automated Adaptive Advisory Systems in oil and gas operations.

3.1.1 Technology

Methods of artificial intelligence have been around for decades. AI has made constant and amazing progress in certain areas, but AI in practise seems to fall short of its expectations. One reason for this might lie in the often too high aspirations. Hollywood subtly suggests that AI is capable of creating machines undistinguishable from human beings or we use the human intelligence as a yardstick for machine intelligence. But we still are far from even being close to these expectations, not even having solved all miracles of the human brain. Technologically it will never be possible to exactly reproduce what evolution has been created in millions of years in all precision.

A good example for the unexplainable and genius sensitivity of the brain is the story of the Getty kouros, a dolomitic marble statue bought by the J. Paul Getty Museum in California in 1985. Since then, the Getty kouros has caused scholars, dealers and scientists to choose sides in a heated public debate about whether the work was carved by a late sixth century B.C. Greek sculptor or by a modern forger.[148] Experts claim that “there is something wrong with the statue” when having a glance at it, but they cannot tell what is wrong with it and why they think so.[149] Even after many years technical analysis could prove or disprove the forgery theory. Its status remains undetermined: today the museum's label reads "Greek, about 530 B.C., or modern forgery”[150]

All pieces needed for an Adaptive Advisory System such as the one described in this chapter are already well published and available. Nevertheless there are challenges in the definition of a comprehensive and reliable ontology of challenges and problems in production operations and reservoir management. Besides, even if a perfect ontology was created, the system still relies on quality and availability of data inputs necessary for the reasoning process. It therefore is recommended to start with a subset of challenges, starting with the most frequent ones and let the system evolve as new information comes in.

Security aspects are of high importance, especially when using the benefits of the cloud service. Data in connection with oil and gas operations is often sensitive in terms of secrecy. The more data is handled via a public system like the Internet, the bigger the possibilities for attacks are and therefor security efforts must be greatly increased. Nonetheless it would

[148] Kimmelmann (1991)
[149] Dijksterhuis (2010)
[150] J. Paul Getty Museum (2011)

be wrong to miss these wonderful opportunities of easy access and sharing knowledge all across the world because of security restraints.

3.1.2 The Human Factor

The technology and scope of research already existing on all forms of artificial intelligence, expert systems, decision support systems and other smart systems is of surprising volume. But in practice, there is little known about doctors using medical support tools with artificial intelligence or companies monitoring their production on a large scale with intelligent systems. Slowly, fuzzy systems are implemented in some web pages and other software, but the full range of its uses for business processes and whole organizations can rarely be seen. So why has no real breakthrough been reached yet? As always, the answer lies in the seemingly most complicated systems of all: the human beings and their interactions with each other.

The *change management* aspect needs to be considered properly in order for an Advisory system implementation to be an organizational success. Workflows and processes need to be in place that support the use of an Advisory System. In contrast to conventional operations workflows are not isolated anymore and highly benefit from intense collaboration and networking among the different disciplines, assets and functions.

The system is designed to *complement* and not to replace the daily work of petroleum experts. It learns from input received from various groups and experts and could neither exist nor evolve without human interaction. Another obstacle in the acceptance of such as system is a possible reluctance to participate in a fully transparent system due to a fear of publication of personal failure. It is therefore extremely important to fully commit to a no blame culture and to establish an environment, where strong competition does not hinder exchange of lessons learned, knowledge and ideas. It will take a while until the system has been calibrated well enough to gain the trust of the experts in an asset team. However, in contrast to other systems, the here described Advisor operates fully transparently and all steps, results and recommendations are fully traceable, which together with the integration of the knowledge sharing tool in a daily used system, may reduce the time to acceptance within an organization.

Another very important issue for the success or failure is the availability of enough good quality data. A system can be designed in a very smart way and promise good results in theory. But in practice the output is worth exactly the same as the input, or in other words: garbage in, garbage out. Although the Advisor is equipped for dealing with uncertainties, missing, noisy or redundant data, its intelligence cannot fully make up for wrong inputs. When it comes to data delivered by sensors, the technical issues can more easily be solved, than problems in areas where it is the employees' responsibility to make the right statements on their operations. On a rig site, for example, people don't have the time, nor the understanding of the importance, to correctly record a huge amount of numbers each day, and not just copy it from previous days, to save a lot of time in an challenging and stressful environment. Easy to handle and supportive systems must be in place to assets and help the employees rather than to "steal their time".

3.1.3 Responsibility

Technological developments always have been sensitive tools that have given its owners enormous *power* over others. History shows that this power often has been misused. The transparency that an Adaptive Advisor would create in an organisation involves some dangers next to its opportunities; therefore developers, owners and users have to act with caution and responsibility to minimize the risks of abuse.

The system theoretically enables the recording of all steps, clicks, and other interaction of the user. This transparency may lead to a "*Big Brother Effect*", which means that a few people supervise and analyse actions of the users and that these actions have an effect on the user's future situation in the organisation. One has to be to be adamantly opposed to the implementation of features in the software that allow the tracking of users without them being aware of this. The developers of the software are especially addressed to show responsibility in this case.

Next to ethical reasons there are also economic reasons for companies to build on trust rather than on control of employees. Only an open, participative and trustful community can tap the full potential of an Advisory System and the resulting changes of business processes. Of employees feel observed and think that mistakes will have large negative effects on their careers, they will always find means to work around a technical software, leading to a culture of secrecy. For mistakes rates we can differentiate a hidden mistake rate and an open mistake rate (for obvious and non-concealable mistakes); if a transparent system is implemented in an organisation with a secretive and blaming culture, the open mistake rate might go down, but the rate of hidden mistakes would strongly increase, which is much more harmful in the long run, since the management is not aware and cannot take actions against it.

Also, there is a general risk in analysing people by the single point signs of an interaction with a system: human beings are much more complex than the algorithms designed to classify them could ever be. If such a system is applied to people (e.g. workforce on rigs, etc.) to predict their actions, it may works well for the bulk of people, but how about the hand full of people that buck the general trends? Often, people that are "out of the box" are then discriminated although their non-conformity would add a lot of value to companies.

