ABB Review

The corporate technical journal
of the ABB Group

www.abb.com/abbreview

4 / 2006

Innovation highlights

Ideas are transformed into innovations through a research and development process. In this issue of ABB Review, we highlight products and solutions that have progressed through the ABB development process to the stage of proven prototypes or even launched products.

Innovation highlights

At ABB, 70 percent of our sales come from products that are no more than five years old. And yet, many of these must remain in operation for 25 years at least! Compare this to companies in the telecom industry, for example, where the portfolio renewal is no doubt faster, but it is coupled to shorter product life and product support expectations. However, no matter the industry, certain factors are the same if a company and its product portfolio are to remain competitive. This includes the right mix of fresh ideas with proven concepts, increased functionality and attractive pricing. And innovation is what this issue of ABB Review is highlighting. From a long and impressive list, our technology management team has selected what they consider the "Top Innovations of 2006". The selection criteria were based on technology and mid-term business impact, as well as merits of creativity and originality.

An innovative manufacturing concept for lean and flexible production is presented in the lead story. Known as FlexLean, it allows multiple products to be produced on the same line using a standardized but flexible cell concept. Although designed for the assembly of car bodies, we believe it will also significantly benefit other applications requiring single-line, multiple-product manufacturing.

Collaboration between humans and robots has been a popular theme for countless science fiction writers. Among the most famous robots in print – as well as on the big screen – are R2D2 of "Star Wars", and Asimov's "I, Robot". While some robots try to emulate the R2D2 type, modern day industrial robotics has remained far from these visions, mainly due to human safety reasons, ie, the risk potential of a sudden movement by an active robot is too high. An innovation by ABB called SafeMove will eliminate this danger and allow humans to work side by side with robots.

Within the power industry, the ability to design self-healing electrical networks is a desirable objective. ABB research has now brought this objective a step closer to reality by developing the functional and architectural specifications of the IT infrastructure necessary for supporting such a grid.

The implementation of IEC 61850 as a single and global standard for substation communication is now ubiquitous. It has led to a number of ABB innovations, which are discussed in two articles. The first deals with hollow waveguides for medium-voltage switchgear, which are used to distribute the vast data-flows needed in such equipment. While flexibility is increased, this technology significantly reduces the amount of switchgear wiring needed, making the overall set-up more immune to wiring errors and electromagnetic compatibility issues. Other innovations stemming from the standard (and discussed in the second article) include several plug and play tools for substation automation applications that help alleviate costly project engineering.

The first of a two-part feature on power semiconductor technology concludes the power segment of this issue of ABB Review. As an essential switching technology for power applications, its current status is discussed in part one. Part two, looking into housing design and the future, will be published in our first issue of 2007.

For many of us, classifying colors is not an easy task. However, the measurement task is even more difficult especially when it has to be accomplished on-line in a harsh environment – such as on a paper machine – and at speeds in excess of 100 km/h! Based on standard color definitions we show in a sequence of articles how this complex issue can be dealt with and how advanced control can reduce the use of chemical dyes to the benefit of the customer and the environment.

Cutting giant reels of paper into smaller rolls according to customer specification is not as straightforward as it might seem. To optimize the individual rolls, the cutting algorithm must take into account quality data from on-line sensors resulting in massive data manipulation tasks.

The next two articles deal with economic optimization and asset management in customer implementations.

The industrial segment concludes with the description of a new principle for the on-line measurement of oxygen, an application used in many industrial processes and utility plants.

The right mix of insight and ingenuity counts at the interface between market pull and technology push. For innovation to succeed, a clear understanding of present and future needs, and opportunities must be sought. Highlighting innovations help communicate our ideas to our customers and suppliers. This continuous dialog is vital to the success of all parties.

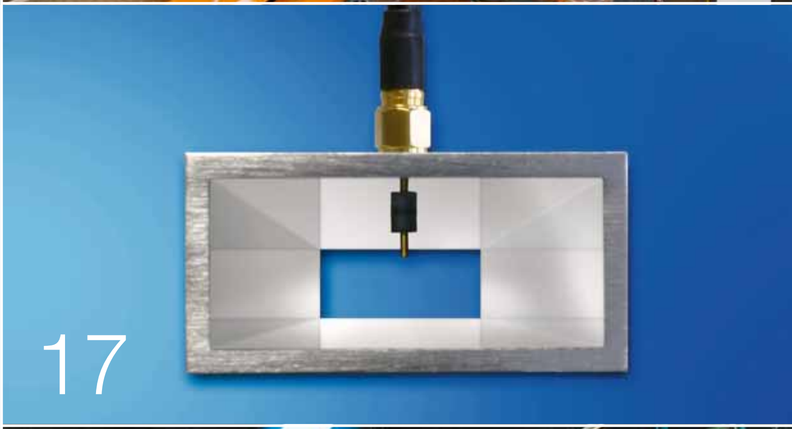
Enjoy your reading

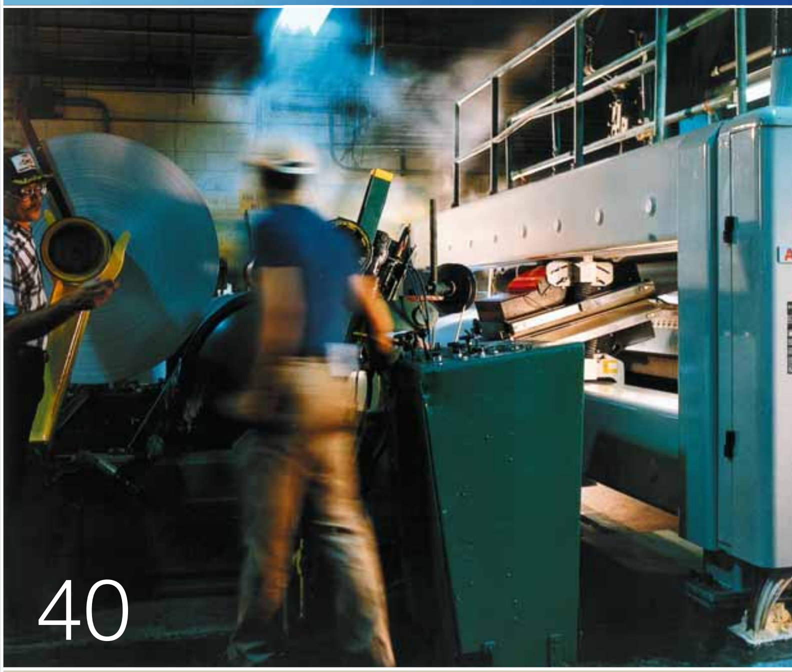
Peter Terwiesch
Chief Technology Officer
ABB Ltd.

Contents

ABB Review 4/2006
Innovations highlights

www.abb.com/abbreview





FlexLean

Robots challenge low cost labor

Bernard Negre, Fabrice Legeleux

The automotive industry is generally credited with having pioneered the large scale use of robots. Long production runs of identical cars were the ideal field of application for these untiring and reliable workers. Today's market is increasingly putting automobile manufacturers under pressure to offer customers more choice, while at the same time demanding lower production costs. To fulfill these apparently contradictory requests, a single line must be able to produce a mix of different models, and must "learn" to make new models without calling for a total re-design of its equipment – and preferably without even stopping production.

ABB has responded to these demands by making robot cells more adaptable, easier to install and more economic on space. The new generation of cells can produce so cheaply that they can compete with manual labor in low-cost countries – providing manufacturers with an alternative to outsourcing while at the same time raising the quality supplied to local markets.

A contemporary production line for small cars completes a vehicle every 45 seconds all day and every day. The naked body shell of a typical car – the so-called "body in white" (BIW) – is typically assembled from 200 to 400 parts (including the parts used to make sub-assemblies). This means the BIW line (including its sub-assembly lines) consumes this number of parts every 45 seconds. Not only must the robots themselves work like clockwork, but advanced logistics are required to keep the line running.

Another challenge facing such production lines is one of equipment reusability. In the past, a production line was specifically designed for a single car model. When the time came to introduce a new model, a new line had to be designed and built. Such a solution required considerable investment and lead time.

ABB, as a robots supplier and system integrator, has always recognized the value of making such "carry-over" operations as smooth and simple as possible. The first element of this lies in enabling cost savings by allowing as much equipment as possible to be reused. However, although the "carry-over" of individual robots and other components may be relatively simple, they must still undergo the same configuration and test cycles as new equipment. To provide greater reusability while simplifying the configuration process, ABB introduced FlexiBase. A FlexiBase is a modular robot cell in which the robots, controllers and cabling are pre-mounted on a platform. This is set up and tested in the ABB factory and delivered to the customer as a working module. It requires only minimal configuration before it can start production. This saving is repeated when the cell is reused.

ABB is now proposing a standard cost-competitive solution to compete with manually operated bodyshops: FlexLean.

In this way, the reusability of installations has changed from being a wish to becoming a "leitmotiv". Several recently released car models are produced on older lines. The "carry-over" of such installations is a real challenge for integrators, especially when the new car model is to be assembled on a line that is already running. In such a situation, the production ramp-up of the new model may not stop or slow down production of the other models. Thanks to its experience as a flexible-line supplier, ABB is considered as a reference for "carry-over" operations by different car makers.

So, in addition to a line being carried over from an old to a new model, such flexibility allows a mix of different models to be produced on the same line. The costs of preparing the line for a new model are decreased dramatically as is, by consequence, the risk to the manufacturer if a new model fails to sell as well as forecast.

The "Far East" challenge…

In the growing Asian market, manually operated installations are still the norm, and robots are only used if heavy duty and accurate operation is required. An example of this is ABB's FlexFramer, a car framing station in which two robots manipulate the 500 kg tooling used to enclose the car body to ensure its geometry. Considering the complete line, however, this arrangement does not require many robots. ABB is now proposing a standard cost-competitive solution to compete with the running costs of a manually operated bodyshop: FlexLean.

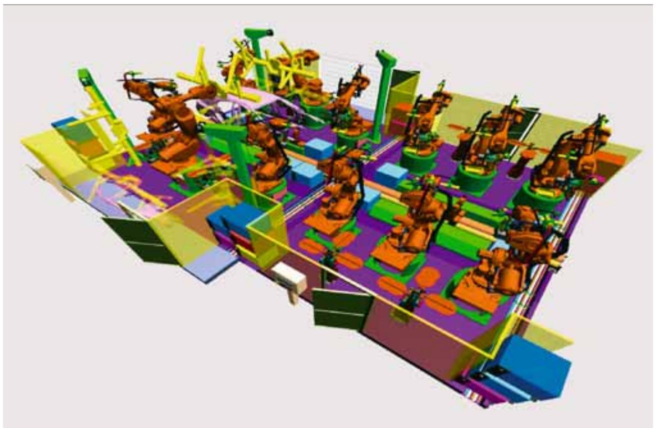
Thanks to FlexLean, it is now possible for the customer to get the flexibility and reliability of a robotized installation at a competitive price compared to the costs of a manually operated one. Moreover, this solution requires up to 40 percent less floor space. The compactness in the modular solution 1 is an advantage from the point of view of quality and logistics (supply of parts). Its scalability is also of advantage when future car models are introduced.

FlexLean

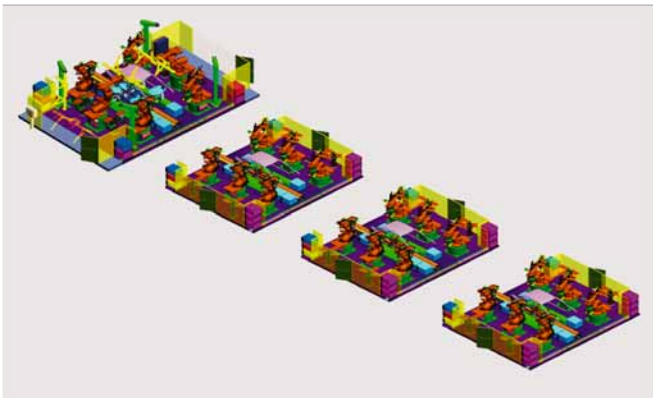
The basic concept of FlexLean is very simple. It builds on the FlexiBase principle, while further raising its adaptability. The "lean" in its name reflects the simplicity, the standardization level, and as a consequence the cost reductions that it enables – making it able to compete with manual labor in low-cost countries, while delivering savings over other robot solutions in higher-cost countries.
At the heart of the concept lies the recognition that customized solutions, multiple technical specifications and dedicated software are a major cause of costs and engineering uncertainty. The answer lies in robot technology and standardization: Two types of such cells are offered 2; one for geo-

3 A robot mounted on a hollow riser. The drive module is contained in the riser, making optimal use of space. Note the absence of cables and pipes from the floor space

1 A high density of robots: Besides saving on floor space, such an arrangement simplifies logistics (supply of parts)



2 An assembly cell and three respot cells. These cells are considered standard products and can be equipped with different process packages



Robotic highlights

Factbox The components of FlexLean – FlexLean incorporates a broad range of new robotic products

IRC5 drive unit

With this robot-control system (which integrates the multimove feature), it is possible to control and coordinate several manipulators, as well as external axes. This is the ideal platform to control the new robotic products such as FlexPLP and FlexGrip.

IRB 6620 robot

A study of the latest OEM projects using ABB robots has shown that value efficiency can be improved by 20 percent. The IRB 6600 is a very good "generic" robot, but offers more capacity than required when it comes to spotwelding applications. The decision was taken to specify, in close collaboration with ABB's robotics division in Sweden, a dedicated spotwelding robot. This new robot would be optimized for its task in terms of cost and performance – the IRB 6620 was born. It has a 2.2 m working range with a 150 kg load capacity – enough to carry a spotwelding gun. The optimization of the arm design is so drastic that the total weight of the robot could be almost cut by half as compared to the IRB 6600! The compactness of this new design makes it possible to increase the density of robots in a cell. Notably, despite the simplified nature of the design (for example, the weight balancing mechanism was omitted) optimization had no impact on motion performance: The new IRB 6620 is as fast and as repeatable as robots of the IRB 6600 family.

FlexPLP: polar

(Flexible Programmable Lean Positioner) This small robot with three positioning axes is used as a positioner to support the car body or workpieces. Previously, jigs would have been used for this. But with different car models being produced on the same line, and each model requiring a different jig, the case for an adjustable solution becomes clear. Typically during work, the car body or workpiece can be supported on four or more FlexPLPs. FlexPLP is available in two versions.

The polar version requires three motors, each of which powers two actuators. The target position is best described in polar coordinates. In the simpler linear version, the three axes are set to the x, y and z coordinates of the target position.

The linear version of such a positioning stage is not new – several models are available on the market – but this one has very specific features. The unit is modular, allowing each axis to be used as a standalone product if required. The design is also highly compact and its motor is enclosed for better protection.

FlexPLP: linear

The process cables run inside the three axes of the unit to its extremity, where they power the actuator and collect information from sensors.

The design presented the challenge of protecting the inside of the slide from the harsh environment caused by dust, fumes, welding splatters etc. that characterize the working conditions of a BIW line. The conventional solutions used for protecting such a linear slide influence the stroke, making the unit more cumbersome. Instead, protection had to be achieved by design. The shape of the sliding section allows a cable cover to run through it. This ensures good protection without calling for more conventional bellows, which are both fragile and require more space.

The customer can reap considerable benefits with this technology: Introducing a new car model in the line is just a matter of programming the new positions for locators. This operation can be achieved off-line, using simulation software (such as ABB RobotStudio), with only a short production break for final tuning.

FlexGrip

FlexPLP is a good solution to the challenge of tooling flexibility. Nevertheless, there still remains the problem of using robots to bring various part types into and out of the production line. Usually, such operations call for several end-effectors with tool-changers. In such a situation, the robot leaves its gripping tool in a docking stand and picks another. The drawback of this approach is the floor space required to store grippers, and the resulting effect on cycle times. The FlexGrip module offers a solution that dispenses with the tool changer. Instead, FlexGrip consists of adjustable grippers similar in principle to the linear FlexPLP. The resulting module is specifically designed to be carried by a robot. It is lightweight and the critical masses (motors) are located around the interface with the carrying robot. Several such units driven by the IRC5 controller can be used on the robot end-effector to build a programmable "gripper".

metrical assembly and one for respot[1]. These come with a choice of pre-defined configurations: several process packages (such as spotwelding, gluing, roller hemming) are available for the assembly cell. The number of robots in the respot cell is scalable from two to six.

All equipment belonging to a FlexLean cell, including the controllers, is located on a FlexiBase (steel frame base). This base contains all piping and cabling, resulting in a clean surface that can be walked on without danger of tripping. The compactness is further optimized through newly designed hollow robot risers 3 permitting the IRC5 drivers to be located just under the robot. After startup and commissioning at ABB's facility, each cell is disconnected from power and adjacent cells and delivered by truck to the customer plant, where re-building of the line is achieved in just a matter of hours.

The components of FlexLean are discussed in the Factbox.

Flexibility in set-up

In a production line using FlexLean, each cell is a standalone system. During startup and commissioning, staff can work on isolated cells, as each such cell is controlled by a complete set of automation modules, including a PLC and a man-machine-interface 4. These modules communicate through a three-level bus system:
FieldBus for real-time process,
SafetyBus for critical I/Os and
Ethernet for generic information

The standardization of products used in the cell is not only advantageous for cost reduction, but is also a big advantage for the control system. Complex PLC programming operations are no longer required: every component of the cell is known and the required PLC code is written once for all.

However, one part of the automation still remains variable – the part that is related to the manufacturing process itself. Here lies what is probably the most innovative aspect of this solution: instead of hard coding the spe-

FlexTrack

As customer requirements for bodyshop flexibility increase, another problem arises: how can sub-assemblies, or even a complete car body, be moved across the production line without restricting flexibility? At present, conveyors use model-specific pallets or tooling to handle the parts during motion. These pallets have to be returned to the beginning of the conveyor system when empty. This return circuit is cumbersome, often requiring an aerial system located on a mezzanine. This makes it expensive. Moreover, the co-existence of several car models in production means several pallet types exist, each requiring different treatment. Under such conditions, dealing with more than two car models in the same line becomes a logistical nightmare. Once again, the solution lies in flexible tooling using three-axis robots to position the part locators. This programmable jig is carried by a linear track motion, which like the robots, is driven by the IRC5 controller. FlexTrack was developed for applications where compactness, protection and cost efficiency are required. The compact width allows optimized layouts where the stationary FlexPLPs are close to the linear track. All internal guidance and transmission components are protected against pollution from welding. This linear track motion can also be used to carry robots for various processes such as gluing and palletizing.

An additional strength of such a linear track motion system as compared to a conventional conveyor system is that each car body is moved individually with a very high accuracy and repeatability. This can greatly simplify certain steps. The transfer time is drastically reduced to less than 5 sec for 6 meters.

FlexGrip in action. This variable gripper has four linear positioners which are coordinated to act as a compact robotic "gripper"

A car body supported on FlexPLPs – those on the left are the polar type and those on the right linear

Footnotes

[1] respot is the process providing the final weld after initial welds are used to hold the parts in position.

cific process, FlexControl integrates a configurable sequencer 5a.

> Complex PLC programming operations are no longer required: every component of the cell is known and the required PLC code is written once for all.

The traditional time-consuming process of programming, compiling, transferring to the PLC and finally debugging the generated code is no longer called for. Now it is just a question of configuring the sequence of operations for each actuator (robots, FlexPLP, clamps 5b etc) by selecting one of the possible operations from a list 5c. The mode switch is then flicked to "auto" and production begins. It is so simple that a PLC programmer is no longer needed to modify the process steps; the person setting up the program merely has to know what he wants FlexLean to do. In production mode, the sequencer is used to display the process status with color codes 5d.

Robots redefine manufacturing

Today, production lines are in existence that can handle a mix of four different models. In future lines, this number may be doubled. There is, however, a limit to the number of models a single line can handle. This comes from the logistics of supplying the parts to the cells: Besides the greater organizational complexity required to handle the increased parts inventory, one fundamental problem is that bringing more part types into a cell demands more space around the cell, and this is already scarce.

> Today, production lines are in existence that can handle a mix of four different models. In future lines, this number may be doubled.

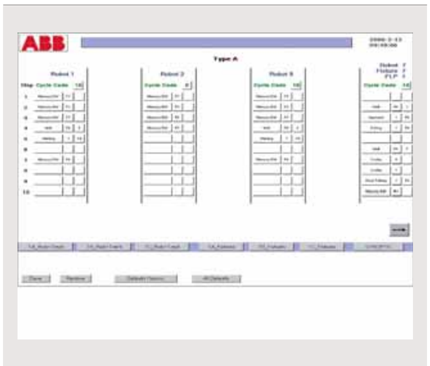
In addition to the different models that are handled, further variants are possible through customization. Each car is defined before production starts, and this information fed to the cell controllers at the appropriate time. This permits, for example, such customization as additional holes or fittings for accessories and avoids expensive subsequent modifications. Manufacturing has come one step closer to the dream of combining mass production with made-to-order individualization.

Innovations in robotics serve to make robots easier to setup, use, and reuse, while cutting back constraints such as space requirements and time lost during operations. These advantages are helping position robots in increasingly challenging applications.

4 A robot cell has a multitude of robot controllers. All are centrally configured and coordinated with FlexControl

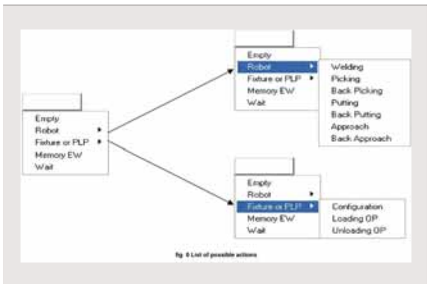
5 The configurable sequencer greatly simplifies the programming of robots

a The sequencer



b Tooling parameters on the sequencer

c A programming example – just a question of selecting the desired action



d In production mode, the sequencer displays the process status with color codes

Bernard Negre
ABB Manufacturing Automation
Auburn Hills, USA
bernard.negre@us.abb.com

Fabrice Legeleux
ABB Manufacturing Automation
Beauchamp, France
fabrice.legeleux@fr.abb.com

Taming the robot

Better safety without higher fences

Soenke Kock, Jan Bredahl, Peter J. Eriksson, Mats Myhr, Kevin Behnisch

Today's industrial robots are fairly well behaved creatures – doing only what they are programmed to do – as long as nobody gets in their way. Since robots move at astonishingly high speeds and carry loads up to 600 kg, humans must keep their distance. The two are usually separated by fences, like visitors and tigers in the zoo. However, the cost of such traditional safety equipment is high and it is slowing the advancement of robot-based automation in highly industrialized countries. ABB now offers a solution that cuts the cost of robot installations by replacing expensive mechanical safety equipment with dedicated electronics and software.