3.2 Conclusion

There exist many successful examples of the implementation of artificial intelligence for decision support systems. In the oil industry, only isolated questions like infill-drilling well locations, optimum completion solutions, or corrosion analysis have been targeted with Bayesian networks or fuzzy logic so far. The Adaptive Advisor follows a higher-level approach: to integrate the whole real-time operations business process for oil and gas assets into one single system and therefore bring together various processes and disciplines. Starting with well management on a general level, a company can in a next step further integrate more and more technical details such as well drilling problems or completion decisions.

The basics of knowledge theory are described in the beginning of this work to see the impact of the application of the principles of learning to a business process. It explains that learning is the accumulation of knowledge. Knowledge is a stock and the flow of learning lets it grow at a certain rate. This concept brings value to a company, because learning how to learn faster is the most sustainable competitive advantage that companies can create in the dynamic and quickly changing environment of today. Data, information and knowledge form a pyramid that can as well be found in the build up of the advisor. The work tries to define these terms for a computational use and to draw borders between them.

After some ups and downs, artificial intelligence has gained a lot of attention within the last years. A breakthrough, and hence a more frequent implementation in software developments could lead to a revolution of information technology (IT), that has not substantially changed in the last decades. A shift towards knowledge technology (KT) means to implement smart software, network structures and collaborative and transparent systems that would help to solve problems of data floods, information overkill and loss of knowledge.

Not all the knowledge the world owns today has been captured in books; a large part of knowledge exists in form of tacit knowledge mainly in our heads. Due to its aging workforce and resulting high fluctuation rates, in particular the oil industry faces difficulties nowadays to keep knowledge in companies. This work addressed how to transfer tacit knowledge into explicit knowledge that is easily accessible, sharable and re-applicable by people or machines.

In order to find out how to do this, we looked at the best-designed and longest proven example of a learning system: the human brain. Our brains are a network of millions of little interconnected cells; one cell by itself cannot do anything, but the sum of all achieves amazing results. Artificial intelligence in general and artificial neural networks in particular aim at modelling the brain and reproducing neurobiological principles to create smart computer programs.

It has been shown that one instrument of artificial intelligence, Bayesian networks, work specifically well for the requirements of the Adaptive Advisory System, because they can very well deal with uncertainties, are capable of inferring hidden influences, are flexible and adaptable and they can easily combine artificial learning from data and capturing the knowledge of human expert. For large and complex BNs it has shown to be easier to use the generative form of the Bayesian network than the discriminative form, since the nodes and probabilities can more easily be split up into different expertise groups.

The Adaptive Advisory System has incorporated findings from knowledge theory, namely a hierarchy of data, information and knowledge and Bayesian network methodologies in its structures and operating modes. Examples demonstrate the application on isolated tasks in well management. The system continuously learns with every interaction between system and user, results will become more and more accurate with time; therefor it is designed to complement and assist humans, not to replace them. Crucial for a successful implementation of the system in an organisation are nonetheless “soft factors”: a no-blame culture and transparent and collaborating structures and processes.

Artificial intelligence with the right computational power is capable of modelling and assessing processes, which complexity is beyond what humans can understand and capture with our brains. There is a lot of room for research in this area but the results and the successful demonstration will amaze us in the upcoming years. Nonetheless, there will be long way to go, if ever, until we can create something that comes even close to what evolution has been created with our brains. Humans have been successful problem solvers because we are capable more than knowledge: our creativity and imagination; or as one of the greatest scientists of all times, Albert Einstein, puts it:

“I believe in intuitions and inspirations. I sometimes *feel* I am right, but do not *know* it. […] I'm enough of an artist to draw freely on my imagination, which I think is more important than knowledge. Knowledge is limited. Imagination encircles the world." [151]

[151] Viereck (1929)

LIST OF REFERENCES

Aamodt, A, Mygard, M, 'Different roles and mutual dependencies of data, information and knowledge – an AI perspective on their integration' in *Data and Knowledge Engineering* 16 (1995) 191 – 222, Elsevier.

Agbon, I. S., et al., 'Fuzzy Ranking of Gas Exploitation Opportunities in Mature Oil Fields in Eastern Venezuela', paper SPE 84337, presented at the SPE Annual Technical Conference and Exhibition, Denver, Colorado, 5 – 8 October 2003.

Ahlert, M., Blaich, G., Spelsiek, J., '*Vernetztes Wissen. Organisationale, motivationale, kognitive und technologische Aspekte des Wissensmanagements in Unternehmensnetzwerken*', Deutscher Universitäts-Verlag, Wiesbaden, Germany, 2006.

AIPA, *Survey* by SPE Communities - Digital Energy Technical Section Artificial Intelligence and Petroleum, retrieved 4 October 2011, <http://research.spe.org/se.ashx?s=705E3F13349DEB59>

Al-Yami, A.S, et al.: 'Expert System for the Optimal Design and Execution of Successful Completion Practices Using Artificial Bayesian Intelligence', paper SPE 143826, presented at the Brazil Offshore Conference and Exhibition, Macaé, Brazil, 14 - 17 June 2011.

Alhajeri, M.M, 'Reservoir Analysis on Applications of Gelled Polymer Technology in the Arbuckle Formation', MS Thesis, the University of Kansas, 2006.

Angelov, P. P., *Evolving Rule-Based Models. A Tool for Design of Flexible Adaptive Systems*, Physica-Verlag Heidelberg New York, 2002.

Beifuss, A., 'Internet and Web 2.0 in Germany', *Goethe Institut 2008*, retrieved 15 September 2011, <http://www.goethe.de/wis/med/ein/en18305.htm>.

Bellhouse, D. R., 'The Reverend Thomas Bayes, FRS: A Biography to Celebrate the Tercentenary of His Birth', *Statistical Science*, 2004, vol. 19, no. 1, 3 – 43. Retrieved of 3 July 2011, <http://www.york.ac.uk/depts/maths/histstat/bayesbiog.pdf>,

Bidyuk, P. I., Terent'ev, A. N., Gasanov, A. S., 'Construction and Methods of Learning of Bayesian Networks', *Cybernetics and Systems Analysis*, vol. 41, no. 4, 2005, p. 587 – 598.

Boahen, K., on 'A Computer that Works Like a Brain', in *TED*, 2007, retrieved 28 July 2011,<http://www.ted.com/talks/lang/eng/kwabena_boahen_on_a_computer_that_works_like_the_brain.html>

Bodendorf, F., *Daten- und Wissensmanagement*, Springer Verlag, Berlin Heidelberg New York, 2006.

Borgelt, C., et al*., Neuro-Fuzzy-Systeme. Von den Grundlagen künstlicher Neuronaler Netze zur Koppelung mit Fuzzy – Systemen*, Vieweg Verlag, Wiesbaden, Germany, 2003.

Brulé, M. 'Using Massively Parallel Processing Databases', *digital energy journal*, September 2009.

Brulé, M., Charalambous, Y., Crawford, M. L., Crawley, C., 'Bridging the Gap Between Real-Time Optimization and Information-Based Technologies', SPE Paper 116758, presented at the SPE Annual Technical Conference and Exhibition, Denver, Colorado, 21 – 24 September 2008.

Cohen, E., ‘Five clues that you are addicted to Facebook’, *CNN*, April 2009, retrieved on 15 September 2011, <http://articles.cnn.com/2009-04-23/health/ep.facebook.addict _1_ facebook-page-facebook-world-social-networking?_s=PM:HEALTH>.

Correa, M., Bielza, C., Pamies-Teixeira, J., ‘Comparison of Bayesian networks and artificial neural networks for quality detection in a machining process’ in *Expert Systems with Applications*, Elsevier Ltd., 2007, retrieved 19 September 2011, <http://www.dec.fct.unl.pt/projectos/SMA_2008/Bib/CBPT08.pdf>.

De la Vega, E., Sandoval, G., Garcia, M., Nunez, G., Al-Kinani, A., ‘Integrating Data Mining and Expert Knowledge for an Artificial Lift Advisory System’, paper SPE 128636, presented at the SPE Intelligent Energy Conference and Exhibition held in Utrecht, The Netherlands, 23 – 25 March 2010.

Dijksterhuis, A., Nordgren, L. F., ‘A Theory of Unconscious Though’, *Perspectives of Psychological Science*, 1, 95 – 109, 2006.

Dijsterhuis, A., *Das Kluge Unbewusste, Denken mit Gefühl und Intuition*, Translation: Hildegard Rohr, Klett-Cotta, Stuttgart, 2010.

Drogseth, D., et al., ‘Gleaning Wisdom from Data, The role of service knowledge management in IT decision-making and ITIL® V3’, *White Paper* by BMC Software and Enterprise Management Associates, retrieved 2 August 2011, < http://documents.bmc.com/products/documents/12/90/91290/91290.pdf>

Druzdzel, M.J, Flynn R.R., ‘Decision Support Systems’, *Encyclopedia of Library and Information Science*, Second Edidition, Allen Kent (ed.), New York: Marcel Dekker, Inc., 2002.

Edeko, F.O, Mbamula, J.E., ‘Application of Bayes’ Model to Oil/Gas Processing Network Failures’, *Corrosion 2004*, paper 04057 published by NACE International, Huston, Texas, 2004.

Engelmann, J. B., Capra, M. C., Noussair, C., Berns, G. S., ‘Expert Financial Advice Neurobiologically “Offloads” Financial Decision-Making under Risk’, PLoS ONE 4(3): e4957. doi:10.1371/journal.pone.0004957, retrieved 21 September 2011, <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0004957>

Facebook, Timeline, retrieved 15 September 2011, <https://www.facebook.com/press/info.php?timeline>.

Fienberg, S.E., ’When Did Bayesian Inference Become “Bayesian”?’ in *Bayesian Analysis*, 2006, vol. 1, no. 1, p. 1- 40.

Friedman, N., Geiger, D., Goldszmidt, M. ‘Bayesian Network Classifiers’, *Machine Learning,* 29, p. 131 - 163, Kluwer Academic Publishers.

Flinn, S. D., *The learning layer: Building the next level of intellect in your organization,* Palgrave MacMillan, New York, 2010.

Fuller, R., ‘Neural Fuzzy Systems’, *Lecture Notes*, Abo Academi University, Finland, 2005. Retrieved 26 July 2011, < http://users.abo.fi/rfuller/ln1.pdf >

Fyfe, C., ‘Artificial Neural Networks’ in *Do Smart Adaptive Systems Exist?,* Gabrys, B., Springer Berlin Heidelberg, 2005, p. 57 – 80.

Gabrys, B. et al., *Do Smart Adaptive Systems Exist? A Best-Practice for Selection and Combination of Intelligent Methods*, Springer-Verlag Berlin Heidelberg, 2005.

Garibaldi, J. M., ‘Fuzzy Expert Systems’ in *Do Smart Adaptive Systems Exist?*, Gabrys, B., Springer Berlin Heidelberg, 2005, p. 105 – 132.

Garvin, D. A., *Learning in Action: A Guide to Putting the Learning Organizations to Work*. Harvard Business School Press, Boston, 2000.

Gayle, R. ‚The DIKW Model for Innovation', on *spreadingscience.com*, retrieved 2 August 2011, <http://www.spreadingscience.com/our-approach/diffusion-of-innovations-in-a-community/1-the-dikw-model-of-innovation/>

Ghahramani, Z., 'An Introduction to Hidden Markov Models and Bayesian Networks.' *International Journal of Pattern Recognition and Artificial Intelligence*. 2001, vol. 15, p. 9 - 42.

Ghoraishy, S. M., et al.: 'Application of Bayesian Networks for Predicting the Performance of Gel-Treated Wells in the Arbuckle Formation, Kansas', paper SPE 113401, presented at the SPE/DEO Improved Oil Recovery Symposium, Tulsa, Oklahoma, 19 - 23 April 2008.

Giese, M., Bratvold, R.B., 'Probabilistic Modeling for Decision Support in Integrated Operations', paper SPE 127761, presented at the SPE Intelligent Energy Conference and Exhibition held in Utrecht, The Netherlands, 23 – 25 March 2010.