Robotic highlights

According to European and North American regulations for occupational health and safety, even a potential malfunction of a robot's controller hardware or software – however unlikely – is considered a risk and must be anticipated. This means that if the door to a robot cell is opened, for whatever reason, a contact must be tripped and the machine shut down immediately. To avoid even a theoretical failure of this safety device, dual channel switches and circuits are required, as are fitted in all safety circuits of ABB robot controllers. To avoid accidents caused by robot collisions, mechanical cams are used to activate position switches mounted on the robot axes, thereby limiting the robot's range of motion. These are expensive, hard automation methods that curb the efficiency of a machine that was originally intended to provide flexible and affordable autoation.

ABB's answer is not to compromise on the safety of its robotic products – but to provide customers with more cost-effective safety installations. SafeMove™ can even enhance the flexibility of ABB's robotic products.

Workers' safety – a competitive disadvantage?

The fact that accidents with robots are extremely rare suggests that adequate safety measures are already in place. Indeed, some argue that safety has been taken too far and that the strict regulations imposed on European and North American factories make them less competitive than their rivals working under less exacting safety standards. Others ask why robots should have higher safety requirements than overhead cranes. Such cranes carry substantially heavier loads than robots and their manual operation makes them subject to human error; robots perform repetitive, pre-programmed tasks and generally make no mistakes.

ABB's answer is not to compromise on the safety of its robotic products – but to provide customers with more cost-effective safety installations by making use of the latest technological advances. By replacing expensive mechanical safeguarding equipment with more efficient and re-configurable electronic motion safety for robots, the SafeMove™ concept can even enhance the flexibility of ABB's robotic products.

Next generation robot safety

SafeMove builds on the latest developments in redundant software, electronics-based safety technology and modernizations in safety regulations (ISO 10218). It allows the reliable, fault-tolerant monitoring of robot speed and position, and the detection of any unwanted or suspicious deviation from the norm. If a safety hazard is detected, SafeMove executes an emergency stop, halting the robot within fractions of a second.

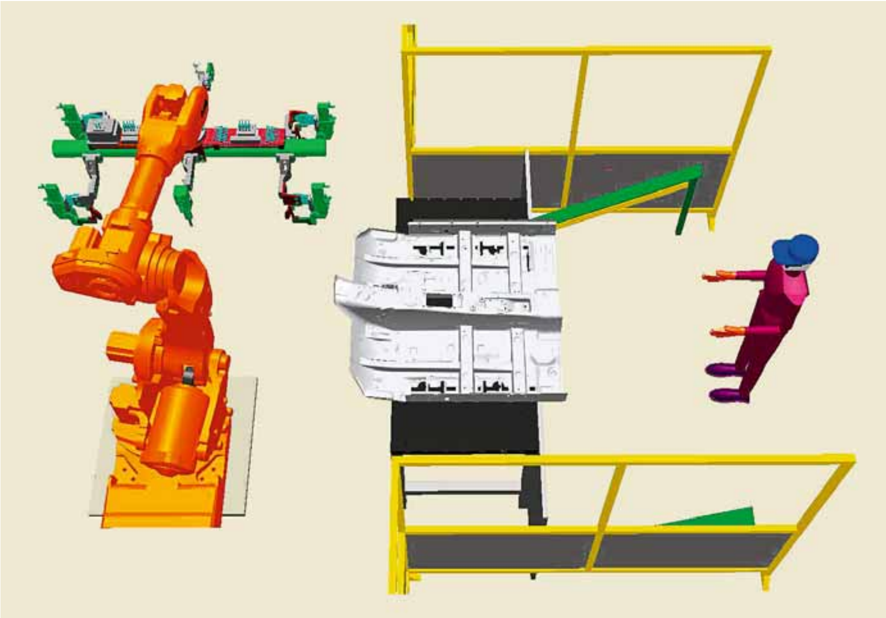
SafeMove also offers new functions such as electronic position switches, programmable safe zones, safe speed limits, safe standstill positions and an automatic brake test, which allows more flexible safety setups.

Programmable safe zones can be used to ensure that the robot stays out of protective, three-dimensional zones. These zones can have complex shapes, adapted to specific installations. Alternatively, the robot can be confined within three-dimensional geometric spaces, allowing significant reductions in the size of robot installations. The fences can now be moved much closer to the robot, saving valuable floor space.

Of course, it is also possible to limit axis ranges by mimicking conventional electromechanical position switches using software – but this is no longer restricted to the three principal axes of the robot. Instead, all 6 axes can be safely limited. Axis limits can be combined logically, and work-piece positioners, linear tracks and other external axes can be restricted without any additional effort.

In "safe standstill" mode, robot movement is inhibited completely, yet all drives are powered and the motors are actively controlled. The purpose of this operating mode is to allow the worker to approach the robot in safety, and even to load a work piece into the gripper or carry out maintenance on the tooling without the need to shut the robot down. This not only saves cycle time when operation is resumed, it also reduces wear on the brakes and contacts needed to achieve the shut down.

In "safe speed" mode, the robot may even be allowed to move – complete-



ly or partially – at a speed that is slow enough to pose no threat to the worker, eliminating the need for a separating fence altogether. In combination with other supervision – such as confined space – workers and robots can now perform manufacturing tasks together, something that has not previously been allowed.

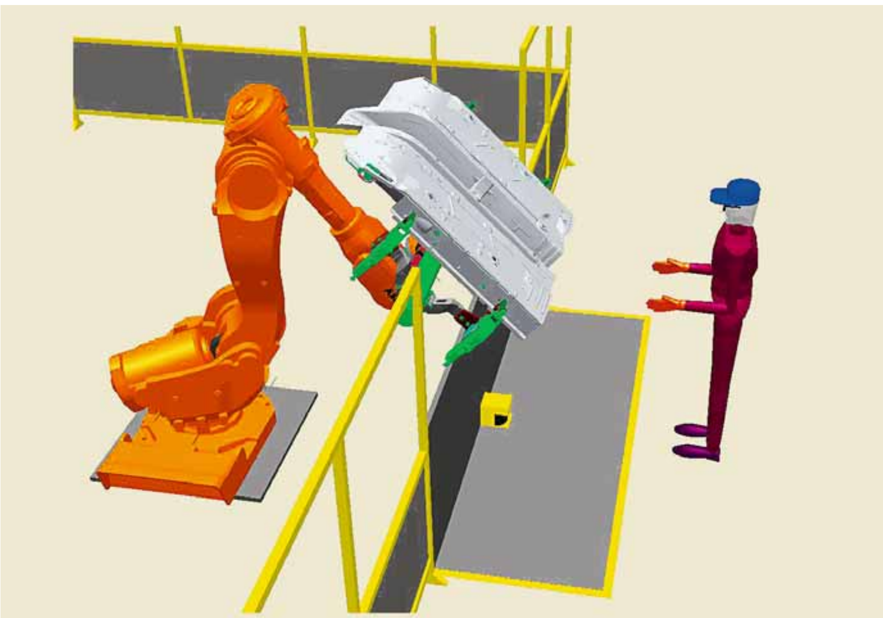
The safety of a robot ultimately relies on its ability to stop, or be stopped, when a hazardous situation arises. This stopping capability is determined by the function of mechanical brakes on the robot motors. SafeMove therefore contains an automatic brake test procedure, that periodically checks the mechanical brakes of the robot – something that would be very useful in a car!

Technical solution

SafeMove uses sensors that are already used for motion control to monitor the position of the motor. It then computes the robot position in a safety-rated computer that works independently from the robot controller.
In addition, the sensor signals are checked for sanity. A separate model of the robot mechanics and extra reasoning about the nominal behavior of the servo control loop further enhance the safety level (patent pending). Even though SafeMove is an independent computer that sits in the cabinet of ABB's fifth-generation industrial robot controller, the IRC5, from a user perspective, it is seamlessly integrated. Events, alarms and changes of state are logged on the robot controller's flash disk. The state of the safe inputs and outputs can be read just like normal robot I/Os and used in the robot program, even though there is no physical wiring between the I/O systems. Instead, SafeMove and IRC5 communicate over an internal network link.

SafeMove can also impose speed limits on a robot and ensure that it stays out of protective, three-dimensional zones.

Finally, synchronization between the safety computer and the robot controller must be checked after a power outage and at the beginning of each shift. This is achieved by a simple switch mounted in the cell, where it is easily accessible by the robot. The switch is visited and regularly activated by the robot – typically every 24 hours. Since this procedure can be easily combined with regular automatic tool service operations like cleaning, dressing or wire cutting, it does not normally add to the cycle time of the installation.



Process safety

Robots often handle dangerous process equipment – such as weld guns, laser heads, water jet guns, or even radioactive sources. Such equipment needs special attention in case a fault develops. It may be necessary to provide a protective enclosure around the complete robot cell that can withstand the process energy in case of a robot malfunction. Imagine, for example, the consequences if a robot were to point an ultra-high-pressure water jet horizontally rather than downwards, and the jet is accidentally turned on. This is the kind of scenario that must be considered when planning a water jet cutting cell. Use of the SafeMove function now allows safety checks to ensure that the orientation and position of the robot tool are within a defined tolerance before the tool can be activated. During operation, the robot is monitored continuously, while the tool orientation stays within the tolerance band. As soon as this tolerance is exceeded, a safe shutdown of both the robot and the process equipment is initiated. This can lead to drastic reductions in the cost of protective enclosures.

Workers and robots can now perform manufacturing tasks together, something that has not previously been allowed.

Access security

Most accidents with machine installations occur as a result of disabled safety equipment. Safety is often seen as an obstacle to productivity, and workers will sometimes take calculated risks if time can be saved. It is therefore in the best interest of both worker and employer, especially now that safety functions can be moved from hardware to software, to limit access to configuration data. This can be done by providing password-access to specially trained, authorized personnel only. Industrial practice shows, however, that it is difficult to keep passwords secret on the shop floor and this leaves the system open to abuse. ABB's scientists and engineers have therefore developed and patented a mechanism that protects the safety set up of SafeMove by a combination of an access-restricted configuration tool and a public activation code. The mechanism makes the safety configuration as secure as a bank account and very convenient to use.

A separate model of the robot mechanics and extra reasoning about the nominal behavior of the servo control loop further enhances the safety level.

Safe, compact, fast and flexible
By exploiting SafeMove's features, it is possible to reduce significantly the number of safety devices employed, including light curtains, safety relays, mechanical position switches, protective barriers etc. Replacing mechanical position switches for robots and additional axes, means that there is no longer any need to maintain these devices, which are often exposed to severe environmental conditions and therefore have a limited lifetime. This allows robot cells to become more compact. Flexibility is increased as safety configurations can be reset easily using software. Replacing broken-down robots equipped with dedicated cams and position switches used to be a lengthy procedure. Today, the time required for such repairs is significantly reduced, since the safety parameters are handled by the controller and limit switches no longer exist. This can even lead to more compact robots, since the cam rings of the past required a significant amount of space; robots without position switches provide reductions in cost.

Planning and engineering safety
ABB offers RobotStudio, an off-line programming tool that allows the visualization, programming and testing of a robot installation on an office PC, and the SafetyBuilder, a secure tool for setting the parameters of, and activating, the SafeMove controller. The combination of these powerful tools allows the engineer to design and test the safety zones in a virtual environment during the planning phase, and later use the data for engineering and commissioning. All these advantages can be fully exploited by implementing them into the initial cell concept. Of course it is also possible to retrofit IRC5 with SafeMove, so that new functions can be introduced into existing IRC5 installations.

Flexible manufacturing
In the future, SafeMove will enable completely new manufacturing concepts with ABB robots. Since humans and robots are now able to work closely together, they will team up to become real colleagues. The powerful robot can lift and present heavy work pieces to the worker, and the worker can perform tasks that are harder to automate. Or the worker can load small parts from a container box directly into the robot gripper, without the need for separating turntables, receiving fixtures or roll doors, and the robot can then do the work – perhaps even in cooperation with another robot or another worker.

ABB is working to convert the technological advantages into cost savings.

ABB is working with partners and end users on new flexible manufacturing concepts to convert the technological advantages provided by applications like SafeMove into cost savings for the customers' operations. While the market is embracing these new possibilities, ABB's researchers and developers are already thinking about what will come next.

Soenke Kock
ABB AB, Corporate Research
Västerås, Sweden
soenke.kock@se.abb.com

Jan Bredahl
Peter J. Eriksson
Mats Myhr
ABB Automation Technologies AB
Västerås, Sweden
jan.bredahl@se.abb.com
peter.j.eriksson@se.abb.com
mats.myhr@se.abb.com

Kevin Behnisch
ABB Automation GmbH
Friedberg, Germany
kevin.behnisch@de.abb.com

Best Innovations 2006

2006

ABB has been at the forefront of technical innovation for many years. The following set of short stories provides a taste of the company's achievements in 2006, which range from a novel sensor for measuring oxygen concentrations to a programmable light switch. ABB's interests in communication, automation and safety are illustrated in these stories, showing that ABB's reputation for excellence in innovation is well deserved.

Safe robots don't need higher fences

Today's high-speed industrial robots can lift and swing payloads of up to 600 kilograms, so their fragile human co-workers must keep out of the way.

This is usually done by fencing off robot work stations, but these traditional methods are expensive and not very flexible. Now there's a better way to keep the workplace safe. ABB's SafeMove is a reconfigurable software and hardware package that cuts costs and increases the flexibility of robot installations, without sacrificing safety.

SafeMove is based on the latest advances in redundant software, electronic safety technology and safety regulations. It reliably monitors robot speed and position, instantly detecting unwanted or suspicious deviations. When it detects a safety hazard in a robot, SafeMove executes an emergency stop, shutting the machine down within milliseconds. SafeMove also offers other new functions, including electronic position switches, programmable safe zones, safe speed limits, safe standstill and automatic brake tests that make setting up robot safety functions much easier.

By eliminating the need for traditional safety enclosures, SafeMove software encourages new workplace configurations that allow man and machine work closer together, safely.

For example, a robot might present a heavy object to a human worker, who could then perform tasks on it that are difficult to automate.

Conversely, a worker might load small parts directly into the robot gripper, eliminating the need for separating turntables, receiving fixtures or roll doors. The robot could then do the work – perhaps in tandem with another robot, or another human worker.

The configurations are endless, and ABB is working with partners and customers on flexible manufacturing concepts to turn SafeMove software into cost savings for customers. SafeMove will be launched by ABB Robotics in the first quarter of 2007.

Soenke Kock
soenke.kock@se.abb.com

For more information on SafeMove, see "Taming the robot" on page 11 of this issue.

ABB data-transmission system sets new records

Power network operators use many communication channels, including their own power lines, to ensure the safe and uninterrupted flow of power.

Power line carrier (PLC) transmissions are reliable and cost effective ways to send important data over long distances.

In a power network, vital signals are exchanged in real time, between many essential locations, to ensure optimum control and protection of the entire power system. The communications infrastructure supporting this continuous coordination must therefore be fast and reliable.

ABB's latest PLC system, the groundbreaking ETL600, set a new record by transmitting via a 380-kV high-voltage power line over a distance of more than 100 kilometers using 32 kHz bandwidth at a speed of 320 kbit/s, very close to the theoretical limit.

In addition to its speed, the ETL600 also offers flexibility.

Electrical noise in AC (alternating current) transmission lines increases during thunderstorms, rain or snowfall, and affects the quality of PLC links. Many PLC systems are therefore designed for worst-case weather conditions, transmitting at low speeds, even in good weather. This is a waste of scarce communication capacity.

ABB's ETL600 adjusts to weather conditions automatically, which means it will run at maximum speed whenever possible, slowing down only temporarily in response to bad weather.

ABB has more than 60 years' experience transmitting data over power lines. Its PLC systems are currently being used in a 1,000-kV AC power line in Russia and a 1,146-kilometer-long 500-kV DC line in South Africa – two more records!

Stefan Ramseier
stefan.ramseier@ch.abb.com

For more information on the ETL 600, see "Making power lines sing" on page 50 of ABB Review 2/2006.

ABB voltage indicator makes the workplace safer

VisiVolt is a voltage indicator designed to minimize the risk of electrocution for workers on indoor and outdoor medium-voltage systems.

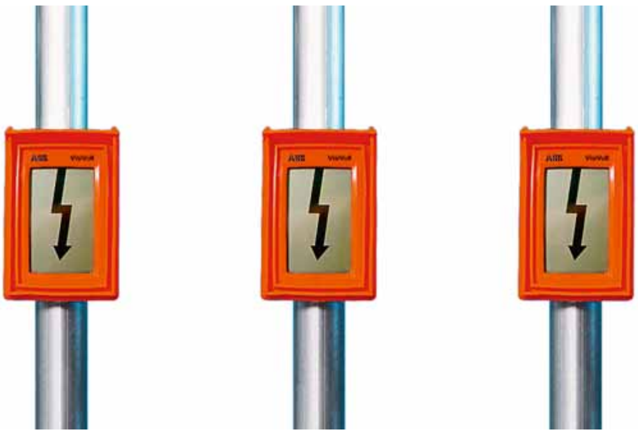
Electrical accidents have many causes; equipment malfunction combined with a momentary distraction can be fatal. Poorly trained workers may approach a live distribution panel instead of one that has been disconnected, with deadly consequences.

The VisiVolt indicator can be permanently installed, directly on current bars and conductors, using simple fittings. This compact device indicates the presence of a voltage by displaying a large and highly visible "lightning" arrow symbol on its liquid crystal display (LCD), providing an active reminder of the potential hazards of working in and around electrical systems.

Due to its unique structure – invented by ABB – the VisiVolt's LCD acts as both the display unit and the sensor element. The LCD detects the electric field around the conductor on which it is installed. The device requires no electronic circuitry, making it extremely robust and durable.

In the dangerous environment of medium-voltage systems, VisiVolt will warn workers that a particular part of the system is live – before it is too late. Since it can indicate the voltage status of every section of a distribution system, VisiVolt can also help to localize faults. It is particularly useful in systems where voltage indicators have rarely been used, such as open indoor switchgear panels and outdoor installations.

Jan Czyzewski
jan.czyzewski@pl.abb.com



Waveguide: simple, reliable, low-cost communication

ABB has developed a simple and reliable wireless method of transmitting data in switchgear installations that is cheaper and needs less maintenance than traditional cable-based switchgear communications.

Switchgear protects expensive electrical equipment by instantly turning off power during unexpected surges.

Reliable grids depend on high quality switchgear communication, and for this ABB Waveguide is an excellent choice.

ABB's Waveguide uses low-power electromagnetic waves inside a closed system to transmit data. This is done with a hollow conductor, an antenna, which receives and transmits modulated electromagnetic waves, and a coaxial connection to the protection and control unit.

Waveguide improves on traditional communication systems by avoiding the use of copper cables, which are sensitive to electromagnetic interference, and fiber optic cables, which have poor mechanical properties and are more complex to install.

Waveguide's rectangular conductor is made of aluminum and its dimensions are defined by the the electromagnetic waves being used. Trapping the signal inside an enclosure avoids radiation and external interference, and the field can be accessed easily by inserting antennas into the Waveguide.

The signal is transmitted with virtually no loss because it is reflected by two parallel surfaces.

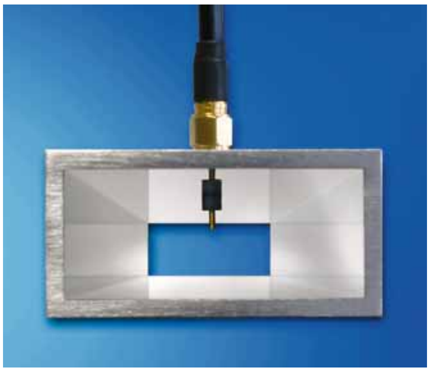
Research shows the Waveguide can transmit up to 22 times more information than cables, making it well suited to the new global standard in substation communication (IEC 61850).

Wireless signals in Waveguide are protected from external interference and the environment is protected from the radio signals. The system is easy to install, virtually maintenance free, and robust enough to withstand the harsh environment of a substation.

The concept was displayed with great success at the Hanover Fair 2006.

Bernhard Deck
bernhard.deck@ch.abb.com

For more information, see "Data pipeline" on page 26 of this issue.



Precision machines hold heavy loads

ABB's new FlexPLP machine can perform precision tasks and hold heavy objects within the tight, narrow confines of a factory production line.

The FlexPLP (Flexible Programmable Lean Positioner) is unlike regular industrial robots that are equipped with extendible arms and handle various materials at high speeds over long, measured distances. These robots have many uses, but they occupy appreciable space and for many production purposes, two arms are better than one.

That is why ABB developed the FlexPLP, a machine that provides precise manipulation and an ability to position heavy loads precisely in constrained spaces.

FlexPLP can support three times its own weight, yet is small enough to operate in an automotive production line. It could, for example, do the fine detail work of positioning locator pins and then carry a car underbody to the next point in the production line. FlexPLP is thus a highly flexible bodyshop all in itself.

ABB used the idea of parallel kinematic machines (PKM) as a basis for creating the FlexPLP, and worked for 18 months on the invention. The best known PKMs are ABB's three-armed Flexpicker and the six-legged motion simulator platforms known as Hexapods. FlexPLP was designed to address the PKMs' most serious disadvantage: the amount of space they need.

By clever arrangement of newly designed servo-cylinder pairs, ABB created a modular positioner for motion along three axes. Despite its unusual appearance, FlexPLP surpasses existing machine concepts of similar cost in terms of payload, repeatability and space efficiency.

Automotive factories are the first of many possible applications for these versatile machines, which have the potential to improve production in any industry.

Soenke Kock
soenke.kock@se.abb.com

Further information on FlexPLP can be found in the factbox on page 8 of this issue.

New switch cuts losses and increases power flow

Capacitor banks are commonly used to compensate for reactive power in medium-voltage distribution systems. They are, for the most part, fixed or breaker-switched, without being synchronized to the voltage and current waveform. They are adjusted only seasonally or, at best, weekly

By comparison, sub-transmission level capacitor banks are often equipped with synchronized switching to reduce over-voltages, and may be switched daily or even hourly.

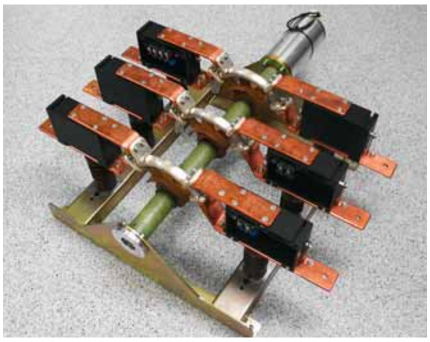
The ideal solution for capacitor banks at the distribution level is to be adjusted more frequently, like those at the sub-transmission level. An ability to follow hourly load variations would further minimize power losses and increase the maximum power flow in the distribution system. However, this is not current practice because circuit breakers are not designed for a high number of switching operations. Furthermore, the current solution can cause over-voltages and high in-rush currents in applications where a malfunction may have expensive consequences.

ABB has now developed and patented a novel switching mechanism designed for stepwise-controllable capacitor banks. It consists of diodes, rotating contacts, a motor drive and a control system. The switch is arc-free, and features synchronized make and break. It maintains more than a million operations, allowing frequent switching operations with negligible switching over-voltages and in-rush currents.