Hatch, D., 'Operational Business Intelligence. Five Things Need to Know About Improving Your Costumer-Facing Performance', March 2009, *Aberdeen Group*, retrieved 10 August 2011, <https://www1.vtrenz.net/imarkownerfiles/ownerassets/1076/5392-RA-Operational_BI-DH-04-NSP1.PDF>

Hite, J. R., Crawley, C., Deaton, D. F., Farid, K., Sterneskey, M., 'Barriers to Implementation of Real-Time Optimization Technology' Paper SPE 110236 presented at SPE Annual Technical Conference and Exhibition, Anaheim, California, 11 - 14 November 2007.

Hite, J.R., Crawley, C., Deaton, D. F., Farid, K., Sternesky, M., 'Barriers to Implementation of Real-Time Optimization Technology', Paper SPE 110236 presented at the SPE Annual Technical Conference and Exhihition, Anaheim, California, 11 – 14 November 2007.

Hollifield, B. R., Habibi, E., *Alarm management: Seven effective methods for optimum performance*, ISA, USA, 2007.

Hooimeijer, M., Azmi, M., 'Advanced Production Monitoring', Paper SPE 104161, presented at the SPE International Oil & Gas Conference and Exhibition, Bejing, China, 5 – 7 December 2006.

Horvitz, E., Berry, M., 'A Decision-Theoretic Approach to the Display of Information for Time-Critical Decisions: The Vista Project', in *Proceedings of SOAR-92*, NASA/Johnson Space Center, Houston, TX, August 4-6, 1992.

IDC, *2010 Digital Universe Study*, April 2010. Retrieved 3 January 2011, <http://www.emc.com/collateral/demos/microsites/idc-digital-universe/iview.htm>.

J. Paul Getty Museum, *Statue of a Kouous*, retrieved September 24 2011, <http://www.getty.edu/art/gettyguide/artObjectDetails?artobj=12908>

Jain, A. K, Murty, M. N., Flynn, P. J., 'Data Clustering: A Review', in *ACM Computing Surveys*, vol. 31, no. 3, September 1999.

Jeffreys, H., *Theory of Probability*, Oxford University Press, 1939.

Jensen F. V., Nielsen T. D., *Probabilistic decision graphs for optimization under uncertainty*. Springer – Verlag, New York, 2001.

Jensen, F. V., Nielsen, T. D., *Bayesian Networks and Decision Graphs*, Springer New York, 2007.

Keller, T. A., Just, M. A., 'Altering Cortical Connectivity: Remediation-Induced Changes in the White Matter of Poor Readers'. *Neuron,* vol. 64, 2009, p. 624 – 631.

Kerman, M. C., Jiang, W., Blumberg, A. F., & Buttrey, S. E., '*Event Detection Challenges, Methods, and Applications in Natural and Artificial Systems.'* Lockheed Martin, Naval Postgraduate School, Monterey, CA, 2009.

Kimmelmann, M., 'ART; Absolutely Real? Absolutely Fake?', The New York Times, 4 August 1991, retrieved 24 September 2011, <http://www.nytimes.com/1991/08/04/arts/art-absolutely-real-absolutely-fake.html?src=pm>

Kuzmich, B., 'Thinking business in the oil field', in *Next Generation Oil And Gas*, Q4 2009, retrieved 5 August 2011, <http://www.ngoilgas.com/article/Thinking-business-in-the-oil-field/>

Leavitt, H. J., Whisler, T. L., 'Management in the 1980's', *Harvard Business Review*, Nov. – Dez., 1958, p. 128 - 136.

LeVine, S., 'Exxon: Juggernaut or Dinosaur?', *BusinessWeek*, p. 1, 5 February 2009, <http://www.businessweek.com/magazine/content/09_07/b4119034606428_page_1.htm>, retrieved 15 April 2010.

Loredo, T.J., 'From Laplace to supernova sn 1987a: Bayesian Inference in Astrophysics', *Maximum Entropy and Bayesian Methods*, Kluwer Academic Publishers, 1990, Dordsrecht, The Netherlands, p. 81 – 142.

Luckham, D., *The power of events: An introduction to complex event processing in distributed enterprise systems*, Addison Wesley, Boston, 2002.

MacKay, D.J.C, 'Bayesian Methods for Adaptive Models', *Thesis 1992*, California Institute of Technology

MarkCC, 'Fuzzy Logic vs. Probability', under Fuzzy Logic, 2011, *Scientopia.org/blogs*, retrieved 26 July 2011, <http://scientopia.org/blogs/goodmath/2011/02/02/fuzzy-logic-vs-probability/>

Meyer, M. E., Gudykunst, W. B, 'The effect of time pressure and type of information on decision quality', *The Southern Communication Journal*, 1997, 62, p. 280 - 292.

Michalek, G., 'Expert Systems', *Carnegie Mellon University Libraries*, May 2001, <http://shelf1.library.cmu.edu/IMLS/MindModels/video7.html>, viewed on 21 July 2011.

Mitchell, T. M., 'Chapter 1: Generative and Discriminative Classifiers: Naive Bayes and Logistic Regression', in *Machine Learining,* McGraw Hill Companies, Inc., 2010.

Mittal A., Kassim, A.*, Bayesian Network Technologies. Applications and Graphical Models*. IGI Global, 2007.

Moussa, A. S., 'The Implementation of Intelligent QoS Networking by the Development and Utilization of Novel Cross-Disciplinary Soft Computing Theories and Techniques', *dissertation*, submitted to the department of computer science, the Florida State University, 2003.

N.N., 'Introduction to Bayes Nets', *Norsys* 2011, retrieved 10 August 2011, <http://www.norsys.com/tutorials/netica/secA/tut_A1.htm>

N.N., CERA, Cambridge Energy Research Associates, 'The Digital Oil Field of the Future: Enabling Next Generation Reservoir Performance ', May 2003

PATTS – nervous system, retrieved 28 July 2011, <http://webschoolsolutions.com /patts/systems/nervous.htm>

Pearl, J., 'Bayesian Networks: A Model for Self-Activated Memory for Evidential Reasoning', submitted to the Seventh Annual Conference of the Cognitive Science Society, 15-17 August 1985.