Low in-rush current enables capacitor banks switched in parallel to operate without in-rush limiting reactors. It is now possible to compensate for reactive power with a number of smaller parallel capacitor banks operated stepwise, instead of switching one large bank.

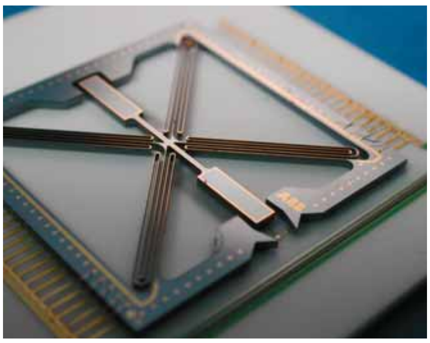
Finally, the switch avoids dangerous restrikes, making it ideally suited to frequent switching operations close to the load and thereby optimizing the system performance.

Mikael Dahlgren
mikael.dahlgren@se.abb.com



Revolutionary change in oxygen measurement

Hartmann & Braun, which became part of ABB in Germany in 1998, has been successfully developing paramagnetic oxygen sensors for more than 40 years. Now the company has returned to the drawing board to improve on classic oxygen sensors.



These manually assembled devices comprise an electromechanical module and an optical readout system. Their performance is good, but it could be better.

In response to this challenge, ABB has developed an innovative new sensor that will expand the application range of gas analyzers in the future.

The new sensor uses a cutting-edge silicon micro-electromechanical chip, developed at ABB's Corporate Research Center in Ladenburg, Germany. The chip is encased in a ceramic housing with integrated electro-optical components and optimized magnetic excitation.

The chip's main advantage over the classical paramagnetic solution is the drastically reduced volume of the sensor chamber. This cuts response time from three seconds to just one. In markets where speed is essential – such as monitoring combustion engines – this is a significant breakthrough.

The planar sensor chip is a key feature of the sensor's layered structure, which allows automated assembly and cost-efficient mass-production. The sensor is also highly resistant to corrosive gases and has negligible sensitivity to gases other than oxygen.

Thanks to low production costs and quicker response times, the new sensor is expected to change the way oxygen is measured– it competes in price and performance with classic, high-performance paramagnetic sensors, as well as the low-cost electrochemical devices.

Peter Krippner
peter.krippner@de.abb.com

For more information on this new device, see "Microsystems at work" on page 68 of this issue.

Variable light control with just one switch

The Busch Comfort Switch can be used to control lighting in a variety of modes from manual to fully automatic.

The switch is mounted flush with the wall and is equipped with an infrared movement detector that provides four modes of operation, from basic manual mode to fully automatic.

Other settings can be programmed into the switch to turn lights on automatically in response to the movement detector and to stay on until switched off manually. Alternatively, the light can be switched on manually and turned off automatically after a pre-determined time, if no movement is detected. In the "maximum comfort" mode, the light comes on if the infrared sensor detects a movement and switches itself off after a pre-set time.

Each mode and pre-set time can be adjusted with the help of potentiometers on the back of the switch's sensor. Manual mode is indicated by the activation of a light-emitting diode (LED) incorporated into the switch. This LED also helps users to find the switch in the dark.

The Busch Comfort Switch has two-wire technology combined with an external input, which means that it can be integrated into any two-way or cross connection, making it particularly useful in refurbishments. It can be combined with incandescent lamps as well as high and low-voltage halogen lamps.

The movement sensor has a detection range of more than 170°, ensuring immediate detection of anyone entering a room. The switch is suitable for use in both private and public areas, such as bathrooms, cellars, hallways and stairways.

Christian Heite
christian.heite@de.abb.com

Compatibility tool boosts grid reliability

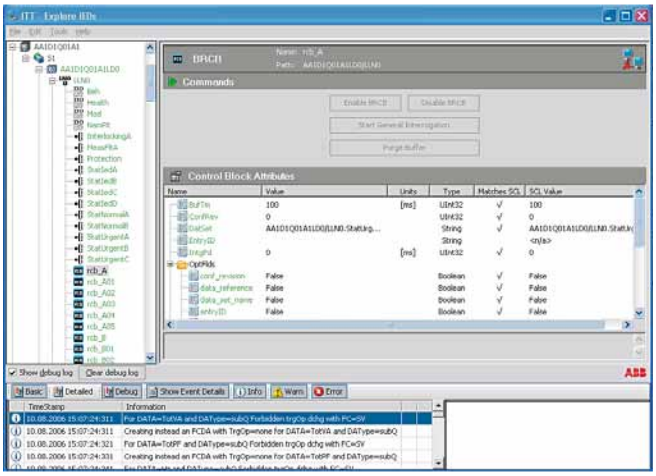
ABB has developed a simple software tool to analyze compliance of substation automation devices and systems with a new global substation standard.

Substation automation devices are like fuses that protect the electricity network and guarantee a reliable power supply. They communicate using a protocol, or language, that is chosen by the manufacturer. Until recently, many different languages were available, but this has now changed with the introduction a new global language – the IEC 61850 standard.

As a system integrator, ABB must ensure that all devices are working harmoniously by measuring key quality figures and demonstrating that the systems meet the required specifications. Building on the experience gained during the implementation of IEC 61850, ABB has developed a set of tools to support the process of testing and commissioning compliant systems.

The software enables substation automation engineers to analyze 61850-based systems quickly and reliably by just plugging a laptop into a substation communication network.

The software searches for automation devices in a network, inspects their configuration and examines network traffic. It helps to detect network problems and inspects the implementation of the 61850 protocol. It always verifies engineering data against actual data loaded on physical devices, checking for inconsistencies, which may manifest during or after the activation of a substation automation system.



Claus Vetter
claus.vetter@ch.abb.com

ABB drive features built-in PROFINET IO industrial Ethernet

In the world of control automation, low-cost Ethernet industrial protocols, capable of handling large amounts of data at very high speeds, are becoming essential

PROFINET IO is an important and open Ethernet communications standard. It focuses on programmable controller data exchange, and connects to higher-level control systems.

ABB now demonstrates the ACS 350, the first compact general-machinery drive with PROFINET IO support. Running on top of the drive's PROFINET IO protocol is a PROFIdrive profile, both of which are standardized by Profibus International. PROFIdrive is a common drive application interface that allows unified device access methods, independent of the physical drive employed.

ABB's Ethernet communication module makes it possible for a general machinery drive to take advantage of enhanced diagnostic features and central engineering schemes available in networked technology.

PROFINET IO is especially important in the European market and meets the requirements of many application areas. It is designed for real-time applications and available controllers usually operate with cycle times of as little as one millisecond. Standard management and engineering tools can be used to configure and maintain any device from any vendor. This makes ABB's ACS 350 with a PROFIdrive profile all the more attractive to customers.

Trygve Harvei
trygve.harvei@no.abb.com

ABB has the right mix for cement makers

The most important aspect of modern cement plant quality control is the Raw Mix Proportioning (RMP) system. The RMP system defines the proportions in which the main raw minerals of the cement-making process (limestone, clay, sand and iron ore) are mixed. ABB's solution to RMP optimization is proving popular.

The goal is to ensure that the resulting mixture (raw meal) has the correct chemical composition, while maintaining the lowest possible material cost. This is extremely important, because variations in the chemical composition of raw meal can increase fuel consumption and decrease cement quality.

The task of RMP control is made particularly difficult by the highly variable chemical composition of the minerals that are dug out of the ground. To meet this challenge, ABB has developed an innovative, state-of-the-art solution for the RMP problem.

The solution is part of the Optimize[IT] Expert Optimizer suite. It executes online control of the cement kiln feeders, where the various raw minerals are stored. This guarantees an optimal trade-off that balances deviations in quality targets and material cost.

The system's control algorithm is based on the latest control technologies, such as Model Based Control, which allow the cement plant's dynamic behavior to be simulated using mathematical models of the feeders, conveyor belts, mills, silos, etc. This allows the effect of different control actions to be predicted and appropriate action to be taken. The crucial mixing operation becomes predictive, rather than reactive.

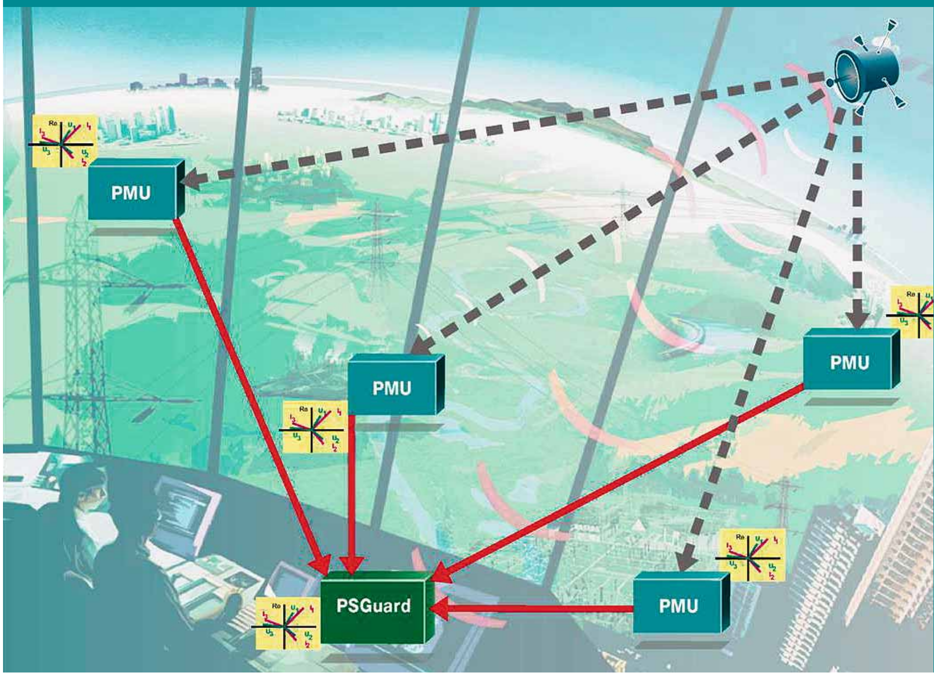
ABB was the first to come up with an advanced model-based application for the RMP process. A successful pilot installation has led to dozens of orders from cement manufacturers around the world.

Dario Castagnoli
dario.castagnoli@ch.abb.com

See also "Cementing profitability" on page 59 of this issue.

Vision for a self-healing power grid

Khosrow Moslehi, Ranjit Kumar



Power system reliability has been thrust into the limelight by recent blackouts around the world. The social and economic costs of these failures can add up to billions of dollars every year. As the digital age prevails, more efficient manufacturing processes, based on computers and power electronics, have come to dominate industry. As the portion of electricity in the total energy consumption continues to grow, the value of power system reliability increases. This article discusses a vision for state-of-the-art solutions to improve power system reliability through improved monitoring and control.

The likelihood of blackouts has been increasing because of various physical and economic factors. These include (1) the demand for larger power transfers over longer distances, (2) insufficient investment in the transmission system, exacerbated by continued load growth, (3) huge swings in power flow patterns from one day to the next that render classic off-line planning studies ineffective, and (4) the consolidation of operating entities. These factors result in larger operational footprints and greater demands on the operator to deal with smaller error margins and shorter decision times. These circumstances have created a less reliable operating environment by pushing power systems close to their physical limits. Such an environment requires more intensive on-line analyses to better coordinate controls across the entire grid. Wide-area monitoring and control tools, eg, Phasor Measurement Units (PMU) and Flexible AC Transmission System (FACTS) devices, and distributed generation and storage devices are the primary technologies used to address such problems. The role of FACTS devices in measures to counter blackouts is described in [1].

ABB is a leading provider of such innovative and field-proven products and services to the utility industry. This article presents the results of an investigation by the authors (supported by the EPRI IntelliGrid consortium) into the requirements for the next generation of power system monitoring and control technologies. The evolution of these technologies is envisioned to lead to the realization of self-healing power grids. A self-healing grid is expected to respond to threats, material failures, and other destabilizing influences by preventing or containing the spread of disturbances [2]. This requires the following capabilities:

- Timely recognition of impending problems
- Redeployment of resources to minimize adverse impacts
- A fast and coordinated response to evolving disturbances
- Minimization of loss of service under any circumstances
- Minimization of time to reconfigure and restore service

In order to realize a self-healing grid, a high performance IT infrastructure will be needed to address gaps in the geographical and temporal coordination of power system monitoring and control. Current practices require considerable improvements at various hierarchical levels, including substations, control areas, regions and the grid. Temporal coordination will require improvements in adapting the faster and often local controls to the slower global controls.

ABB has developed the functional and architectural specifications of the IT infrastructure necessary for supporting a self-healing grid. This work included an assessment of its technical and financial feasibility [3, 4]. The remainder of this article comprises a brief overview of the results of the work.

Infrastructure for a self-healing grid

To attain a self-healing grid, it is essential to address a comprehensive set of operating concerns (in normal and abnormal conditions) that are associated with performance enhancement, adequacy of resources (market procurements, etc), and equipment and system operating limits (stability, sustained oscillations etc), as well as primary and back-up protection of systems and components.

ABB has developed the functional and architectural specifications of the IT infrastructure necessary for supporting a self-healing grid.

Existing on-line analytical capabilities are envisioned to continue playing their respective parts in the proposed infrastructure to address operating concerns. In addition, current off-line capabilities (eg, forecasting, dynamic analysis, transmission capability analysis) will migrate into the on-line environment. The details of their implementation will differ, as will the interdependencies in each of the functional areas of data acquisition and maintenance, monitoring, performance enhancement, and control actions. These functional areas have to provide non-stop service in terms of

- Providing situational awareness throughout the grid
- Predicting, preventing and containing problems
- Enforcing operational plans and required margins
- Supporting system restoration

These capabilities require the use of on-line decision support tools with intensive computational and communication requirements. The envisioned infrastructure calls for a distributed system in which the locations of hardware, software and data are transparent to the user. This enables autonomous intelligent agents distributed throughout the system to realize the required functions and to support local, global and/or cooperative processes through timely and effective information access throughout the system.

Architecture

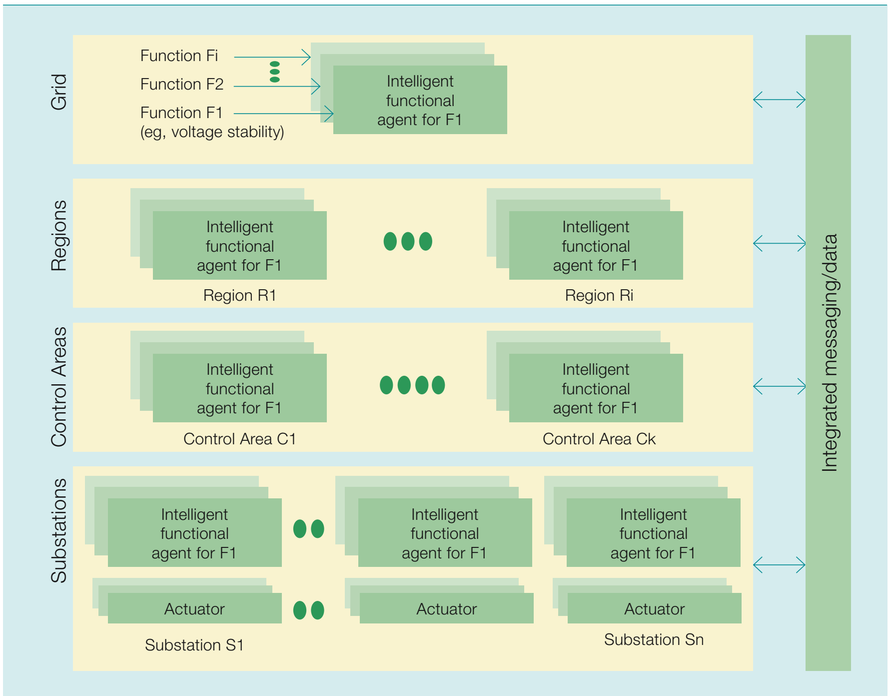
The IT infrastructure 1 must be modular, flexible, and scalable to meet global operational needs and to allow for evolutionary implementation on a continental scale.

The computing and communication systems of the infrastructure support a large number of computers and embedded processors scattered throughout the system. These must communicate with each other through networks with standardized interfaces that use message-oriented middleware and web services. The network would be dedicated to local and global data exchanges and decision processes using distributed databases that are integrated through open interfaces. The system would be constructed of plug-and-play hardware and software components.

The infrastructure supports a complex of software applications, including autonomous intelligent agents distributed throughout the system in a virtual hierarchy. They adapt to events and environments, and act both competitively and cooperatively for the good of the entire system. They can improve control performance by responding to problems faster than a human operator [5]. Thus the system supports more intelligence at all levels, especially at lower levels such as substations, to provide timely and accurate control responses.

The agents are distributed in a three-dimensional system to take into ac-

1 Infrastructure for distributed autonomous intelligent agents



count the geographical distribution and control hierarchy of the power system, as well as the diversity of functional areas. Various users and software components at different locations access and maintain the data (static and dynamic) that is distributed across the system in virtual relational databases.

Determining the degree of autonomy at each level, and the protocols for resolving conflicts between the levels, can be a major design challenge. Generally, the software at the higher-level needs to consider data for a larger portion of the power system. The software at the lower level can provide timely feedback and quick responses to local information, according to the most recent directives from the higher levels.

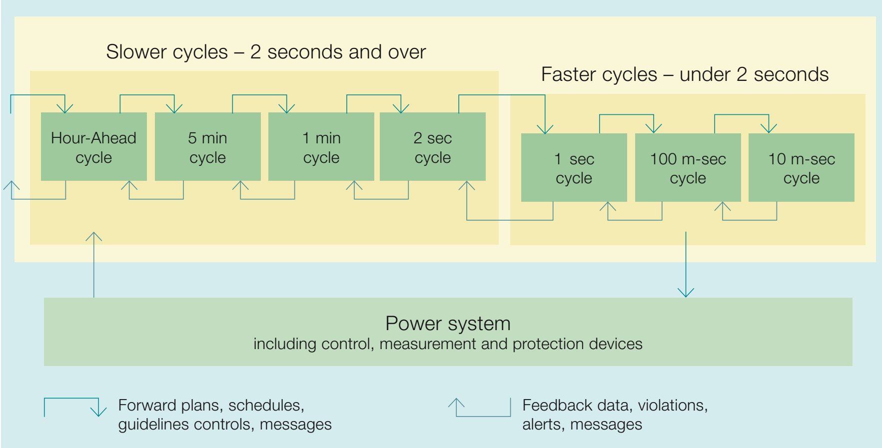
Each cycle can be tailored to the required control response times, computational burden, and historical practices.

Some existing Special Protection Systems and Remedial Action Schemes (SPS/RAS) may be seen as precursors of intelligent agents. Their effectiveness is expected to be improved by frequent tuning from a higher-level, as well as through better local analyses.

Coordination of tasks via execution cycles

There is also a temporal dimension in which the various agent tasks can be distributed, based on the time-scales of the relevant physical phenomena in the power system. This temporal coordination can be accomplished via several execution cycles. (An execution cycle refers to a set of related tasks executed in a temporally coordinated manner.) The execution cycles and their periods are defined according to operating needs and engineering judgment. Each cycle can be tailored to the required control response times, computational burden, and historical practices. The specific periods

2 Organization and interactions of execution cycles



Factbox Execution cycles for temporal coordination

Cycle	Purpose
1-hour-ahead	■ Assure adequacy of resources ■ Identify system bottlenecks
5-minute	■ Assure reliability, efficiency ■ Update control parameters and limits ■ Look-ahead (about 10 to 20 minutes) ■ Alert system operator and/or hour-ahead cycle
1-minute	■ Maintain efficiency and reliability, as per the 5-minute cycle. ■ Adapt the more recent models
2-second	■ Collect/validate data for use by control area or interconnection including data acquired in the 10-millisecond cycle (PMUs). ■ Perform closed loop controls (Area Generation Control, etc.) ■ Adapt control parameters and limits for faster cycles
1-second	■ Control extended transients (secondary voltage control, etc.). ■ Adapt control parameters and limits for faster cycles
100-millisecond	■ Control imminent system instabilities including execution of intelligent Special Protection Schemes (iSPS) based on adaptive models or criteria identified by slower cycles.
10-millisecond	■ Perform faster intelligent protection actions (load shedding, generation rejection, system separation)

Autonomous intelligent agents

In computer science, a software agent acts "on behalf of" a user or a program in a relationship of agency with the authority to decide when (and if) action is appropriate. The idea is that agents are not strictly invoked for a task, but are persistent so as to activate themselves, depending on the perceived context.

Agents can be intelligent, ie, possess learning and reasoning skills, and autonomous with the capability to adapt the way in which they achieve their objectives without human intervention. They can be distributed on physically distinct machines, as necessary, and could be mobile so that their execution would be transferability to different processors. Multi-agent systems consist of distributed agents that achieve an objective cooperatively. They can execute their tasks synchronously or asynchronously, and access decentralized databases as needed.

The design of agent-based systems should consider means to provide the capability to a) prioritize, schedule and/or synchronize tasks, b) facilitate communication and collaboration using appropriate ontology for representing knowledge and metadata in a hierarchical organization, and c) detect and respond to all possible changes in the environment.

and activities of the execution cycles can be configured according to the relevant operating concerns. These cycles cover time-scales ranging from 10 milliseconds to an hour. The exact periods of the cycles may be different in each implementation. A representative set of execution cycles is presented in the Factbox on page 23.

Based on the latency of real-time data acquisition, the cycles can be categorized as slow or fast. In the foreseeable future, the communications technology will impose a qualitative dichotomy at about 2 seconds 2.

Each functional agent is composed of plug-and-play building blocks.

The slower cycles perform the intensive computations required for system-wide coordinated controls, performance optimization, and control strategies. The faster cycles address local (substation and vicinity) analytical needs to respond to rapid events using the control strategies developed by the slower cycles. Higher-level intelligence is more prominent in the slower cycles, while the lower level intelligence is dominant in the faster cycles. The execution cycles interact with each other through exchange of event triggers, control parameters, problem indicators, contingency alerts, etc.

Each execution cycle includes a number of functional agents. Each functional agent is composed of plug-and-play building blocks called components, which may be reused in other contexts.

State estimation: a pre-requisite for self-healing capabilities

In contemporary control centers, most analytical functions are limited to slow cycles. For example, the state estimator (SE) is a key function that provides a refined snapshot of the steady state operating condition by minimizing the effects of errors in available data. The results of a state estimator are used not only by the operator, but also by various analytical functions in the slower cycles. These analytical functions need solutions for ever-larger networks with little time skew to support the emerging needs in market operations, in addition to meeting traditional reliability requirements. Similar needs arise in faster cycles to provide self-healing capabilities.

In order to meet these emerging needs, SE should be implemented as a cooperative solution by distributed agents. Each agent can inform other agents of the state of its own portion of the power system at any specified time, with an accuracy of a few milliseconds.

An SE agent at a substation retrieves data from the substation and other substations within an "electrical" vicinity, defined in real-time by an agent at a higher level.

An SE agent at a control-area receives data for a prescribed time from all substation agents in the area and assembles a solution for that area. This requires dealing with issues of geographical and temporal coordination at boarders of the different areas. Similarly, SE Agents at regional and higher levels have to coordinate solutions from the various control areas.

Such implementation of the SE function as a cooperative solution limits the time-skews to a few milliseconds, regardless of the size of the system (assuming PMUs are utilized for all measurements). The sub-second SE capability is essential to support the required faster local (ie, substation, etc) control. Local SE validation improves quality of the SE solutions at the higher levels.

Enhanced graphics should seamlessly combine navigation and presentation of information using animation and three-dimensional capabilities.

Effective visualization of information should allow the operator to understand the state of the system at a glance, and respond in a timely manner. Such situational awareness is an integral part of analysis and control. In addition to the current status, projection of trends, forecasted changes and look-ahead scenarios should be presented. Various views of the same object may be needed to present different aspects of the system to numerous users with diverse needs.

Enhanced graphics should seamlessly combine navigation and presentation

Prevailing viewpoints on the prevention of blackouts

The essence of a self-healing grid is its ability to prevent or contain major disturbances in power supply and to recover from problems in a timely manner. There are three prevailing philosophical viewpoints on major disturbances and the efforts to mitigate them [6].

Disasters are bound to happen: Power system disturbances are just random events that cannot be controlled by human intervention.

Strengthen the weakest point: Each disturbance exposes the "weakest link" in the system at the time it occurs. With the subsequent reinforcements, the next disturbance would have to be greater in scope to expose the next "weakest link" in the reinforced system. Thus, the cycle repeats.

Contain problems via better engineering: Proponents of this viewpoint believe that the complexity of the power system can be managed so as to predict and prevent (or contain) the problems before they become too large.

of information using animation and three-dimensional capabilities. Over time, such features may evolve to create a "virtual reality" environment.

In conclusion

In accordance with the approach of containing problems via better engineering, the envisioned framework requires deployment of autonomous intelligent functional agents across an interconnection-wide system to support a self-healing power grid. This would enable the system to adapt to the varying operating conditions of the system for analyzing and maintaining its reliability in real-time and in the near future. The interactions between the intelligent components of the infrastructure would be orchestrated through a set of execution cycles, tailored to the physical phenomena and operating concerns in the power system.

The immediate benefits are improved economics through lower congestion and unserved energy resulting from relieving operational limits and reducing interruptions.

More robust monitoring and control capabilities, realized through coordinated local and global controls, provide the resilience needed to deliver non-stop service and a greater degree of automation. The split-second local control decisions made under extreme emergency conditions would be faster and more robust than would be possible using only the operators or higher level controls.

This new infrastructure can be realized using existing technologies. All enabling technologies called for are either in use or proven in concept. Most of the necessary analytical techniques are already being used in various off-line and on-line design processes, eg, protection systems, generator controls, and system operating limits, though some improvement in speed, degree of automation, and level of distribution and coordination will be required.

The proposed system could be implemented in an evolutionary manner, starting with the realization of self-healing capabilities for the "backbone" of the grid. These capabilities could then be extended to additional portions of the transmission system as business needs and budgetary constraints allow. The development of the state estimation function discussed above would serve as a foundation for the overall realization of the required infrastructure.

Our work in this area has analyzed the functional, architectural, and financial feasibility of the proposed infrastructure. The authors have developed a methodology for evaluating the return on investment for this infrastructure, considering the costs related to hardware and software. The immediate benefits are improved economics through lower congestion and unserved energy resulting from relieving operational limits and reducing interruptions. ABB's efforts to enhance products and technologies for improving the reliability of power systems will continue to advance the industry towards the realization of a self-healing power grid.

ABB Network Management products

ABB is a leading provider of state-of-the-art power equipment, systems and services to improve power system reliability. Advanced monitoring and control is an essential part of efforts to improve power system reliability and cost effectiveness.

ABB business units for Network Management and Substation Automation.
These business units provide devices and turn-key systems for monitoring and control. They have been at the forefront of IT system development to facilitate seamless integration of transmission and distribution (T&D) operations. The resulting integrated platform supports systems for distribution and outage management, Supervisory Control and Data Acquisition (SCADA), energy management, and energy market operations. This includes a group of advanced technologies for Wide Area Monitoring Systems (WAMS). WAMS utilize phasor measurement readings of grid conditions at strategic points across a very large area. With precise time-stamps at their points of origin, the readings can be used to provide an accurate picture of the grid, far beyond any single control area, and to support faster coordinated control actions.

Khosrow Moslehi
ABB Network Management
Santa Clara, California
khosrow.moslehi@us.abb.com

Ranjit Kumar
Project Consultant
ABB Network Management
Santa Clara, California

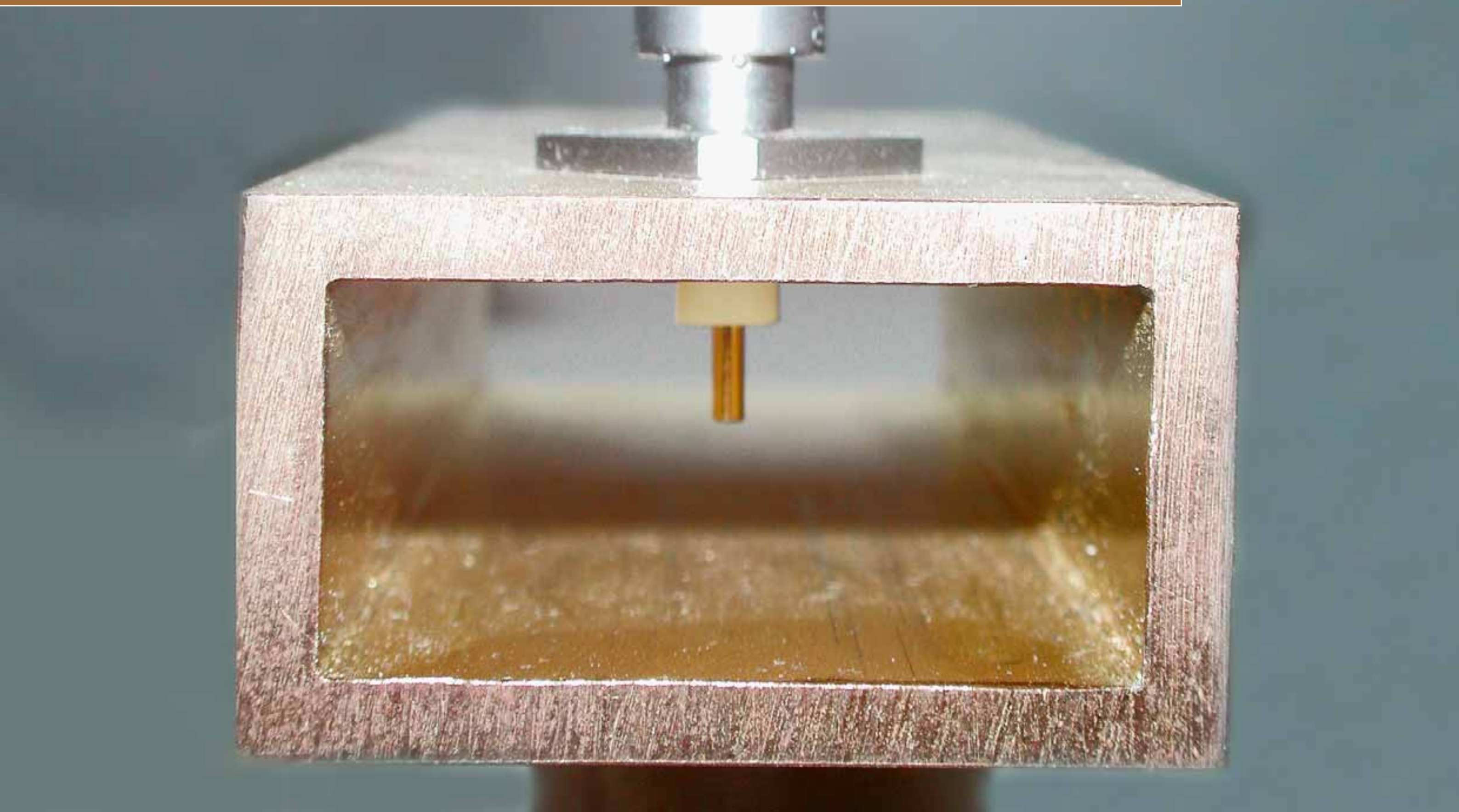
References

[1] Pourbeik, P., Bahrman, M., John, E., and Wong, W. (2006) Modern Countermeasures to Blackouts, IEEE Power & Energy Magazine, 4 (5), pp. 36–45.
[2] Amin, M., S. and Wollenberg, B. (2005) Toward a Smart Grid, IEEE Power & Energy Magazine, 3 (5), pp. 34–41.
[3] Moslehi K., Kumar, R., Chiang, H-D., Laufenberg, M., Bose, A., Hirsch, P., and Beard, L. (2004) Control Approach for Self-Healing Power Systems: A Conceptual Overview, Presented at the Electricity Transmission in Deregulated Markets: Challenges, Opportunities, and Necessary R&D, Carnegie Mellon University, Dec. 15–16, 2004
[4] Moslehi K., Kumar, R., Hirsch, P. (2006) Feasibility of a Self-Healing Grid – Part I Methodology and Cost Models, Part II Benefit Models and Analysis, Presented at IEEE PES General Meeting – Montreal, Jun. 2006.
[5] Rehtanz, C. (2003) Autonomous Systems and Intelligent Agents in Power System Control and Operation. Springer-Verlag, New York.
[6] Fairley, P. (2004) The Unruly Power Grid, IEEE Spectrum, Aug. 2004, pp. 22–27.

Data pipeline

Transmission technology for communication in MV switchgear

Kornel Scherrer, Bernhard Deck, Andreas Reimüller



From the point of view of electromagnetic compatibility, switchgear presents a highly challenging environment. The correct functioning of the equipment, however, depends on reliable communication between its components. Copper wiring is prone to electromagnetic interference whereas optical fibres are fragile and expensive to install. Both also present a risk of cabling errors. As an alternative, ABB is proposing broadband radio-based communication using hollow metal conductors – Waveguide. Radio signals are transmitted through these conductors where they are immune to external interference and do not themselves interfere with other devices. Furthermore, such conductors can be used by multiple communication channels simultaneously and are robust and easy to install.

In view of the introduction of the IEC 61850 series of standards, ABB considers that the time has come to encourage a paradigm shift in switchgear communication. This change will not stop at the transmission medium. The solution proposed is a closed-system, low-power broadband communication using Waveguides.

Wireless LAN radio technology in combination with IEC 61850

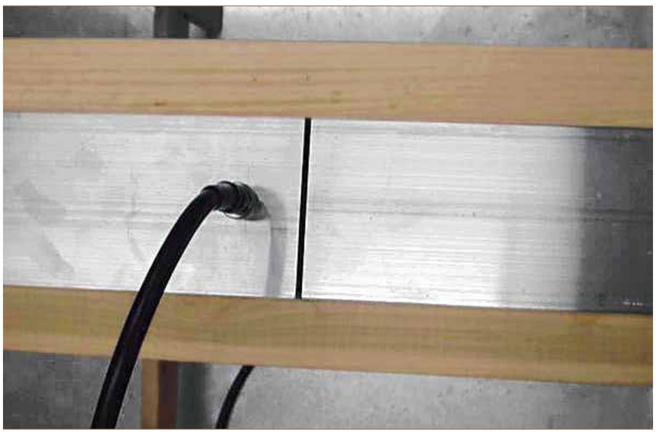
The world of standards has changed and expanded in recent years. The last major innovation appeared on the market with IEC 61850, and more are awaited with IEC 62271-1. The IEC 61850 series of standards, "Communication Networks and Systems in Substations", is being increasingly adopted in the global power engineering market. Initial pilot projects – still limited in their scope – are already in progress to implement these standards. Furthermore, an extensive exchange of experience on current solutions and present limitations of these standards is taking place. The future will require further rethinking in this context.

In a joint development effort involving several divisions of ABB, an innovative (although proven in other industries) data transmission method has been established. This will allow data within switchgear installations to be transmitted with high reliability and simplicity. The approach relies on the use of high frequency electromagnetic waves of low power in a closed system. The principle is simple and versatile. What is needed is a correctly dimensioned Waveguide, a probe that can both receive and transmit, and a coaxial connection to the protection and control device used (see title picture and 1).

Radio technology has been an indispensable part of modern life for a long time. There are hardly any areas in which data is not transmitted through the air. Wireless communication technology is in use everywhere, for radio itself, television, telephones and network connections.

1 Waveguide is set to become an integral part of medium-voltage switchgear.

2 Where Waveguide sections are joined, a narrow sleeve is used. This assures galvanic isolation without any degradation of transmission.



The present state of the art in transmission systems for all these applications is the digital, wireless network. The first radio signal was transmitted as early as 1886 by Heinrich Hertz using a spark gap. The development of antennas replaced the spark gap in the course of time, but radio operators are still nicknamed "sparks" today.

If radio signals are to be propagated in a Waveguide with low attenuation, a few simple conditions must be fulfilled. The dimensions of the conductor and the frequency of the radio signal, for example, have to be matched. This means that signals are transmitted with low attenuation only when a certain limit frequency is exceeded. Wavelength and frequency are inversely proportional, and so the wavelength λ has to be smaller than the limit wavelength λ_{Limit}. The following formula is used for the dimensioning of the system:

$$\lambda \leq \lambda_{Limit} = 2 \cdot x \qquad (1)$$

where x is the width of the rectangular Waveguide.

State of the art

In today's switchgear, the internal communications links (substation bus) are normally established from panel to panel – irrespective of whether serial or binary signals are to be transmitted. With parallel wiring, this is appropriately implemented using loop lines that are plugged into the terminal strip of each panel. Depending on the size of the installation, the complexity of the interlock system and the operator's need for control functions and information, it can involve cable harnesses with over 60 individual cores. Adding or changing signals in the loop lines immediately implies complex rewiring. The work required and the corresponding tests entail high costs for the operator. Only the use of serial communications techniques allows the number of loop line cores to be reduced. With this transmission method, the signals and measured value data are as a rule transmitted serially to a central point. Furthermore, the control commands can be transmitted to the relevant medium voltage panels.

In response to various ambient influences in switchgear applications, optical fibers (optical waveguides) have become established as the preferred transmission medium. In contrast to copper alloys, optical fiber cables are insensitive to electromagnetic interference, but they are also more costly due to their mechanical properties, the greater amount of work involved in installation (eg, fitting of plugs, protected routing in conduits) and the need for specialized tools and skilled personnel.

> Waveguide relies on the use of high frequency electromagnetic waves of low power in a closed system.

Indoor switchgear with Waveguides

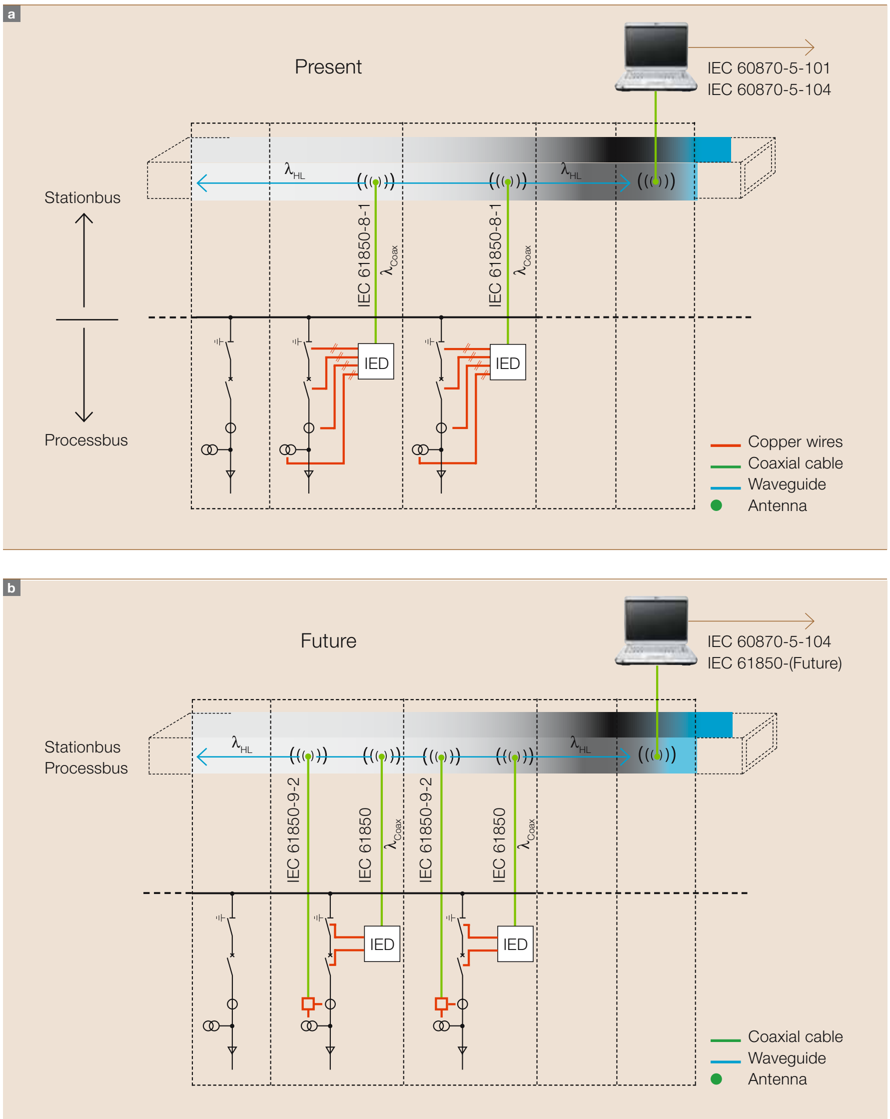
The demands for reliability and durability of systems are continously increasing. Therefore, it became an objective of ABB to find a more suitable transmission medium for panel to panel communication that would fulfill the criteria outlined above, and at the same time, would be simple to use. In addition, this medium had to satisfy the new requirements of the IEC 61850 series of standards. The achievable bandwidth had to be in the range of an optical waveguide (fiber optical cable), but installation had to be significantly easier. The advantage of electrical isolation between the data transmitters and receivers, which is guaranteed by the material of optical waveguides (but not copper conductors), was also to be retained in the new system.

The principle of the Waveguide is very simple and versatile in its application. Similarly to the optical waveguide, reflection from one, or rather two, parallel boundary surfaces is required for quasi loss-free transmission of the signal. To achieve this, the two walls (boundary surfaces) must be at a defined distance from each other, which is calculated from the wavelength used (see equation 1). Input and output of the signals are accomplished by spherical antennas. Shielded coaxial cable is used to bridge the short distances between the protection and control devices. and the Waveguide. Factbox shows a comparison between the various transmission media.

> The dimensions of the conductor and the frequency of the radio signal, have to be matched.

The electrical energy which is injected into a Waveguide by means of an antenna (probe) builds up an electromagnetic wave with E and H (electrical and magnetical) fields inside the conductor. As soon as the limit frequency for the particular system is exceeded, an electromagnetic wave spreads out in the Waveguide and propagates at almost the speed of light. On input, first an E field is created, which results in an H field. Waveguide antennae are in principle reversible, ie, can be used both to transmit and receive HF energy. If the Waveguide is correctly dimensioned,

3 Waveguide solutions – a present, and b future.



the electromagnetic waves are propagated almost without losses (attenuation approx. 2 dB/km). With the form selected, a low power 5 GHz signal is used. The technology employed corresponds to that of modern wireless local area network (LAN) systems. With a Waveguide, the radio signals are optimally protected from external interference and vice versa – the environment is protected from the radio signals.

The Waveguide segments are arranged in such a way in the low voltage compartments of the switchgear that they are automatically connected together when the panels are installed. Using Waveguide technology, the work involved in establishing panel to panel connections during site installation of a switchgear system is reduced to a minimum, in comparison with conventional loop line systems (typically up to 60 cores). When the panels are joined together, the sections of Waveguide are lined up with each other. The small gaps between the Waveguide sections in each panel unit are hermetically sealed with sleeves 2 so that no contamination from outside can enter the conductor. At the wavelength used, the gap has no adverse effects on the attenuation of the transmission system. In direct comparison with a conventional shielded Ethernet cable, the Waveguide is mechanically more robust, shielded from high frequency interference and, in contrast to cables, electrically isolated panel by panel (similarly to an optical waveguide connection). With this "plug and play" system, the entire communications system can easily be tested during inspection at the works.

Two Waveguide systems a few meters apart, eg, a switchgear system installed on opposite sides of the substation, can be connected by means of a passive system consisting of antennas and coaxial cables. From the point of view of network topology, a redundant network would have to be structured in such a way that failure of either a switch or the connection could be tolerated (n-1 principle). In direct analogy to copper or optical waveguide communications systems, this problem is reduced to duplicating the Waveguide access point, as it may be assumed in this case that a communications link via the Waveguide can be regarded as highly robust and therefore safe from failures. In this respect, a highly dependable network can be achieved at comparatively low cost 3.

With a Waveguide, the radio signals are optimally protected from external interference.

Prospects

The new IEC 61850 series of standards not only describes a simple communications interface on the basis of a substation bus. It also describes a process bus, which permits the connection of intelligent primary devices. These can, for example, be voltage and current sensors or transformers, or switching devices, which have a communications interface according to IEC 61850. When current and voltage measurements (sampled measurement values) according to IEC 61850-9-2:2004-04 [9, 10] are to be transmitted in real time from the sensor/instrument transformer to the Intelligent Electronic Device (IED), or distributed horizontally among the IEDs in a substation (eg, for busbar protection), a robust communications link is of decisive importance. Furthermore, it must be ensured that the physical connection provides sufficient bandwidth for fast transmission, so that no delay that would have an adverse effect on the protection system can occur. With a Waveguide connection, a large bandwidth is achieved using multichannel technology. In this way, up to 24 independent channels, each with 56 Mbit/s, can be connected to the Waveguide. This design permits not only transmission of vertical and horizontal information to IEC 61850, but additionally allows further services to be implemented via the system. Electricity meters installed in the switchgear can, for example, be read via the Waveguide, or web-based services implemented. Coupling of other active components can be achieved with a corresponding media converter. Especially for sampled measurement values, the Waveguide provides a connection that satisfies the safety demands of a protection system and the technical requirements of IEC 61850-9-2:2004-04 [9, 10].