Pearl, J., 'Networks of Plausible Inference', *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann Publishers, 1988, ISBN 1-55860-479-0.

Polanyi, M., *Personal Knowledge*. The University of Chicago Press, Chicago, 1958

Pourret, O., 'Introduction to Bayesian Networks', in *Bayesian Networks, A Practical to Applications*, John Wiley & Sons, England 2008.

Rajaieyamchee, M. A., Bratvold, R. B., 'A Decision Analytic Framework for Autonomous Geosteering', Paper SPE 135416, presented at the SPE Annual Technical Conference and Exhibition, Florence, Italy, 19 - 22 September 2010.

Rajaieyamchee, M. A., Bratvold, R. B., 'Bayesian Decision Networks for Optimal Placement for Horizontal Wells', Paper SPE 129984, presented at the SPE EUROPEC/EAGE Annual Conference and Exhibition, Barcelona, Spain, 14 -17 June 2010.

Rostand, A., Soupa, O., 'Benchmark Survey Reveals Looming Talent Shortage', *E&P Magazine,* 1 May 2011, retrieved 8 August 2011, <http://www.epmag.com/Magazine/2011/5/ item81982.php>

Rowley, J., 'The wisdom hierarchy: representations of the DIKW hierarchy' in *Journal of Information Science*, vol. 33, 2007, p. 163–180.

Russel, S., Norvig P., 'The Foundation of Artificial Intelligence', in *Artificial Intelligence – A Modern Approach*, 3rd Edition, Pearson, New Jersey, 2010.

Saeedi, A. et al.: 'Using Neural Networks For Candidate Selection and Well Performance Prediction in Water Shutoff Treatments Using Polymer Gels – A Field Study', paper SPE 101028 presented at the SPE Asia Pacific Oil & Gas Conference and Exhibition, Adelaide, Australia, 11 – 13 September 2006.

Schlumberger, Schlumberger Well Completions and Productivity Chevron Main Pass 313 Optimization Project, Schlumberger Gas Lift Design and Technology, 2000

Shenoy, C, Shenoy, P., 'Bayesian Network Models of Portfolio Risk and Return', 1999, retrieved 9 August 2011, <http://kuscholarworks.ku.edu/dspace/bitstream/1808/161/1/ CF99.pdf>

Skillings, Jonathan, 'Getting Machines to Think Like Us', *CNET*, 3 July 2006. Retrieved 25 July 2011, <http://news.cnet.com/Getting-machines-to-think-like-us/2008-11394_3-6090207.html>

Smith, D. L.: 'The Evolution of Communities of Practice in a Petroleum Service Organization', Paper SPE 108314 presented at the SPE Annual Technical Conference and Exhibition, Anaheim, California, 11 – 14 November 2007.

Solesa, M., 'Case Study K Field', *Lecture Notes for Well Problem Analysis*, University of Leoben, 2004.

Solesa, M., 'Problem Diagnosis Part 503', *Lecture Notes for Well Problem Analysis*, University of Leoben, 2004.

Sperling, E., 'Too Much Data', on *forbes.com*, 16 November 2009, retrieved on 3 August 2011, <http://www.forbes.com/2009/11/14/data-pitney-bowes-technology-cio-network-bourland.html>

Stigler, S.M, 'Who discovered Bayes' Theorem?', *Statistics on the Table: The History of Statistical Concepts and Methods*, Harvard University Press, 1999, Cambridge, MA.

Strasunskas, D., Rajaieyamchee, M. A, 'Decision Analysis Approach to Valuing Integrated Drilling Operations in the North Sea', Paper SPE 136023, presented at the SPE Production and Operations Conference and Exhibition held in Tunis, Tunisia, 8 – 10 June 2010.

The World Bank Group 2011, 'Internet users are people with access to the worldwide network.', International Telecommunication Union, World Telecommunication/ICT Development Report and Database, retrieved 15 September 2011, <http://data.worldbank.org/indicator/IT.NET.USER>.

Thrun, S., & Norvig, P, '*Introduction to Artificial Intelligence*', the Standford University, retrieved August 29, 2011 from Standford Engineering: http://www.ai-class.com/.

Tucci, L., 'Five Tips for A New BI System', *Next Generation BI Strategies*, searchcio.com e-book, 21 October 2009.

Van Gerven, M., 'Bayesian Networks for Clinical Decision Support: A Rational Approach to Dynamic Decision-Making under Uncertainty', *Dissertation*, Radboud University Nijmegen, 2007.

Vanek, L., 'Introduction to Bayesian Networks', viewed on the 17 July 2011. <http://www.fit.vutbr.cz/study/courses/VPD/public/0809VPD-Vanek.pdf>

Viereck, G. S., 'What Life Means to Einstein', Saturday Evening Post, 26 October 1929, retrieved 28 September 2011, <http://www.saturdayeveningpost.com/wp-content/uploads/satevepost/what_life_means_to_einstein.pdf>

Voß, S., Gutenschwager, K., *Informationsmanagement,* Berlin: Springer, 2001

Watthayu, W., Peng, Y. '*A Bayesian Network Based Framework for Multi-Criteria Decision Making',* University of Maryland, Department of Computer Science and Electrical Engineering, Baltimore, 2004.

West, G., 'The surprising math of cities and corporations', in *TED* 2011, retrieved 3 August 2011, <http://www.ted.com/talks/geoffrey_west_the_surprising_math_of_cites_and_corporations.html>

Zadeh, L. A., 'Fuzzy Sets', in *Information and Control*, 1965, vol. 8, p. 338 – 353.

Zadeh, L. A., 'What is Soft Computing', in *Soft Computing*, Springer Verlag Berlin, New York, 1997.

Zins, C., 'Conceptual Approaches for Defining Data, Information, and Knowledge', in *Journal of The American Society for Information Science and Technology*, vol. 58, 2007, p. 479 – 493.

APPENDIX

AIPA Overview Methods of Artificial Intelligence

"*Data mining:* the process of discovering new patterns from large datasets involving methods from statistics and artificial intelligence but also database management.