Factbox Waveguide in comparison with other media

medium	system availability	EMC	thermal resilience	mechanical resilience	installation	extendebaility
electrical connection (Twisted Pair CAT 5, RJ 45)	+	-	+	+	0	0
optical fiber (Multimode 2G 62.5/125µm)	+	+	0	-	-	-
waveguide	++	+	+	++	+	++

Kornel Scherrer
ABB Management Services Ltd.
Zürich, Switzerland
kornel.scherrer@ch.abb.com

Bernhard Deck
ABB Medium Voltage
Baden-Dättwil, Switzerland
bernhard.deck@ch.abb.com

Andreas Reimüller
ABB AG, Calor Emag Medium Voltage Products
Ratingen, Germany
andreas.reimueller@de.abb.com

More than meets the eye

Beyond IEC 61850 as a pure communications standard

Christian Frei, Tatjana Kostic

In a typical energy IT environment, how do you get several components and devices from different manufacturers to operate seamlessly together? The answer lies in a cost-effective standardized communication interface that enables open infrastructures and the interoperability of devices.

One such standard, IEC 61850, is regarded as an important international standard for substation automation systems because it defines the communication between devices in the substation, as well as the related system requirements and the underlying data model.

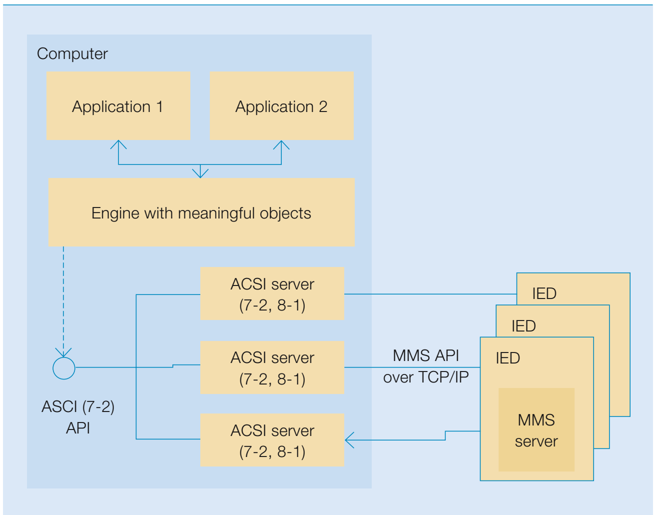
ABB Corporate Research is making use of these IEC 61850 features to provide plug and play technology for substation automation applications.

Electrical substations are complex distributed systems – containing heterogeneous primary equipment, such as switchgear, transformers or lines – and are controlled by Substation Automation Systems (SAS). An SAS is composed of all the electronic equipment[1] needed to continuously control, monitor, and protect the network. This equipment is interconnected and has to communicate at different levels in and outside the substation (eg, towards the network control centre). Traditionally, the engineering and configuration efforts for an SAS have been intensive. These efforts have also proven costly in that (a) customer delivery costs are driven by the missing standardized configuration description among devices from different vendors, and (b) product development costs are driven by the number of proprietary or semi-standard communications protocols that must be supported.

Thanks to the recently adopted IEC 61850 standard [1] – and, in particular, two unique features of the standard – SAS engineering and configuration costs have been significantly reduced. These unique features include self-describing Intelligent Electronic Devices (IEDs)[2] at run-time (because of their data and communications service models), and a standardized substation configuration language (SCL).

These two features allow ABB to employ a plug and play[3] approach to substation automation applications to help alleviate costly project engineering for many customers.

1 A simplified view of the architecture of a plug and play application based on IEC 61850. Numbers in brackets denote the corresponding part in the standard. Examples of applications include basic monitoring or control systems



The IEC 61850 standard

IEC 61850[4] is a substation automation communication standard. Its main objective is to enable seamless interoperability of IEDs (especially from different original equipment manufacturers) within the substation. It is divided into 10 parts, which circle around four major issues: a functional model of the substation automation application domain (Part 5), a data model for SAS (Part 7), communications protocols and their services (Parts 7, 8 and 9), and an XML based Substation Configuration description Language (SCL – Part 6) [2].

Thanks to the substation automation communication standard IEC 61850, SAS engineering costs have been significantly reduced.

A model generally provides some information about a piece of equipment or a process. In substation communications, a model that lists the input and output data of a transformer, for example, is known as a data model and for each function, IEC 61850 has a data model. A data item has a name, which comprises three standardized parts namely, the Logical Node (LN), the data object and the attribute[5].

Since the IEC 61850 standard defines the entire data model in a text-and-table format, Unified Modelling Language (UML)[6] [3] has been used to develop a formal model [4,5] of it. Throughout the rest of this article, this is referred to as *the UML model*.

In a substation automation system, there are 2 types of data exchange: run-time application data exchange and (off-line) configuration data exchange. The run-time data exchange is enabled through communications services. IEC 61850 part 7-2 defines a set of abstract communication services (ACSI) that address the basic requirements for the process of exchanging information. These services can be implemented as an Application Programming Interface (API) specified in a given programming language, and/or services of the application layer of a given communications stack. Examples of this kind of exchange include reading the current position of a circuit breaker, or an operator command, or a protection function trip to change the breaker's position at run-time. These services and their implementation are the "how" of the data exchange.

The second type of data exchange concerns configuration data among all

Footnotes

[1] In IEC 61850 talk, an SAS is usually composed of Intelligent Electronic Devices (IEDs) which are connected by a communications network.

[2] Self-description of an IED means its ability to provide certain information such as process status information or configuration information of protection settings (the latter type of data is typically not available in previous SA communications standards).

[3] A plug and play application means it can discover the functionality of one or more IEDs at runtime without any configuration required, and then generate a corresponding graphical user interface based on these findings.

[4] ABB is very involved in the definition and maintenance of the IEC 61850 standard (several editors of the standard are employed by ABB), and as a result it is a leading vendor in that field. Today, this is vital as many new projects, especially in China and India, require its use.

[5] The Logical Node, the data object and the attribute are illustrated by the example of a circuit breaker. The status of the circuit breaker – using IEC61850 terminology – has the name: XCBR.Pos.stVal. XCBR (the circuit breaker) is the logical node. Pos (switch position) is the data object and stVal (Intermediate state, on, off, bad state) is the attribute.

[6] The de facto modelling standard in software engineering.

SA applications and devices. It is performed using XML (eXtensible Markup Language) files that follow the SCL[7] outlined in the standard. These XML files describe the configuration of IEDs in terms of functionality (eg, circuit breaker control, measurements and status values), communication addresses and means (eg, fast multicast messages, reporting), as well as the switch yard layout and its relation to the functions implemented in the IEDs.

Intelligent plug and play applications

In the following paragraphs, It is assumed an IED is fully compliant with the IEC 61850 standard. In other words, it should provide an ACSI-compliant interface to its data. Intelligent, and plug and play applications can rely on the interoperability features normalized by IEC 61850. The discovery and retrieval of actual values for process and configuration data is performed as follows:

1. Plug the computer with the running application into the IEC 61850 network.
2. Provide the IP address of the server residing within an IED. (This step is the only one not considered fully plug and play)
3. The application retrieves the variables using ACSI directory services.
4. The user interface of the application is auto-generated according to these variables and their semantics.
5. The application acquires current values for all variables of interest
6. If updated values of process data are needed, the application subscribes to adequate events.

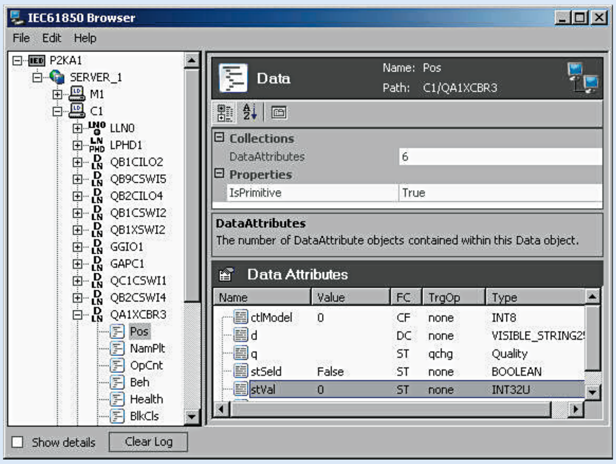
Steps (2) to (6) can be repeated for as many servers as is needed. Because step (2) is not considered fully plug and play, a network scan can be performed to detect different servers – which respond on a predefined port – provided security mechanisms in each IED allow this.

1 shows the outline of a plug and play system architecture. The ACSI servers provide services as defined in Part 7-2 of the standard, and they allow services to be abstracted independently of their specific implementation. They also permit the "browsing" of variables, the retrieval and setting (where applicable) of certain values, and finally, updated process data through an eventing mechanism can be received. The engine is the "semantic" interface between ACSI and the application. It takes care of transforming MMS[8] variable names into meaningful objects according to Parts 7-4 and 7-3. The access rights – read, write, or read/write – are automatically identified for each data attribute. The engine is also responsible for catching any event providing live process data and then updating the corresponding values. An application, for instance a basic monitoring system, relies on the engine to acquire the servers' contents and to get and set values. As the type of data is clearly identified, the engine can provide views for process data (current values, analogue or digital), or parametrization (of protection functions). Furthermore, events can be logged and displayed to the user, and some can even be automatically interpreted as alarms.

> Key components of the plug and play architecture are becoming an integral part of existing ABB SAS products.

Some prototype applications[9], developed at ABB's corporate research center in Switzerland as part of the "Zero-Configuration Substation Monitoring System" project, are described briefly below.

2 An IEC 61850 browser: an application that needs no configuration, but can discover the contents of an IED, analyze it semantically and auto-generate the display.



A plug and play IEC 61850 browser

A screenshot of a plug and play IEC 61850 browser is shown in 2. A very basic monitoring system can be built using only the IP address(es) of the device(s). Equipped with this IP address, a user can access the contents of the IED automatically and autonomously, ie, without prior configuration of the software. In the example in 2, the contents of the Logical Node *QA1XCBR3* (reflecting the information of a circuit breaker) is shown together with its data object *Pos* (the position of the breaker) and the data attributes (with their names, types, values and other properties) belonging to that data object. None of these properties is hard-coded but is instead generated from the semantic information the application possesses, and which comes from the formal data model expressed through the UML model.

A plug and play IED configuration comparator.
Another immediate application built on the top of the engine is a plug and play IED configuration comparator. It compares the actual contents of an IED with its configuration, as described in the corresponding SCL file. The usefulness of this tool is particulary evident during commissioning or maintenance, when the engineer needs to verify that the actual IED configuration matches that provided by the corresponding SCL file.

A plug and play IED configuration generator.
Normally, SCL files are generated by the engineering tool. However, it is possible to generate an SCL file automatically. This is particularly useful during retrofit as most substation automation systems do not have an SCL file describing them (or if one does exist, it may be outdated). The generated SCL file can then be imported into the engineering tool for further processing, thereby avoiding tedious and error-prone work.

The example applications described above illustrate the possibility of building plug and play applications based on the self-describing capability of IEDs at run time, using only the IP address (or a range of them) as input. However, there are limitations, from a system perspective, on what can be done using the purely data model and ACSI services (as defined in Part 7 of the standard), since they focus only on individual IEDs. As a result, it is impossible to bind the discovered functions (ie, logical nodes) to primary equipment automatically or to deduce the substation layout, or even to analyze the inter-device communication network. In other words, given solely the IP address (or their range), only a very basic monitoring system focusing on individual IEDs can be built.

IEC 61850 is much more than just a communications standard as it also defines a data model and services to operate on that data.

To extend the capabilities of the monitoring system, an additional complete and up-to-date SCL file is required if the substation automation system, as a whole, is to be understood. If, however, an application requires a user to provide such an SCL file, this application is no longer considered fully plug and play. A possible solution might be to host the extra SCL file at a predefined address on a gateway. If that were the case, then the application described below would be possible.

A plug and play (basic) substation monitoring system.
The IEC 61850 browser described above can discover a substation automation configuration of one or more IEDs. Thanks to the data model defined in the standard, status and measurement data points are easily identified, and therefore a graphical user interface can be built to automatically display all process (status and measurement) data. Additionally, subscription to an eventing service that provides updated values can be done automatically. Furthermore, if an SCL file with switch yard description (ie, a substation section) is provided, a simple, single-line diagram incorporating process data can be built and displayed.

Conclusion

Despite its name, IEC 61850 is much more than just a communications standard as it also defines a data model and services to operate on those data. ABB's Corporate Research has exploited these features to develop the underlying technology that enables plug and play applications (for instance, deducing IED configuration automatically). Of course, very complex systems such as a SCADA can never be fully plug and play and would still require some engineering. However, the engineering required can be simplified using the proposed approaches, resulting in reduced factory acceptance tests, commissioning time, and maintenance.

Key components of the plug and play architecture are becoming an integral part of existing ABB SAS products. This, together with the complexity and the increasing importance of IEC 61850, has triggered the joint development of tools across ABB.

Christian Frei
Tatjana Kostic
ABB Corporate Research
Baden-Dättwil, Switzerland.
christian.frei@ch.abb.com
tatjana.kostic@ch.abb.com

ABB publications about the IEC 61850 standard can be accessed at: http://www.abb.com/cawp/seitp202/C1256A8C00499292C1256D4100388F27.aspx

Footnotes

[7] The SCL defines the interrelationship of the substation equipment to each other and to the substation itself.

[8] Manufacturing Message Specification: ISO 9506-1 and ISO 9506-2: Industrial automation systems – Manufacturing Message Specification; first edition, 2000-08-15

[9] Some of these applications are currently in the process of being transferred to ABB's Business Units for product integration.

References

[1] IEC 61850: Communications Networks and Systems in Substations, International Standard, 2003.

[2] O. Preiss and A. Wegmann, "Towards a Composition Model Problem Based on IEC61850", The Journal of Systems and Software, Vol. 65/3, Elsevier Science, 2003, pp. 227–236.

[3] OMG, Unified Modelling Language Specification, Version 2.0, July 2005. http://www.omg.org/uml

[4] T. Kostic and O. Preiss, "UML model of the IEC 61850 (v6, May 2004) and the data mappings between CIM 10 and IEC 61850". Not yet published (contact authors).

[5] T. Kostic, O. Preiss, C. Frei, "Understanding and Using the IEC 61850: A Case for Meta-Modelling," Elsevier Journal of Computer Standards & Interfaces, vol. 27/6, pp. 679–695, 2005.

Power semiconductors

Part one: Basics and applications

Stefan Linder

Over the last 10–15 years, in the wake of rapid progress in semiconductor technology, silicon power switches have developed into highly efficient, reliable, and application-friendly devices. These devices have firmly established themselves in high-voltage and high-current applications, handling power outputs ranging from one megawatt to several gigawatts. Power semiconductor devices have started a quiet revolution, in the course of which, electromechanical solutions are gradually being improved by the addition of power electronics, or even completely replaced by power-electronic systems.

This article, which is intended for a relatively knowledgeable readership, is the first of a two-part contribution, in which ABB Review takes a tour of high-power semiconductors. In this first part, different classes of devices are presented, in particular the IGBT and IGCT. Specific advantages and disadvantages of these devices are compared, as are some important aspects concerning their application. In the second part, to be published in ABB Review 1/2007, thermal issues and aspects of housing design are discussed. Moreover, an attempt will be made to predict future developments, and what importance "wide bandgap" materials, such as SiC (silicon carbide), GaN (gallium nitride), and diamond, will gain in the high power arena.

Power highlights

It was the introduction of neutron transmutation technology in the 1970ies that made the manufacture of power semiconductor devices with blocking voltages of greater than circa 1000 V possible. Only this technique permits silicon to be produced with the required doping homogeneity.
At that time, the thyristor was the only device in this voltage class that had been properly mastered technologically. However, the number of applications was restricted, as this device did not permit forced current interruption at an arbitrary point in time. In the 1980ies and 1990ies, the thyristor was joined by devices with turn-off capability, namely the GTO (Gate Turn-Off Thyristor), and later the IGBT (Insulated Gate Bipolar Transistor) and the IGCT (Integrated Gate Commutated Thyristor). These devices increased the spectrum of efficiently operable task definitions significantly. Nowadays, thanks to these devices, variable speed drives in the megawatt range are state-of-the-art and it would be hard to imagine the power transmission and the grid stabilization sectors, where applications extend well into the gigawatt range, without solutions based on power semiconductor components.

In the last 10 years, the IGBT and the IGCT (which replaced the GTO) have been developed further with regard to losses, voltage withstand capability, current carrying capacity (SOA = Safe Operating Area), and ease of use. Therefore, the old paradigm that IGBTs are for "small" power outputs and IGCT for larger ones, which was still generally accepted in the mid 1990ies, no longer applies. IGBTs are now used successfully in applications with an output of over 300 MW [1]. The conclusion that the IGCT will lose its raison d'être as a result of the advance of the IGBT is incorrect, however, as is witnessed by the strong growth of applications, mainly in the medium-voltage range. The decision as to which component is most suitable for a desired application depends on a number of technical factors, which will be illuminated somewhat in this article. However, the know-how and experience of the user in making the selection should not be underestimated in this context. Since the performance and the reliability of semiconductor devices is strongly dependent on the service conditions and the physical design of the system (electric, thermal, mechanical), users will, whenever possible, make use of platforms where they have the most extensive experience.

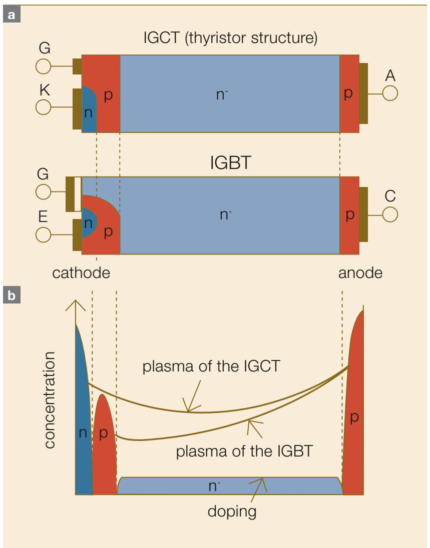
The old paradigm that IGBTs are for "small" power outputs and IGCT for larger ones no longer applies.

Design objectives of the IGBT and the IGCT

Introduction

The doping of the silicon body of power semiconductors, ie, the conductivity of the substrate, must be reduced continually as the targeted breakdown voltage increases. As a result, components that rely during on-state on their substrate conductivity (so-called unipolar or majority carrier components, eg, the power MOSFET and the Schottky diode), exhibit at blocking capabilities exceeding 200–1000 V an on-state voltage drop that is too high for economic operation (the limit depends on the type of component and application). Consequently, silicon power semiconductors above 600 V are usually designed as conductivity-modulating (plasma) devices. The interior of such a device is flooded with a large number of positive and negative charge carriers (holes and electrons) during the conductive phase, lending the device a strongly enhanced conductivity with respect to the substrate. Such components are frequently termed "bipolar components" in the power semiconductor industry, although the use of this expression is not strictly correct from the technical point-of-

1 Component structures and doping zones of the IGCT and the IGBT a and qualitative comparison of the plasma distributions in the conductive state b



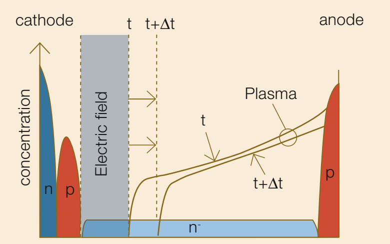
Semiconductors have become ubuiquitous in a wide range of applications, including power transmission a traction applications b and industrial drives c



view (this will be discussed further in part two of this article, to be published in the next edition of ABB Review).

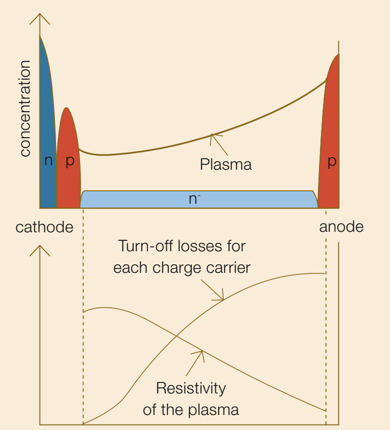
During device turn-off, in order to recover blocking capacity, the plasma must be removed. This is accomplished by the recovery voltage, whereby an electrical field builds up, driving the negatively charged electrons to the anode and the positively charged holes to the cathode. As a result, current still flows as the voltage increases, ie, losses arise in the form of heat.

2 Inside an IGBT/IGCT during turn-off



An electric field is built up at the pn-junction on the cathode side and sweeps out the plasma. The nearer the charge carriers are to the cathode, the lower the voltage at which they are removed.

3 Geometric dependency of the plasma conductivity and turn-off losses inside the IGBT.



By comparison with 1, it becomes evident that the IGCT has a better plasma distribution.

Optimization of forward-power losses and turn-off losses by adjustment of the plasma distribution

The common design objective of high-voltage power semiconductor switches (whose best-known types are the IGBT and the IGCT) is the optimization of the combination of conduction losses and turn-off losses. In practical terms, this means that the semiconductor should have the lowest possible voltage drop in the on-phase (ie, a dense plasma should build up) without high turn-off losses arising when the excess charge is removed.

> The minimum thickness of a power semiconductor is predetermined by the desired blocking capacity and the breakdown field strength of the silicon.

1 shows the typical plasma distribution of IGBT and IGCT components. The main difference between them is that the IGCT builds up a dense plasma near the cathode, while the excess charge density in the IGBT drops relatively sharply from the anode to the cathode. The reasons for this are explained later in this section.

The importance of this charge carrier distribution is illustrated by considering the turn-off process: during switching off, the component regains its blocking capability through the build-up of an electric field from the pn-junction on the cathode side into the n⁻ zone 2. The recovery voltage sweeps out the plasma from the cathode to the anode. The charge carriers near the cathode are removed at a low voltage, and therefore, generate low turn-off losses, whereas the carriers near the anode flow out of the device at a high voltage, causing high losses.

This consideration makes clear why the plasma distribution of the thyristor is also generally regarded as a desirable ideal for the IGBT: the voltage drop in the conducting mode is determined mainly by the region of the lowest plasma density, which explains why an IGBT has higher conduction losses than a comparable thyristor. Therefore, if the plasma of the IGBT at the cathode can be increased successfully, the on-state losses can be reduced without considerably higher turn-off losses arising 3.

The primary cause for the low plasma density at the cathode of the IGBT is a weak "carrier storage effect": the holes originally injected by the anode can enter the p-zone on the cathode side relatively easily, and from there, leave the component via the (required) contacting of the emitter to the p-zone without hindrance (see 1). In contrast to this, due to the non-existent contact to the p-zone, the thyristor has a pronounced carrier storage effect. The potential barrier of the pn-junction at the cathode contact prevents holes from entering the n-zone.

Two different concepts for the improvement of the plasma distribution in the IGBT have been proposed: one very effective option consists in the application of the trench principle [2], in which the holes are prevented from "finding" the p-zone by a skillfully designed geometric cathode structure. As an alternative, a weak potential barrier can be generated by means of a doping layer in front of the p-zone to keep the holes away from this zone [3]. A detailed explanation of these methods can be found in literature, eg [4].