Ruled-based case reasoning: A particular type of reasoning, which uses "if-then-else" rule statements. Rules are behavior patterns and an inference engine searches for patterns in the rules that match patterns in the data. The "if" means "when the condition is true," the "then" means "take action A" and the "else" means "when the condition is not true take action B."

Expert Systems: are software solutions that use a knowledge base of human expertise for problem solving, or to clarify uncertainties where normally one or more human experts would need to be consulted.

Artificial Neural Networks: is a mathematical model or computational model that is inspired by the structure and/or functional aspects of biological neural networks. A neural network consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation.

Genetic Algorithms: is a search heuristic that mimics the process of natural evolution. This heuristic is routinely used to generate useful solutions to optimization and search problems. Genetic algorithms belong to the larger class of evolutionary algorithms (EA), which generate solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover.

Fuzzy Logic: is multi-valued logic; it deals with reasoning that is approximate rather than fixed and exact. In contrast with traditional logic theory, where binary sets have two-valued logic: true or false, fuzzy logic variables may have a truth-value that ranges in degree between 0 and 1.

Machine learning: is a scientific discipline concerned with the design and development of algorithms that allow computers to evolve behaviors based on empirical data, such as from sensor data or databases.

Intelligent Agents: an autonomous entity, which observes and acts upon an environment (i.e. it is an agent) and directs its activity towards achieving goals (i.e. it is rational). Intelligent agents may also learn or use knowledge to achieve their goals.

Automatic process control: Engineering based discipline (architecture, mechanisms, algorithms) for maintaining the output of a specific process within a desired range, by moving field actuators following predetermined error correction algorithm. The objectives are to proactively keep a process in statistical control, maintain certain operating point, keep process safety, or optimize asset performance. The control signal may be computed from field measurements and optimum expected performance target which are derived using physics-based or data driven analytical methods or artificial intelligence techniques such as neural networks, fuzzy logic and others.

Proxy Models: the proxy models are a simplified representation of the response surface of the numerical models, used commonly to make an approximate simulation model of a physical process (reservoir models, well models, surface models, advance process control) in a specific boundary of time and restrictions. Surrogate Reservoir Models, are proxy models that are developed using Machine Learning technology. They are also classified as "Surrogate Models". AI-Based reservoir models.

Workflow automation: is a set of methodologies and technologies which aims to integrate data and applications into automated workflows, which reflects the business processes developed in a company over a well structured information management platform. Workflow automation has been one of the main areas of interest of the Oil & Gas Industry in the last 5 years, since is one of the key elements of the Digital Oil Field and Integrated Production Operations trends.

Virtual environments: is the combination of simulation, computation and visualization technologies in order to reach partial or total immersive environments for the analysis of production and reservoir data."[152]

[152] quoted from AIPA (2011)

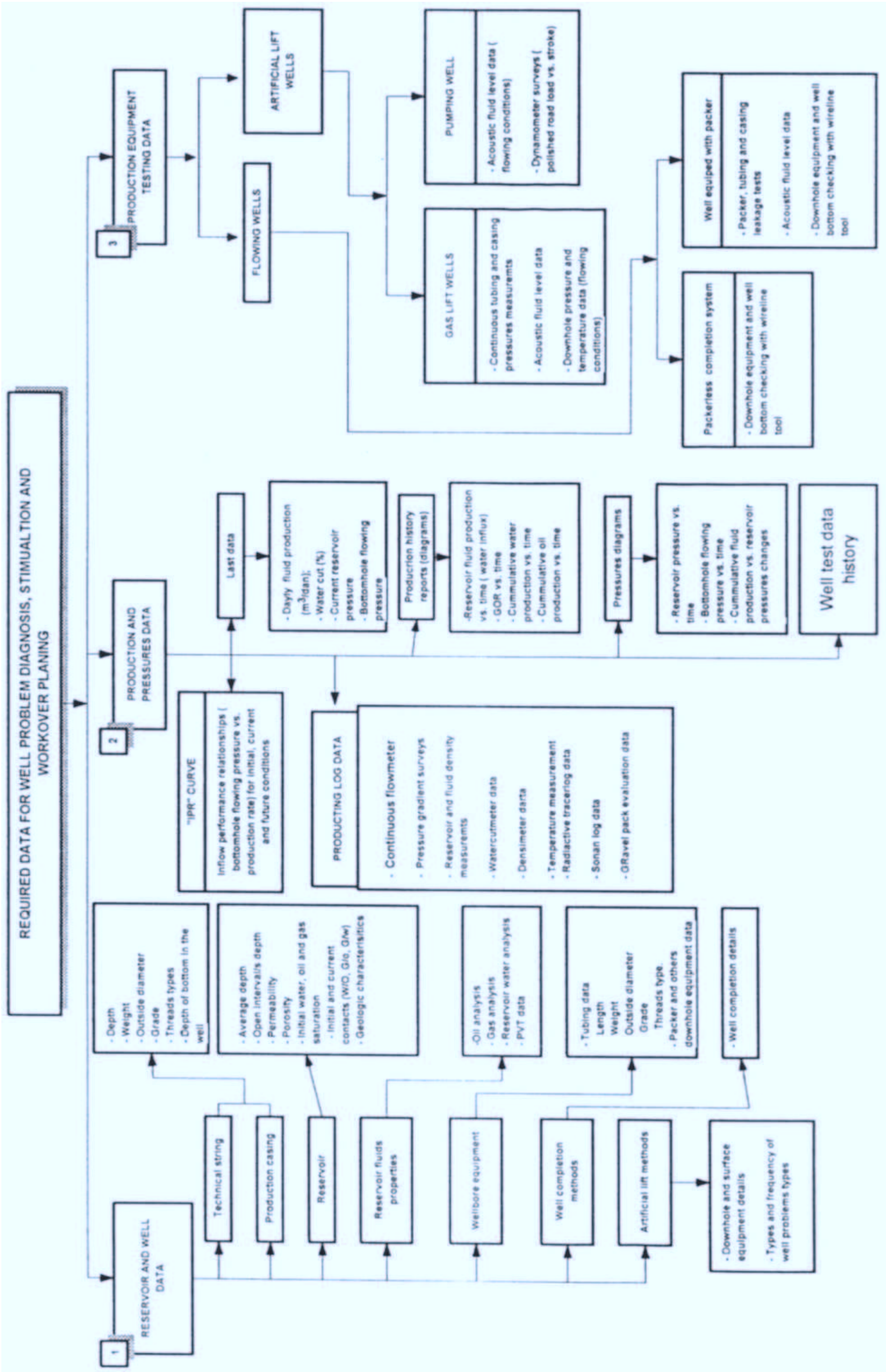


Figure 40: Data required for well problem diagnosis, stimulation and workover planning[153]