Modern IGBTs, which are designed in accordance with one of these basic approaches, have correlations between conduction losses and breaking losses, which come very close to those of IGCTs. Although future improvements are possible, the latest designs (eg, the SPT+ from ABB [8]) have already been optimized so well that no great steps can be expected.

Lower losses through a reduction in the thickness

A reduction in the thickness of the components is the most effective parameter for reducing total losses. The reasons for this are simple: the resistance of the device in the conducting state decreases as a result of the lower thickness, and at the same time, there is less plasma in the device during the conductive phase, thus resulting in lower losses during turn-off.

The minimum thickness of a power semiconductor is predetermined by the desired blocking capacity and the breakdown field strength of the silicon. 4 shows two different aggressively designed devices with the same blocking capability.

It is apparent that the maximum blocking capacity for a given element thickness is obtained by means of a field strength distribution being as close as possible to the breakdown limit over the entire thickness.

The gradient of the field strength *dE/dx* 1 can be adjusted by the doping concentration in the silicon.

In practice, there are limits to the aggressive design of the field strength distribution, and therefore, to the minimum thickness of the devices:
1. If the doping concentration in the semiconductor is set at a very low level, the electric field extends over the entire thickness of the component, even at low voltage. The entire plasma can therefore be removed at a low voltage during turn-off. Although this is theoretically desirable (since turn-off losses decrease as a result), it also causes the current to break off abruptly on reaching a certain voltage (the point at which the device is cleared of plasma). This effect is referred to as snap-off. The high *di/dt* generates overvoltages in stray inductances and can initiate undesirable oscillations in combination with capacitances. 5 shows examples of a desirable ("soft") and a bad ("hard") turn-off waveform.

The stray inductance makes a greater difference in power semiconductors for high currents than in small discrete components. Firstly, the leakage inductance is higher on account of the physically larger assemblies, and secondly, the semiconductor experiences a much higher stress through a given stray inductance. To illustrate this, a hypothetical discrete 50 A IGBT chip is compared with a 1000 A module, the latter being assembled from 20 discrete 50 A chips. The stray inductance in the circuit with the discrete chip is assumed to be at 20 nH, and that of the module, 100 nH. The calculation of the stored inductive energy ($E_{ind} = LI^2/2$) shows that, at rated current, each chip of the module experiences an inductive load 100 times greater than the discrete chip (2.5 mJ, compared to 25 µJ). This shows that components for high power outputs have to be dimensioned for a much softer switching behavior than chips for small assemblies on printed circuit boards. In practical terms, engineers must make the components thicker than would be theoretically necessary. This naturally implies extra losses, as is shown in the example in 5.

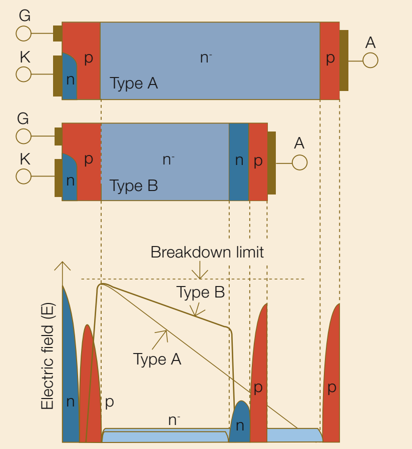
The semiconductor should have the lowest possible voltage drop in the on-phase without high turn-off losses arising when the excess charge is removed.

In addition to the design-in of a certain extra thickness, the snap-off can be reduced by a skilful arrangement of dopings on the anode side of the component. Manufacturers use different names for concepts that are similar (at least in their action), eg, SPT (Soft Punch Through) [5] or FS (Field Stop) [6].

It should also be pointed out that it is more important than ever for users to restrict the stray inductances in their systems as much as possible, on account of the more offensive design of modern components.

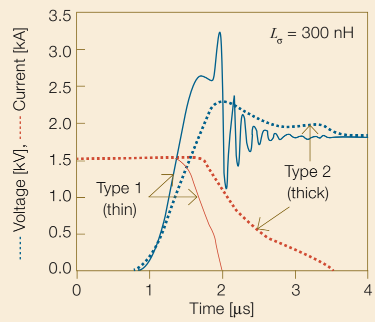
2. The second limitation is attributable to cosmic radiation. If a high-energy nuclear particle from space, eg, a proton, strikes a silicon nucleus, the released energy generates a very high quantity of electrons and holes. If the device is in blocking mode at high voltage, these carriers are multiplied in avalanche-like mode due to the high field strength in the component. This causes a highly localized breakdown of the component, which may result in the destruction of the device. Manufacturers have, therefore, developed dimensioning rules, according to which components must be designed with respect to thickness and field strength distribution to ensure that the probability of destruction by cosmic radiation is restricted to an acceptable degree. This is specified at approximately 1–3 FIT (Failures in time) per cm² of component surface area. This corresponds to 1–3 failures per billion operating hours and cm². Proof of the failure rate of new components is

4 Different vertical designs of a power semiconductor, through the example of a thyristor-structure.



The voltage across the device is proportional to the area under the electric field. The mid-section (n-) is usually referred to as the drift zone in unipolar components and the n- base in bipolar components.

5 Effect of snap-off during turn-off of a large 3.3 kV/1500 A IGBT module under the influence of a high stray inductance.

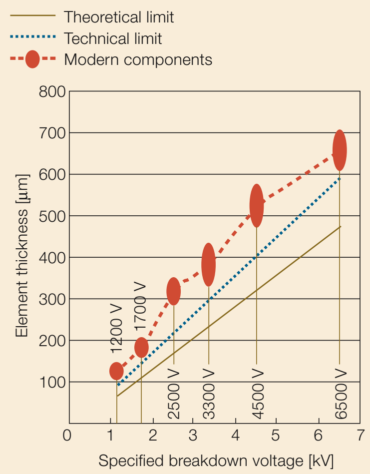


The "Type 1" IGBT is significantly thinner than the "Type 2" (340, resp. 380 µm, see also 6).

nowadays usually obtained by proton or neutron bombardment in accelerators, which simulates the effect of natural cosmic radiation with sufficient accuracy.

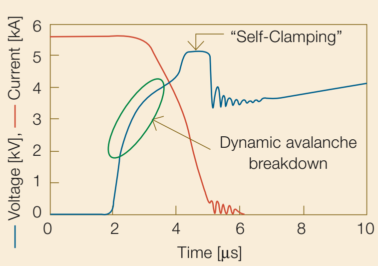
High-voltage components of the latest generation are already close to the practical limits with regard to thickness. 6 illustrates the position of the latest components in relation to the calculated theoretical limits. Although a further reduction in thickness below the current level would be theoretically possible, it would be at the expense of a more severe snap-off or significantly higher turn-off losses. At present, it is questionable whether users will accept such devices.

6 Comparison between the calculated theoretical minimum thickness of the component



Assuming no extra thickness, and breakdown occurring at room temperature – approximate technical minimum thicknesses (practically achievable values, without consideration of electrical bahavior), and the thicknesses of modern components (the red areas represent the different values of individual manufacturers).

7 Turn-off of an IGCT with an active area of 40 cm² under SOA conditions, without a protection (snubber) circuit.



The switching power density is more than 500 kW/cm². Avalanche breakdown first reduces the steepness of the voltage ramp, and then limits the overvoltage.

An IGBT can be controlled by the gate voltage during turn-on/turn-off, whereas the switching transients in the IGCT are governed only by the internal dynamics of the component.

Increase in turn-off capability (Safe Operating Area, SOA)

The useful output current of a power semiconductor is restricted both by the capability of the housing technology to dissipate power losses, and by the maximum current that can be safely controlled during turn-off. Part two of this article will deal with the housing technology in detail, whereas SOA aspects are covered here.

It was generally assumed in the 1990ies that the occurrence of a dynamic avalanche breakdown was an unsafe operating condition. Such a breakdown occurs if the power density (which is calculated as the current that can be turned off multiplied by the DC link voltage) reaches approximately 150 kW/cm².

From theoretical considerations, the conclusion that dynamic avalanche breakdown is unsafe cannot be maintained. On the contrary: the effect is self-limiting [4] and can, therefore, be considered harmless. Consequently, it makes sense for manufacturers to raise the destruction limit of the components to the highest possible level. Power densities of more than 1 mW/cm² have already been successfully demonstrated in all modern components (IGCT, IGBT and diodes). An example, showing that large components can safely control very high power outputs, is shown in 7.

Today, due to thermal limitations, it is hardly possible to operate components at more than approximately 100 kW/cm² RMS power. Therefore, the question is legitimate whether an SOA margin much beyond this limit is of practical importance. The answer is affirmative for the following reasons:

- In large-area power semiconductor devices, it cannot be assumed that the current flows uniformly through the semiconductor. Irregularities in the cooling, different coupling inductances and slightly varying semiconductor properties can lead to substantial temperature differences and inhomogenous electrical loads, the latter particularly during turn-on and off [7]. Substantial power margins can save components from failing under such conditions. Several large equipment manufacturers were able prove a causal connection between the power margins and the field reliability, even when the components were nominally operated within the specification limits.
- A high tolerance for dynamic avalanche breakdown prevents overvoltages beyond the specified nominal voltages (see 7).
- A high SOA power margin can be used to cope with rarely occurring overload conditions (eg, fault conditions). The high dissipated energies during such events can most often be tolerated, since a turn-off usually occurs only once.

Increase in the maximum junction temperature

The extension of the temperature limits is closely connected with the properties of the housing technology and is, therefore, discussed in more detail in part two of this article.

The IGCT and IGBT in comaparison

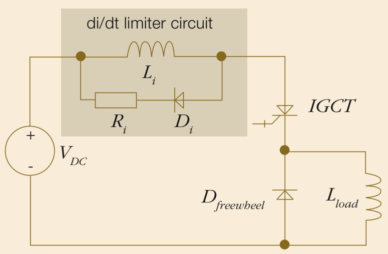
The lower driver power of the IGBT is frequently cited as a key advantage of this device as compared to the IGCT. The difference in driver power is attributable to the fact that the IGBT is controlled by a MOS input, whereas the IGCT is a current-controlled device. In practice, however, the differing power requirement is crucial only in a small number of applications, since the IGCT's driver power is low enough that it can usually be provided at an acceptable effort.

On the other hand, the most important application-related difference between an IGCT and an IGBT lies in the fact that the IGBT can be controlled by the gate voltage during turn-on/turn-off, whereas the switching transients in the IGCT are governed only by the internal dynamics of the component. This difference, which may not seem significant at first glance, has far-reaching consequences for the circuit topology and for applications where parallel and/or series connection is required.

Differences in circuit topology

On account of the IGCT's internal thyristor structure, the device can build up current very quickly during turn-on, ie, it produces a steep *di/dt*, which generates an unacceptably high stress in auxiliary diodes. Because of this, the *di/dt* in IGCT circuits must always be restrained by a limiter circuit. In voltage source inverters, this usually consists of a small inductance in series with the switch 8. Although this increases the complexity of the circuit, it has several beneficial characteristics:

8 IGCT phase test circuit



The current gradient permitted by the inductor L_i during turn-on reaches a maximum di/dt_{max} equal to UDC/L_i. The elements D_i and R_i form a freewheeling circuit for L_i, thus limiting the overvoltage during the turn-off of the IGCT.

1. In voltage source inverters without external *di/dt* limitation (typical IGBT circuits are examples of this), the *di/dt* must be limited by controlling the switching device itself. This causes substantial turn-on losses. In inverters with high voltage ratings, the combination of turn-on losses of the switch and recovery losses of the diode make up 40 to 60 percent of the total inverter losses, depending on the switching frequency. Significantly lower turn-on losses occur in a silicon switch upon the use of a passive *di/dt* limiter, relieving the device of thermal load and, by consequence, enabling in principle, a higher output power for the inverter. However, it should be noted that the losses occur nevertheless, since they are merely transferred to the freewheeling circuit of the *di/dt* limiter (they occur in the resistance R_i and the diode D_i of 8). The interpretation that an inverter with a *di/dt* limiter circuit always generates lower total losses than a conventional IGBT inverter is therefore incorrect.

In large-area power semiconductor devices, it cannot be assumed that the current flows uniformly through the semiconductor.

2. The second benefit is that, as a result of passive *di/dt* limitation, the current can increase only relatively slowly when a fault occurs (eg, a short-circuit in the inverter bridge or in the load). Hence, there are two effective strategies available for coping with such events: (a) If the fault is detected in time, it is possible to initiate a normal turn-off. (b) The energy stored in the DC link can be discharged by firing all switches, dispersing it evenly in all semiconductors (the inductance L_i can be dimensioned to mainatin the short-circuit current in the safe area).

Parallel and series connection

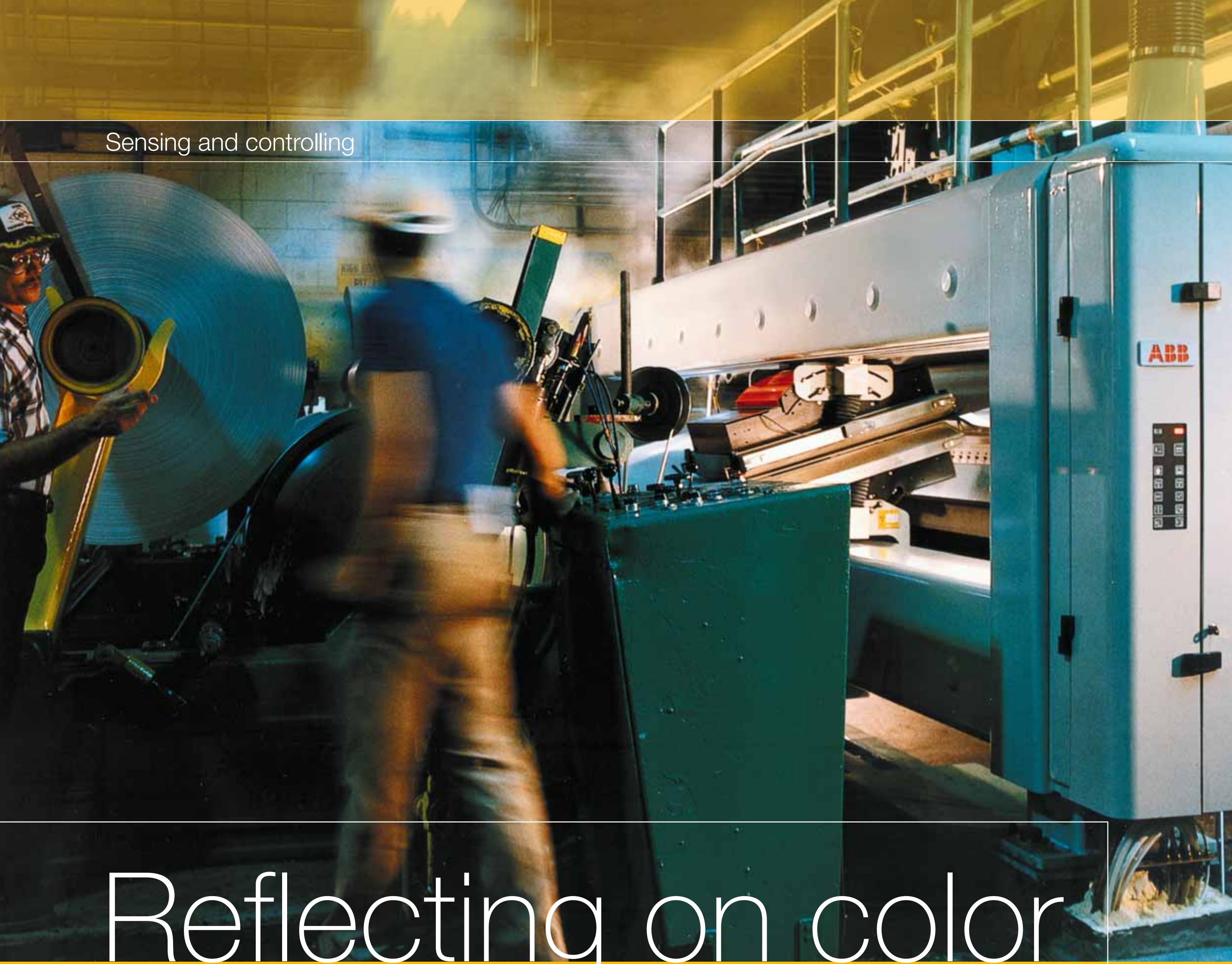
As the switching transients of an IGCT cannot be externally influenced, the gate control circuit must drive the whole device absolutely simultaneously to guarantee a homogeneous and therefore safe turn-off process. The tolerable time difference is less than 100 ns, which means that IGCTs can only be operated in parallel or series with a relatively large effort. In both cases, passive or active snubber circuits must compensate for even the smallest differences in switching times between IGCTs (caused by control timing errors and local conditions, eg, temperature). If this cannot be achieved, individual IGCT devices can be overloaded. The cost and complexity of such snubber circuits is, in most cases, too high when compared to the IGBT alternative.
To conclude, IGCTs are best used in applications in which each switching function is carried out by a single device.

In the second part of this article on high-power semiconductors, which will be published in ABB Review 1/2007, aspects of the housing design are discussed. The article will also look at the potential of "wide-band-gap" materials.

Stefan Linder
ABB Switzerland Ltd, Semiconductors
Lenzburg, Switzerland
stefan.linder@ch.abb.com

References

[1] K. Eriksson: "HVDC Light™ and Development of Voltage Source Converters". Proc. IEEE/PES T&D Latin American Conf., Sao Paolo, 2002
[2] T. Laska, F. Pfirsch, F. Hirler, J. Niedermeyr, C. Schaeffer, T. Schmidt: "1200V-Trench-IGBT Study with Short Circuit SOA". Proc. ISPSD '98, 433-436, Kyoto, 1998
[3] M. Mori, Y. Uchino, J. Sakano, H. Kobayashi: "A Novel High-Conductivity IGBT (HiGT) with a Short Circuit Capability". Proc. ISPSD '98, 429-432, Kyoto, 1998
[4] Linder, Stefan: Power Semiconductors. EPFL Press / CRC Press, 2006. ISBN 2-940222-09-6 (EPFL Press) or 0-8247-2569-7 (CRC Press)
[5] S. Dewar et al: "Soft Punch Through (SPT) – Setting new Standards in 1200V IGBT". Proc. PCIM Nuremberg, 2000
[6] T. Laska et al: "The Field Stop IGBT (FS IG-BT) – A New Power Device Concept with a Great Improvement Potential". Proc. ISPSD 2000, 355–358, Toulouse, 2000
[7] D. Cottet et al: "Numerical Simulations for Electromagnetic Power Module Design". Proc. ISPSD '06, 209–212, Naples, 2006
[8] M. Rahimo, A. Kopta, S. Eicher: "Next Generation Planar IGBTs with SPT+".Power Electronics Europe, Ausgabe 06, 2005



Reflecting on color

Color control technology for papermakers

Anthony Byatt, Steve Sturm

For a papermaker, getting the color of paper right is an important element of quality assurance. Customers expect envelopes to match their writing paper, and expect the paper they buy today to match the envelopes they bought months or even years ago. Different grades of paper manufactured with different machine settings can be required to have identical colors. Look carefully at this issue of ABB Review for example; the cover is slightly thicker than the other pages, yet the color matches precisely.

Reproducing color is no simple task and would not be possible if it were not for advanced color-control algorithms and highly sensitive instruments. But it is not just sensitivity and reproducibility that these instruments must tolerate, they must perform under harsh conditions such as high moisture levels and thermal and mechanical shocks.

In papermaking, even an unwritten sheet has a tale to tell. In this and the following articles, the amazing story of color control in paper is unfolded.

Television sales outlets normally have a broad selection of different types of TV sets on display. The electronics inside a color TV represents the pinnacle of high-performance, low-cost technology usage in a consumer appliance. An observer looking at one such screen cannot fail to be amazed by the brilliance of the colors displayed. But taking one step back and looking at the whole row of TVs on display, a subtle, and sometimes not so subtle, difference in color between identical TVs becomes clear.

To add to the complication, color is not absolute – perception of color is influenced by an individual's biology (some are color-afflicted or color-blind) as well as other effects. Perceived color, then, is a product of a light source output, an object's reflectance, and an observer's optical sensitivity.

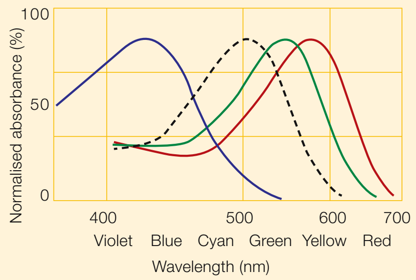
> To add to the complication, color is not absolute – perception of color is influenced by an individual's biology as well as other effects.

If achieving color consistency in this high-technology TV, built specifically for the purpose of true color reproduction is so difficult, then achieving it in a very humble, and apparently low-technology, product such as paper is very challenging indeed.

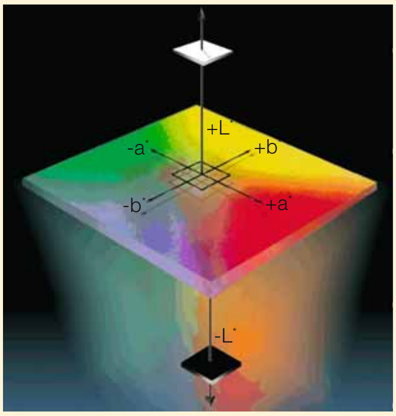
In paper manufacture, control of color – often actually whiteness – is critical. The continuous shift in the color of the dye in paper produced over a period of hours may be imperceptible to the human observer, but the difference in color can become very obvious when two sheets, one from the start and one from the end of the production run, are placed side-by-side. It would be unfortunate if these were to wind up as adjacent pages in a book. Paper appearance properties such as color, brightness, whiteness, opacity, and gloss have increasingly become the differentiating quality parameters in paper products.

1 Color perception

a Spectral absorbtion in the three types of cone cells of the human retina, and rod cells (dashed line)



b Representation of the L^* a^* b^* color space. Along the L^* axis, lightness varies from black to white



The retina of the human eye has three types of color-sensitive cells (known as cone cells). Each of these cell types is sensitive to a different range of wavelengths. These ranges peak at approximately 440, 544 and 580 nm (for blue, green and red light respectively). There is a considerable overlap **a** between the ranges, rendering the retina responsive to all frequencies between 400 and 700 nm.

An individual cone cell only reacts to the intensity of the stimulation. It can neither resolve the exact wavelength of the stimulating light, nor can it differentiate between monochromatic light (of a single wavelength) and polychromatic light (with a combination of different wavelengths). This represents a significant difference between color vision and hearing. In the latter case, humans can differentiate between wavelengths with considerable accuracy and the trained ear can even resolve the individual notes of a musical chord.