[153] Solesa (2004)

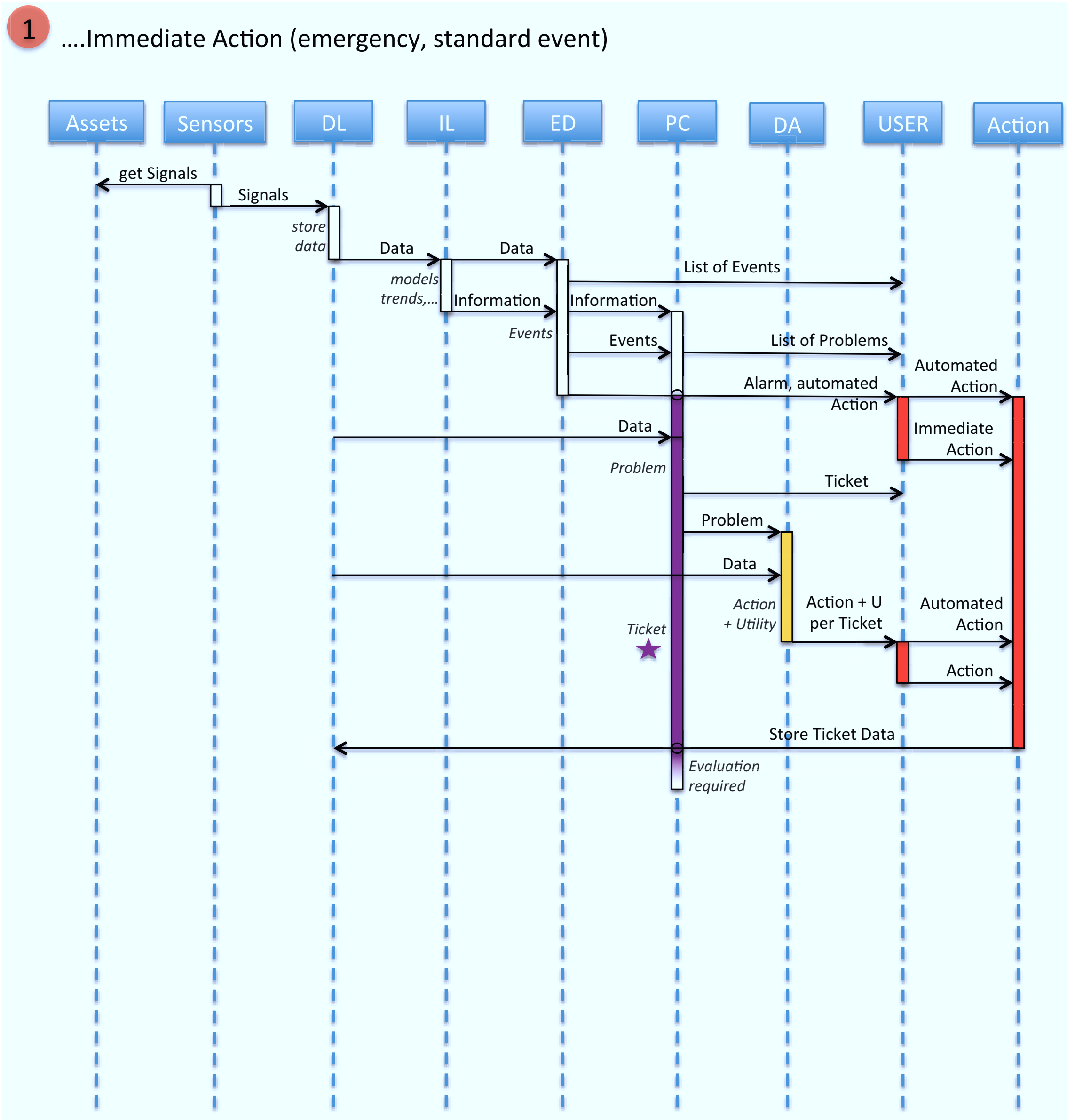


Figure 41: Sequence diagram for an emergency or standard action

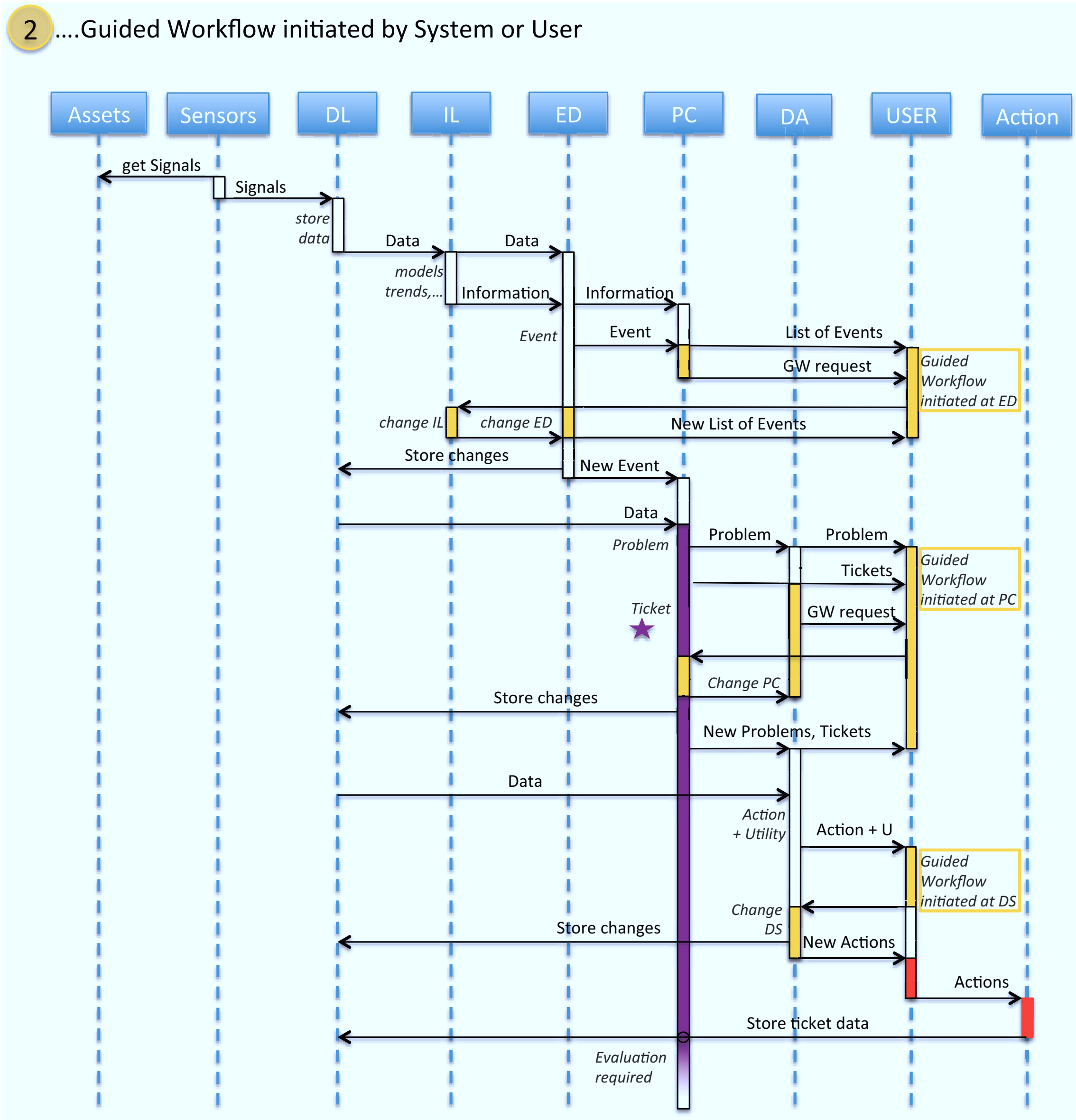


Figure 42: Sequence diagram for guided workflow

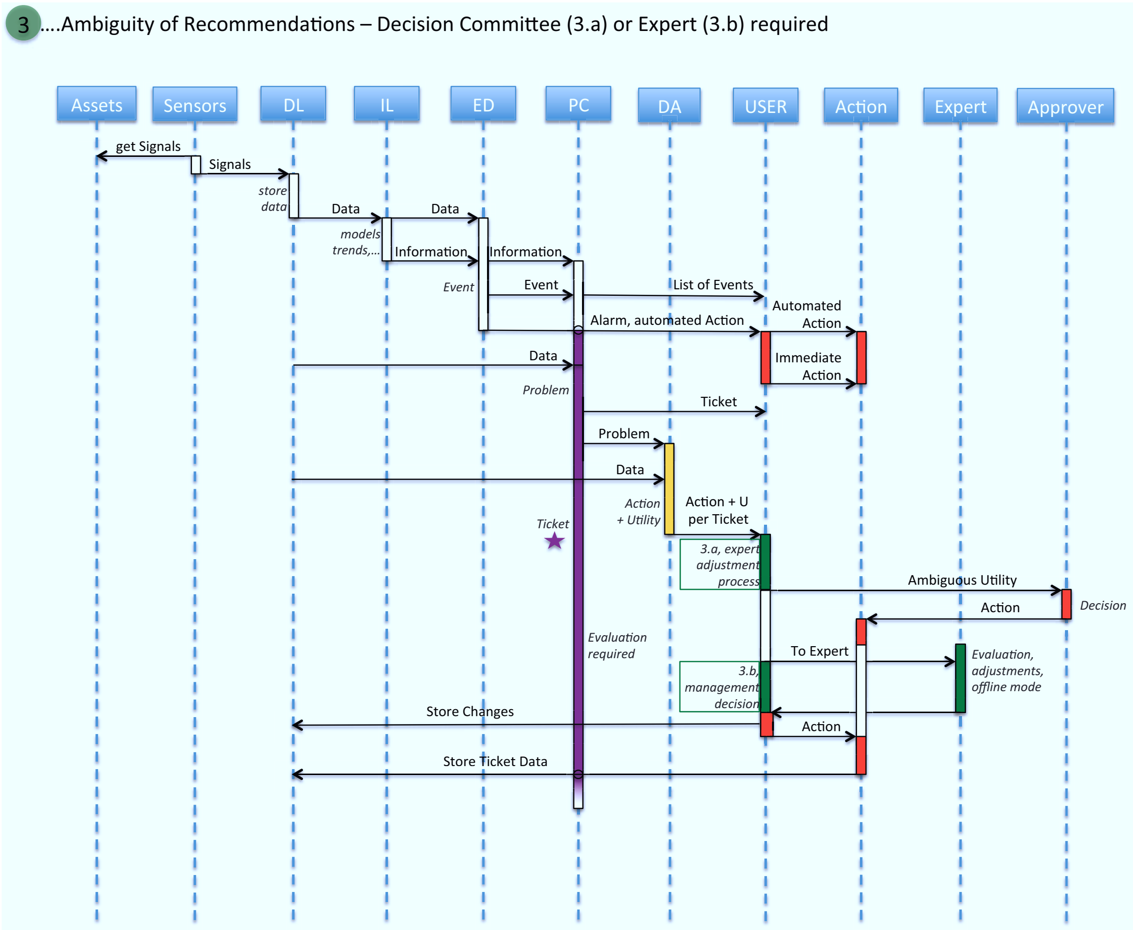


Figure 43: Sequence diagram for ambiguous recommendations