Despite this, humans can distinguish a broad range of colors. This is due to the cortex (the area of the brain concerned with vision) combining the signals from the three types of cone cells and interpreting every combination as a different hue.

In contrast to this perceived color system, the L^*, a^*, b^* color-space represents the real color of an object. L^* maps the lightness and a^* and b^* the magenta-green and yellow-blue variability respectively **b**.

(Figure 1a is taken from the Wikipedia encyclopedia and is subject to the GNU Free Documentation Licence.)

Manufacturers, then, go to great lengths to control the color of their product – and ABB provides them with the technology to do so, with advanced on-line color sensors and software to control the addition of colorants to the process.

Quantifying color appearance

Color is perceived via red, green and blue eye stimuli (tristimulus) **1**. The description of color is not intuitive and is therefore subject to a rigorous scientific CIE[1)] standard: L^*, a^*, b^*. There are many ways to quantify color appearance, but L^*, a^*, b^* is one of the more universal schemes.

The three parameters in the model represent the lightness of the color (L^*, L^*=0 yields black and L^*=100 indicates white), its position between magenta and green (a^*, negative values indicate green while positive values indicate magenta) and its position between yellow and blue (b^*, negative values indicate blue and positive values indicate yellow).

When light hits a surface, it can be reflected, absorbed or scattered. Smooth surfaces reflect and rough surfaces cause diffuse scattering. A surface that

Footnotes

1) CIE (International Commission on Illumination – the abbreviation actually stands for its French name, Commission Internationale de l'Eclairage) is a Vienna-based scientific organization that is broadly accepted as an international authority on light, illumination, color, and color spaces.

diffusely reflects all wavelengths equally is perceived as white, while a surface that absorbs all wavelengths equally is perceived as black. Besides this diffuse reflection, specular reflection can also occur (as in a mirror). A good mirror reflects all wavelengths equally, but is not perceived as white because of its smoothness. Similarly, a black object can reflect light if it has a smooth finish.

ABB metrology experts have succeeded in finding ways to deploy the measurement principles in environments too harsh for constant human habitation.

This, then, is the arena for the on-line color control produced by the ABB Quality Control Center of Excellence, part of the Paper business unit, based in Dundalk, Ireland. This unit specializes in the on-line measurement of paper properties and implementing complex control schemes to optimize these properties automatically. Color measurement is a popular on-line control, utilized on 40 pecent of all quality control systems sold.

2 The scanning platform

ABB's clients expect similar precision, reliability and ease of use from ABB's on-line instruments 2. These must operate without a technician for 24 hours a day, 365 days a year at temperatures of 60 °C and 100 percent humidity! Sometimes the instrument package vanishes into a cloud of water vapor as it traverses the sheet. Every few hours, the instrument spends some minutes off-sheet, in a suddenly 15 °C cooler environment, for test and calibration. In addition, shock of up to 4 gravities along any axis is common and vibration equivalent to 2 gravities between 5 to 500 Hz must be tolerated.

From the "any axis" shock point of view this is similar to dropping the laboratory technician onto a tiled floor from a height of about 390 mm onto his head – and still expecting quality measurements!

Where ABB excels, color measurement

Rising to these challenges, ABB metrology experts have succeeded in finding ways to deploy the measurement principles in environments too harsh for constant human habitation. The instruments must not only survive, they must operate continuously and provide data that rival laboratory instruments in accuracy and precision.

Modern instrument design

The color instruments used in quality control laboratories have become extremely sophisticated over the past 15 years. Nearly all high-quality instruments measure reflectance spectra of samples presented to the sensor. This generally requires a stable source of visible radiation and a complex optical system to gather reflected energy.

Laboratory color instruments are operated by technicians: They test the calibration, select the type of color coordinates required, align samples and collect colorimetric data. These technicians perform this task every hour or so in a benign laboratory environment with few vibrations or shocks to the instrument, or, indeed, the technicians.

Anthony Byatt
ABB Ltd.
Dundalk, Ireland
anthony.byatt@ie.abb.com

Steve Sturm
ABB Automation Technologies
Westerville, OH, USA
steve.sturm@us.abb.com

The Land effect

The deviousness of color is amply demonstrated by the so-called Land Effect. This effect was described in 1977 by Edwin Land (better known as inventor of the Polaroid camera).

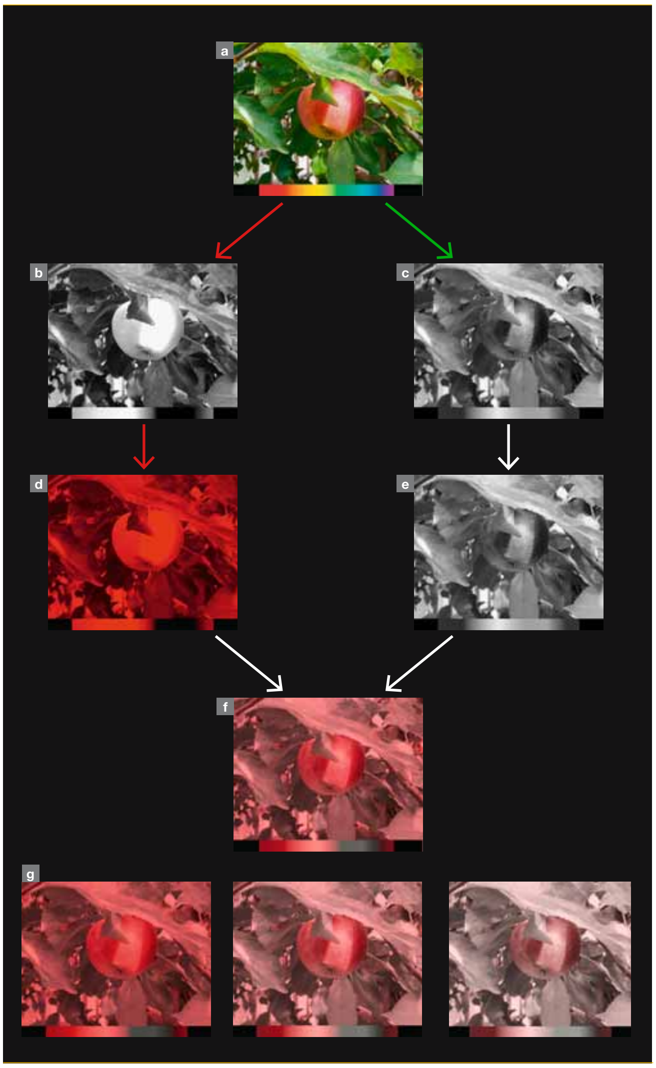
1 What colors does this picture contain?

Humans interpret objects as having constant color, regardless of the illuminant (daylight, artificial light, etc). For example, grass that appears green in intense sunlight retains its green color under a street light at night, despite the different intensity and spectral composition of these light sources. Land postulated that we perceive an object's color by comparing the trimstimulus input from its reflectance (see also figure one on page 41) with that of adjoining objects.

In photograph 1, several colors are visible. Or are they? In fact, this is a monochrome red picture with a black and white overlay. The only "real" color present is red. However the brain introduces other colors, including various shades of green and brown. (The picture is best viewed in artificial light which is slightly yellow.)

Land demonstrated[1] this effect by taking two photographs of the same scene 2a using black and white transparency film – one photograph was taken through a red filter 2b, and one through a green filter 2c. He then used two projectors to superimpose these pictures. He placed a red filter in front of the projector with the corresponding picture 2d, but left the picture that had been made with a green filter in black and white 2e. The resulting projection is shown in 2f. Whereas neither of the composing pictures has any green, this color appears to be contained in the resulting scene. The appearance of the colors can be fine-tuned by changing the relative intensities of the projectors 2g. The perception of these scenes varies with the light under which they are viewed.

2 The steps that went into the creation of 1



The Land Effect is an interesting illustration of how easily the brain is "fooled" into seeing colors, underlining the value of an absolute system of color measurement.

Footnote

1) Today, it is far simpler to recreate this experiment using the channel mixing functions of software such as Adobe Photoshop (as indeed was done here).

Color kitchen

Challenges of coloring in paper-making processes

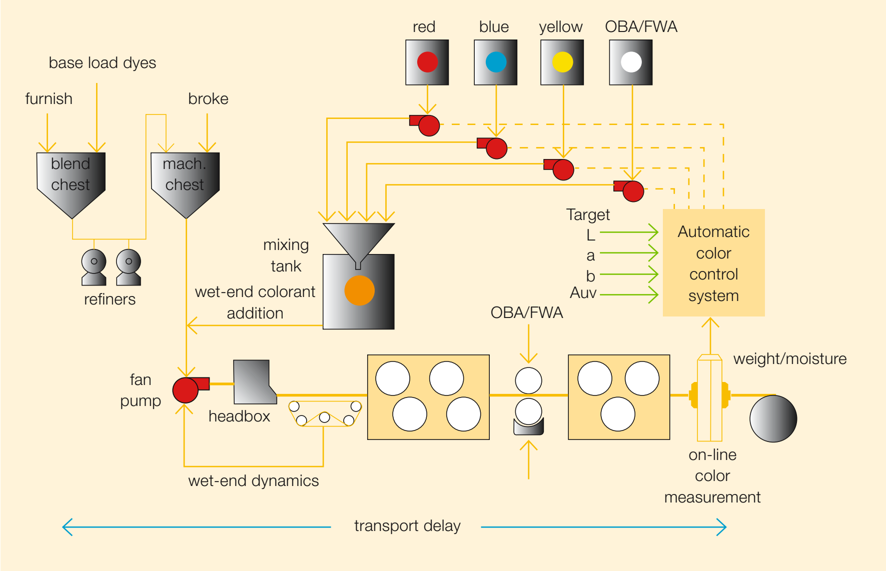
Anthony Byatt, Shih-Chin Chen

The color of paper is adjusted through the addition of various colorants during the production process. These colorants are dyes, fluorescence whitening agents and pigments. To permit the precise reproduction of color and to keep its varaibility within strictly defined bounds, a high-performance control system is required. Its implementation is made all the more technically challenging by to the long time delays and multiple process steps that can affect the color of the finished product.

Paper color is the result of chemical reaction between fibers and colorants; the reaction degree determining the color shade and depth. This, in turn depends on the base color of fiber furnish, fiber retention agent, the pH of carrying water and other chemistry in the wet-end circulation. Fluctuations in these conditions frequently give rise to unexpected disturbances to the uniformity of paper color shades.

A typical color process plant is shown in 1. The amounts of colorant per unit of fiber, also known as the dye rate between colorants and fiber, is the key contributor to paper color. Colorants can be added at different stages of the papermaking processes.

1 Coloring process of papermaking



2 Color kitchen

3 Well labeled, or color coded valves and connections are intuitive. They speed setup, and reduce mistakes.

To make a very deep shade of color, the majority of colorants are loaded into the blend chest in order to allow the colorants a longer time to bond with paper fibers. This operation is called base-loading and, typically, is not automated.

To adjust paper color, the colorants are usually added somewhere between the inlet to the fan pump and the coating addition on the sheet used to compensate for color disturbances. This is also known as color-trimming or color-dosing and is precisely metered to achieve the subtlety of color shade.

The automated control system needs to model the dynamic response of dosing colorant accurately and implement a complete multivariable feedforward and feedback control scheme.

Color kitchen

Central to the entire dying process is the color kitchen 2 3, where colorants are stored, prepared, metered, and delivered. The color kitchen is usually located so far upstream from the paper-forming process that the response of the colorant has very long time delays and slow dynamics, rendering manual control difficult. The challenge the operators face is further increased by the fact that color disturbances can originate in all parts of the paper machine and also that the chemistry of wet-end, size-press, and coating materials influence the coloring reaction at multiple stages. Compounding these challenges are the transient dynamics of grade change, speed change, and shade change, which induce additional shade variations

An automated color control system utilizing on-line color measurement has become necessary technology to the modern paper coloring processes.

Color control objectives

To achieve a uniform color shade, a precise delivery of colorants is needed. The key goals followed in the automation of a coloring process are the reduction of color variation from reel to reel for the same shade, the reduction of color breaks during shade changes, start-ups, and normal production and the reduction of dye-usage and the costs involved in matching a target shade during production (ie, to reduce metamerism Factbox and two-sidedness[1)] – without resorting to excessive trial and error).

In order to achieve these goals, the automated control system needs to model the dynamic response of dosing colorant accurately and implement a complete multivariable feedforward and feedback control scheme.

Factbox Metamerism

Metamerism is the phenomenon that two samples appear to be the same color under one light source, but different under other sources. Metamerism is caused by different dyes, levels of fluorescence whitening agents, fiber and filler types, etc, and is the second major cause for customer rejects.

There is almost always potential for metamerism between two production cycles. The better the variables that contribute to metamerism are controlled, the more consistent the paper products will be.

Anthony Byatt
ABB Ltd.
Dundalk, Ireland
anthony.byatt@ie.abb.com

Shih-Chin Chen
ABB Automation Technologies
Westerville, OH, USA
shih-chin.chen@us.abb.com

Footnotes

1) The term two-sidedness is used to refer to differences in the thickness or consistency of the coating on either side of the sheet

Shedding light on color

Sensing what customers want

Steve Sturm

In theory, the measurement of color is a question of shining a light with a known spectrum onto the paper and observing the spectrum of the reflected light. In a real production environment, however, the sensor is working under hostile conditions with ambient dirt and moisture and temperature variations, not to mention shocks and vibrations. The measurement apparatus must withstand these as well as possible while still delivering the most accurate results.

The complexity of perceiving color is reflected in the sensor that must measure it. A schematic of the ABB color sensor principles is shown in 1 and 2. The source of optical radiation is a Xenon lamp 2b. Approximately 5.6 Joules of energy are released via a 400 microsecond discharge. This produces a luminous flux equivalent to a 14 kW watt lamp radiating at a color temperature between 6500 and 7000 °K. The illuminator provides a short, high intensity pulse of mostly visible and ultraviolet radiation. This is collected in integrating cavities 2c. The exit aperture of the lower cavity is focused onto the sheet by the toroidal mirror 2e, providing a homogenous 12 mm diameter spot of illumination on the process (paper) 2i.

The accuracy and precision of these measurements rival modern laboratory instruments, but operate continuously and in very harsh environments.

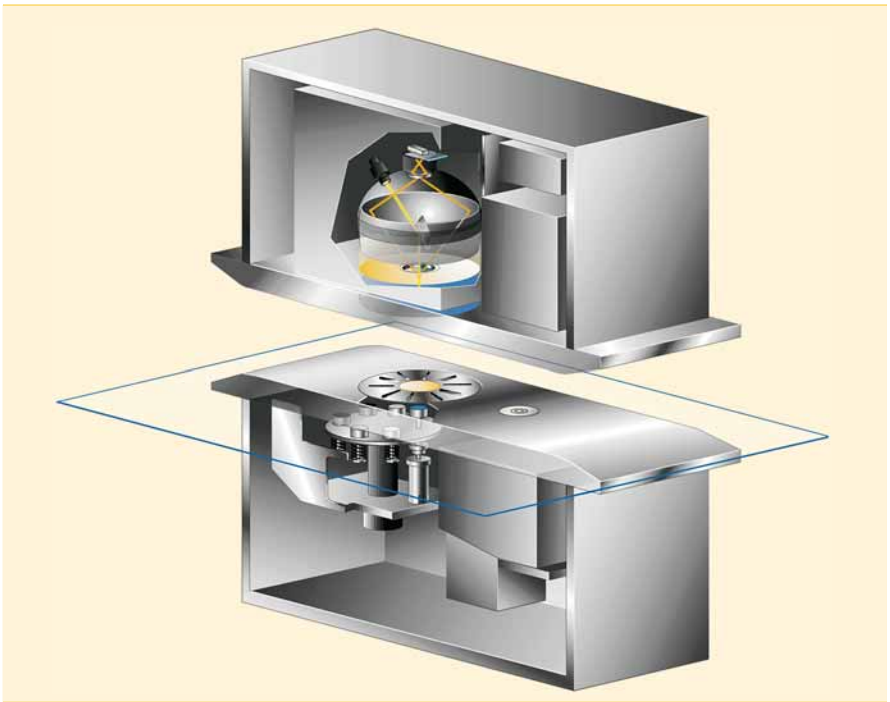
Radiation reflects from the process in all directions – the perpendicular part of this is detected. The detection optics consist of an angular mirror 2f, a collimating lens and a fiber-optic pick-up. The collected energy is delivered to an imaging spectrometer 2l.

The optical system incorporates a holographic diffraction grating 2n that spatially disperses detected radiation according to wavelength (from 340 nm to 780 nm) spreading it on a 256 element linear photodiode array 2o.

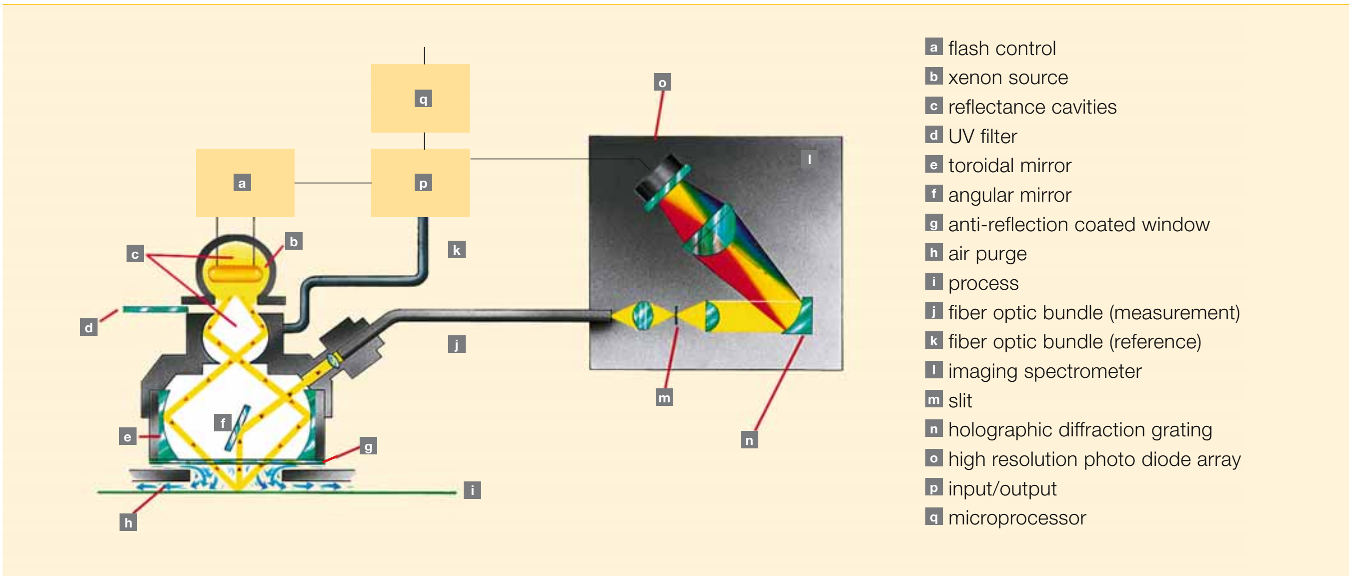
The reference radiation from the lamp is measured in two separate measurements with the help of two filtered detectors: One measures the lamp intensity in the red portion of the spectrum, the other in the blue. These signals normalize the detected signals for source intensity shift and color temperature shift from one flash to the next. This instrument is called an "abridged" dual beam spectrometer.

The voltages captured in the photodiode array 2o are serially shifted to an analog to digital converter where they are stored as an array of 256 digitized values. These are wavelength-corrected by evaluation of Xenon emission lines, whose center-wave-

1 A peek inside a color sensor

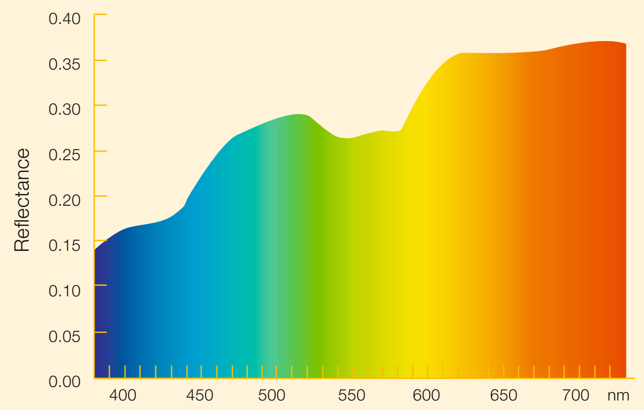


2 Principle of ABB's on-line color measurement



lengths are known. In this way, the values are corrected for the intensity and color temperature of the illuminator and normalized on the basis of previous measurements of a reflectance standard. The processed array accurately represents the reflectance spectrum of the process.

■3 Spectrum of the color of human complexion



Colorimetrics are computed on the basis of the measured spectrum. These are based on psychophysical measurements of the way a human brain interprets colors.
A representation of the reflectance spectrum of the color of human complexion is shown in ■3. This spectrum is very different from what the human brain interprets! (see also the factboxes on color perception on page 41 and the Land Effect on page 43).

The backing tile

The color sensor module is positioned on the top or bottom of the sheet (or both) and a so-called backing-tile module is positioned on the opposite side of the sheet ■4. This backing tile module maintains the sheet in the focal plane of the color module mechanically. The module also permits one of its tiles to be positioned behind the sheet during measurement, or in the focal plane when the instrument is off sheet for automatic calibration.

During the on-line measurement, a black tile and then a white tile are used as backing. Using the two resulting spectra, a third spectrum is computed for the sheet as if this were infinitely thick (a stack of sheets). This is necessary because the customer's QC laboratory characterizes their product by measuring a stack of sheets. This is usually the way that the end user "sees" the product.

The other tiles in the carousel are used to normalize the spectrometer for influences of dust on the sensor window (black hole), reflectance calibration (standardize tile), variation in the exact focal plane position (auto-focus tile) and reading a ceramic tile of known color (blue tile).

By algebraic manipulation of measured reflectance spectra, many different colorimetrics of the sheet can be derived. Some of these are shown in the Factbox .

In this manner, the ABB online color sensor derives colorimetrics of the process based on measurement of high-resolution reflectance spectra. The accuracy and precision of these measurements rival modern laboratory instruments, but operate continuously and in very harsh environments.

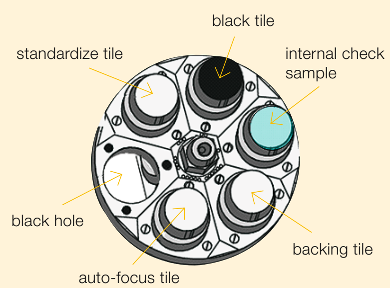
The instruments have been adapted to changing requirements: Deep shade measurement once represented 80 percent of ABB's color sensor business – today white shades make up this percentage. New customer standards have necessitated handling of compensation for fluorescent whitening agent.

But measurement, sophisticated as it may be, is only one part of the story. The use these measurements are put to is an equally important part of the paper machine's control story. This is explored in more detail in the next article.

Factbox Some of the colorimetrics that are computed

- CIELAB L*, a*, b* or L* u*, v*
- Hunter L, a, b
- DWL, Excitation purity and Y
- Tristimulus, X Y Z
- CIE Whiteness
- UV excluded brightness, TAPPI or ISO
- UV included brightness, TAPPI or ISO
- UV brightness (emission due to fluorescent whitening agents)
- Metameric mismatch with the target
- Color difference in dE*, CMC, FMC and others
- Opacity, TAPPI or PRINTERS

■4 Backing tile module, carousel and tiles



Steve Sturm
ABB Automation Technologies
Westerville, OH, USA
steve.sturm@us.abb.com

Color control in real time

Modeling colorant responses

Shih-Chin Chen, Anthony Byatt

In the previous articles in this section, ABB Review explored the challenges of measuring and correcting the color of paper. The description of coloring, of the added dyes and of their effect follows stringent scientific methods. In this article, ABB Review presents some of the mathematical notations involved and explores their implementation in production processes.

The dosing of colorants changes the reflectance of the final paper sheet. This can be expressed in the form of $[L^*, a^*, b^*]^T$ changes. The cause-effect relationship between color measurement $[L^*, a^*, b^*]^T$ and the dye rates u of all colorants are quantified as:

$$[L^*, a^*, b^*]^T = g(u) \tag{1}$$

where g is a nonlinear vector function determined by the selected color coordinate, the standard observer, the standard illuminant, and the colorant characteristics. For a small change in dye rate, δu, the color change $[\delta L^*, \delta a^*, \delta b^*]^T$ in the paper sheet is approximated by the perturbation model from (1) as:

$$[\delta L^*, \delta a^*, \delta b^*]^T = \boldsymbol{G}\delta u \tag{2}$$

For a typical paper machine, usually one to three different dyes are used to control paper color. Therefore if three dyes are used to control color, $\boldsymbol{G}$ is a 3×3 matrix:

$$\boldsymbol{G} = \boldsymbol{E}[\partial R/\partial\ (k/s)]\ [\partial\ (k/s)/\partial u] \tag{3}$$

where $\boldsymbol{E}$ is a matrix of the combined standard observer and illuminant of the coefficients. R is an array of the reflectance at every wavelength of the measured spectrum.

$[\partial R/\partial(k/s)]$ is the partial derivative of reflectance with respect to k/s

Sensing and controlling

evaluated at target color reflectance.

$[\partial(k/s)/\partial u]$ also known as the dye coefficient matrix, is the partial derivative of k/s with respect to dye rates evaluated at every wavelength of the measured spectrum.

k/s is defined as a light absorption to scattering ratio and is related to reflectance by the Kubelka-Munk formula:

$$k/s = (1-R)^2 \;/\; (2R) \qquad (4)$$

$[\partial(k/s)/\partial u]$ reflects how the colorants change sheet reflectance and can be derived from a set of standard sheets, known as dye samples, where each sample sheet is made with a precise dye rate. For each colorant, two to ten dye sample sheets are made with dye rates typically ranging from 0.01 percent up to 2 percent. Measuring these sheets yields a family of reflectance curves as indicated in 1a. The corresponding k/s curves are shown in 1b.

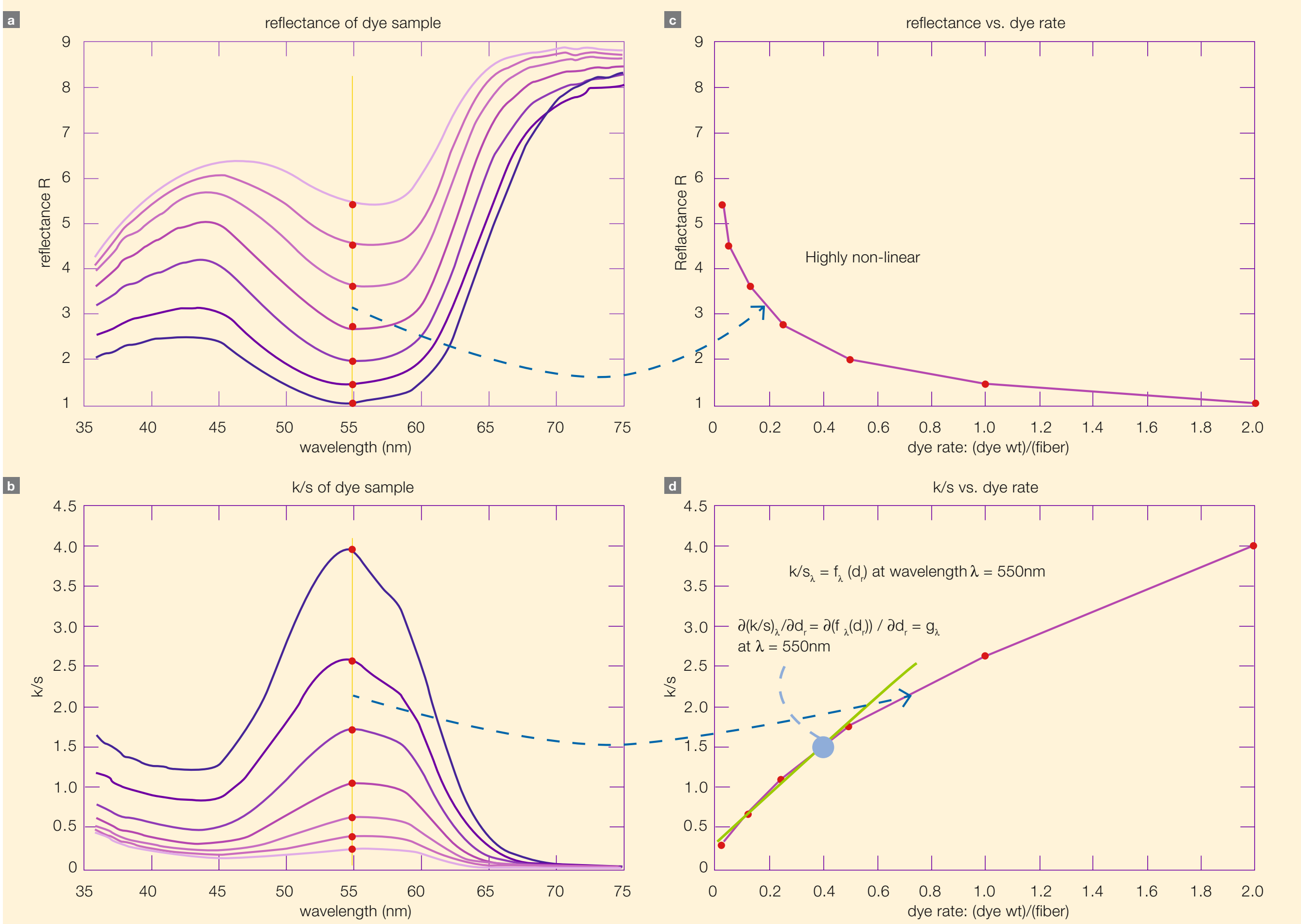
The reflectance and k/s as functions of dye rate at given wavelengths are shown in 1c and d, respectively. Noticeably, the relationship in 1d is closest to linear and can easily be approximated with a lower order polynomial. Additionally, the term $[\partial(k/s)/\partial u]$ from equation (3) at each wavelength can easily be evaluated from 1d. To automate a color control system, it is necessary to establish a complete library of dye coefficients $[\partial(k/s)/\partial u]$ and the $\boldsymbol{G}$ matrix can thus be calculated for any shade of paper from this library.

Dynamics of colorant delivering systems

In a continuous coloring process, the colorant additions have to be precisely measured and delivered at the correct stages of the process, usually by positive displacement, rotary, or peristaltic pumps.

Colorants are usually carried with a constant water flow through a mixing tank and piping before they enter the fiber flow. The volumes of the mixing tank and carrying water should be minimized to reduce response times and delays. Any pulsation of the colorant is to be avoided. Colorants delivered to different points in the coloring process may have different dynamic responses and time delays. The dynamics of each colorant addition to the process can be modeled as the combination of transport (or dead time) delay and a response with appropriate orders. The dynamics of a colorant delivery system are typically formulated as a diagonal matrix $\boldsymbol{D}(s)$ with transfer functions:

1 Dye characteristic calibration



$$\boldsymbol{D}(s) = diag(d_j(s)e^{\theta_j s}) \qquad (5)$$

where $d_j(s)$ is the first order dynamic and θ_j is the transport delay of the j^{th} colorant dosing point in the process.

Each element in the diagonal matrix represents the dynamics of each colorant dosing point in the coloring process. Since colorants are delivered to various locations in a sheet-forming process, the dynamics of colorants are usually represented by groups of transport delays and time constants.

Each colorant addition point may also exhibit different dynamic response behavior.

Multivariable feedback control

According to the Kubelka-Munk theory, if the colorants are mainly light absorbers, then the k/s of a paper sheet is the sum of the k/s of each individual colorant and the k/s of base fiber at every wavelength. This relationship is illustrated in the following equation:

$$(k/s)_{product} = (k/s)_{base\ fiber} + \sum_j (k/s)_{j\text{-}th\ colorant} \qquad (6)$$

Formula (6) allows the use of colorant response from formula (2) for implementing a multivariable feedback color control system. The color reflectance target is usually established by measuring a standard shade sample sheet with an on-line color sensor. The control objective is to reduce the deviation between the measurement and color target. This is defined as:

$$J = w_L\,(L^*\text{-}L^*_{target})^2 + w_a(a^*\text{-}a^*_{target})^2 + w_b(b^*\text{-}b^*_{target})^2 \qquad (7)$$

where w_L, w_a, and w_b are weighting factors to prioritize the importance of each color measurement at each coordinate. In some cases, the color deviations are not controllable under all conditions. For example, if the sheet color is already too dark and too green, it is impossible to simultaneously reduce both shade and lightness deviations by adding further dyes. In this situation, the trade-off between lightness and shade deviations can be made by adjusting the related weighting factors.

To implement a multivariable feedback color control system, the paper color must be measured continuously with on-line color sensors. The measured reflectances are converted to color values [L^*, a^*, b^*] and then compared with the color target. The required colorant adjustments are calculated from the model (formulae 2 and 5) of coloring process while minimizing the control objective (formula 7) and satisfying all control constraint conditions.

The transport distances between colorant addition points and the color sensor location normally appear as dead time delays for a color feedback control system. Each colorant addition point may also exhibit different dynamic response behavior. A multivariable model-based control scheme is implemented to accommodate the long dead time delays and the multiple dynamic responses.

Feedforward compensation

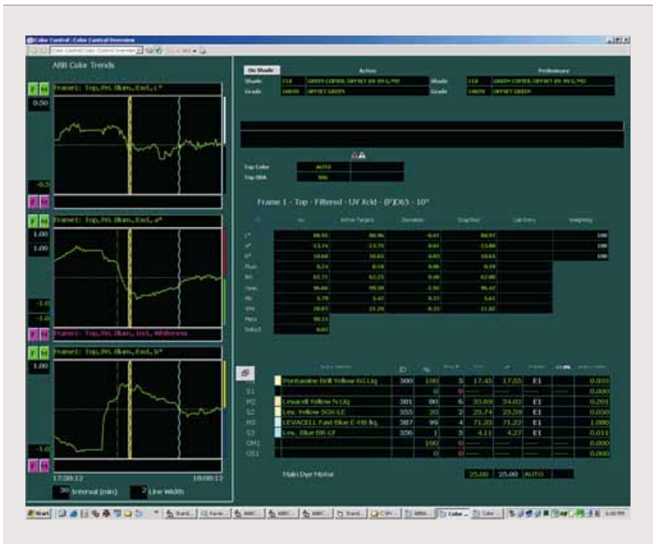
The throughput of the paper machine may change as the result of fluctuations in stock consistency, stock flow rate, and/or machine speed. The colorant flows can also change if the head size of pumps, the metering drive speeds, or the colorant consistencies change. In order to maintain the dye rates at the desired levels, all these changes must be compensated for with a feedforward control scheme. The feedforward control is integrated with the feedback control that adjusts the dye rates in order to reduce the color deviations. The feedforward for color control is calculated with the following formula:

$$\delta u_a = (\delta u\ P) \,/\, (H_p\ S_s\ C_c\ K) \qquad (8)$$

where

δu_a is the desired setpoint changes to dye metering pumps or valves

2 Color deviation correction after a shade change

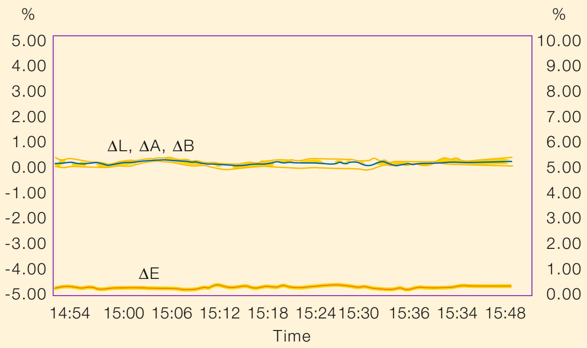


3 Stable color control over a long period of time

Sensing and controlling

P is the production throughput of paper machine
H_p is the head size of metering pump
S_s is the shaft speed that drives metering pumps
C_c is the concentration of colorant if the colorant is diluted
K is a conversion constant for matching up different units

4 Reproducible and reliable color control



Application aspects of color control system

There are several practical aspects that are crucial to the implementation of a closed-loop color control system for production applications. They are:

Handling shade target and dosing information

To enable a paper machine to achieve the required shades, appropriate shade targets are saved in a database. These targets are expressed in reflectance, tristimulus, color coordinates and dosing information for the colorants. The automated color control system retrieves these data before making a shade change. The shade database is crucial for shade repeatability. The characteristics of each colorant are also saved in a colorant library. Both databases are also used for deriving the coloring process models that the control system implements.

The average shade change time is reduced from over 40 minutes to less than 20 minutes.

Shade change coordination and boosting

Shade change transition is the most critical operation in every color production. Without coordination, paper makers can easily generate large numbers of color discontinuities. To coordinate a shade change, both the active and the upcoming (preliminary) shade information should be retrieved from a database and brought to a single user-friendly display for operators to examine and make adjustment before the execution of a shade change.

Man-machine interfaces and Visualization

According to survey responses from a broad range of paper mills, operators prefer to have all process information, such as color trends, plots of *a–b* coordinates, pump setpoint and status, alarm messages and operator entry points, all appear on one display. 2 and 3 show examples of this display. At first glance, this display may appear overcrowded. But after the operator gets used to it, this single display becomes the main interface for operator to manage all aspects of color production.

This sophisticated product has enabled the manufacturer to elevate production levels and quality significantly.

Application results

A color measurement and control system as described here has been implemented and installed in many paper machines. The performance of this control system can be highlighted in the following sample results:

2 illustrates a case where the color control was turned on after a shade change. The color deviations were quickly corrected by the control as shown in the color trends.

3 shows the color trends of a process with and without automatic color control. The long-term color stability and the small magnitude of color deviations indicate the significant improvements made by control.

Performance statistics

Based on the production statistics from machines controlled by the above color control system, the average ΔE is reduced to less than 0.5 for deep shade and 0.1 for white paper, or an average of up to 70 percent reduction. The average shade change time is reduced from over 40 minutes to less than 20 minutes, a reduction of more than 50 percent. The utilization of a color control system is typically more than 92 percent. These results can be further improved if both the process and the controller are fine-tuned.

Color satisfaction

The advanced color sensing and control described in this article are built on the experience of more than a quarter-century and have been supplied to leading papermakers around the world. In all cases, this sophisticated product has enabled the manufacturer to elevate production levels and quality significantly 4. In some cases it has led to month-on-month record production increases. The sales of, and customer satisfaction with, these products speak for themselves.

Shih-Chin Chen
ABB Automation Technologies
Westerville, OH, USA
shih-chin.chen@us.abb.com

Anthony Byatt
ABB Ltd.
Dundalk, Ireland
anthony.byatt@ie.abb.com

The cutting edge

Cutting the inefficiency out of paper re-trimming

Iiro Harjunkoski, Simo Säynevirta

It is a commonly appreciated fact that the success of complex operations often depends on being able to think ahead. As a result, the detailed planning of industrial activities absorbs considerable management focus. Although computer-based planning tools are still far from rivaling the analytical flexibility of the human brain, there are more and more tasks where computers perform as well if not better than humans.

In a paper-making plant, the jumbo-reel that comes out of the paper-making process must be cut into smaller rolls to suit the needs of customers. For a human operator to identify the best way to cut such a reel is both time consuming and offers no guarantee of an optimum in respect to avoiding waste or guaranteeing product quality. The problem gains a further degree of complexity when some areas of the paper sheet must be discarded for quality reasons. No two tasks are identical and the possibilities are endless. ABB provides a software package that identifies the best cutting strategy.

Sensing and controlling

In many industrial processes, there is still much potential for cost-cutting waiting to be tapped simply by implementing better planning strategies in regard to time and raw-material. In this article, an advanced optimization strategy is discussed, which combines offline planning with online quality augmentation. The solution considers quality profiles along the jumbo-reel and the requirements attached to each roll that must be cut from it. A complete geometric solution to this so-called "trim-loss" problem is achieved. The approach presented leads to optimal or close-to-optimal solutions achieving significant savings through tackling quality loss, ie, the economical loss based on degraded product quality. The resulting advantages include reduced energy and raw material consumption, improved reliability in matching customer demands and higher profitability through lower overall production costs.

Cutting paper

A typical paper machine produces a ten meter wide paper sheet at a speed of 120 km/h (or 33 m/s – ie, sufficient paper every second to make more than 5200 A4-sized pages). With a grammage of 80 g/m², this corresponds to 97 tons of paper every hour. The planning of such a process has a crucial impact on the result and its overall efficiency and profitability. ABB already offers an entire, fully-integrated production management suite for paper production including state-of-the-art tools such as leading quality control (QCS[1]) and web imaging systems (WIS). Furthermore, ABB's production planning software is often benchmarked as a top-of-the-line solution. It is therefore natural that ABB is seeking ways of making the existing solutions even better and more economical. This ensures that the paper needed is produced as efficiently as possible in terms of production costs, environmental burden and energy and material consumption.

A typical paper machine produces sufficient paper every second to make more than 5200 A4-sized pages.

The cutting of huge "jumbo-reels" into smaller ones takes place at a winder, immediately after the paper machine. The primary objective in cutting problems is to minimize trim loss, ie, the waste that occurs when the full width of the jumbo-reel cannot be used for producing the target rolls. If, for instance, 1.5 m wide rolls are to be cut from a ten meter wide jumbo reel, a 1.0 m wide strip (or 10 percent of the total produced) will go to waste. The problem of trying to combine different roll widths in order to minimize this waste is commonly called the "trim-loss" or "cutting stock" problem. In solving this problem, cutting patterns are identified that are then applied in a slitter by positioning the knives in the desired positions 1. The two most common problem objectives are:

- To find a cutting strategy that produces the required roll widths using as little material as possible, ie, minimizing the trim loss.
- To minimize the number of different patterns and sequence them in such a way that unnecessary knife-setup actions are avoided, maximizing production efficiency.

The approach taken in solving these sometimes contradictory objectives makes a huge difference to the solution that is finally achieved. As the number of discrete settings is vast, the optimization involves challenging mathematics. There are often millions of potential ways of arranging the target rolls on the jumbo-reel. In analyzing these, it soon becomes evident that there is no practical way of exhaustively testing them, even using the best and most efficient supercomputer. The problem size grows rapidly with the number of rolls due to the large number of alternative cutting strategies. Many heuristic/mathematical approaches to solving such a problem efficiently do exist however, although none of them is guaranteed to deliver the global optimum. Such approaches include rounding heuristics, column generation methods, solving problems only partially and other knapsack-type[2] algorithms, to name but a few. Looking at this collection of known approaches, it might seem that the problem is solvable. But would such an approach alone deliver adequate results for the paper-cutting problem?

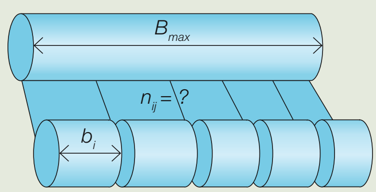
Considering paper quality

In modern paper mills, the trim-loss problem for the paper machine is commonly solved as an integral part of production planning, well ahead of the actual production of the jumbo reels.

This advanced planning would be quite adequate if it could be assumed that the paper has a uniform quality distribution, ie, to be of optimal quality throughout. Unfortunately, this is not always the case and local quality variations may occur. During the paper-making process, a large amount of data on quality – relating to numerous criteria – is collected from different on-line measurement and scanning devices. These data are then carefully processed and analyzed. Most of the resulting information is available shortly after production of the jumbo-reel has been completed and while it is waiting to be cut or trimmed. In 2, the deviation from target quality is represented by a color code.

When comparing the actual quality distribution with the planned cutting patterns of a jumbo-reel, the predetermined cutting plan may turn out to be far from optimal. For instance, the most valuable customer rolls may have been assigned to the worst positions on the jumbo-reel. If the paper were to be cut in this way, these sections would have to be rejected.

1 Jumbo-reel cutting and the trim-loss problem – what is the best way to cut small rolls of specified dimensions from the jumbo reel?



(B_{max} = jumbo-reel width, b_i = roll width, n_{ij} = number of rolls i in pattern j)

Previous modeling approaches for trim optimization did not support quality-based optimization, as the standard trim-loss problem did not take into account the exact position of each roll in a pattern. It merely focused on the total amounts being produced, ie, how many rolls of a given type each cutting pattern contained.

A novel mathematical programming approach

Quality information is collected by the quality control system (QCS) that performs continuous scans along the paper reel. Properties such as moisture, caliper and brightness are measured very frequently – in the cross-direction, typically every 10–50 mm and in the machine-direction every few hundred meters, depending on the speed of the paper machine as well as on the time that the scanning device requires to move across the paper width. In practice, this means that even a small paper machine has tens of thousands of measurement points for each quality criterium.

Another ABB technology is the web imaging system (WIS). In this system, cameras track visual defects (holes, cracks, wrinkles, etc) and the images are efficiently analyzed using neural-network based methods. These ensure fast and reliable processing in the classification and determination of defects.

Besides such online methods and the rapid and accurate information they provide, quality is also analyzed off-line in laboratories. This testing is much more time consuming and is therefore more suitable for tracking certain general quality trends than to observing short-term variations. This monitoring is based on selected samples and can lead to the rejection of a whole jumbo-reel.

Embedding all these quality aspects into a standard mathematical model for the trim-loss problem would increase their complexity further, soon rendering them intractable due to additional, non-implementable requirements and the large number of discrete decisions. An alternative modeling approach for a quality-based trim optimization is needed. ABB has developed a novel mathematical programming-based approach for automatic computation of an optimized solution of the trim-loss problem. The model assumes an existing cutting plan and is able to cover the various quality profiles of a jumbo-reel through a geometric representation when performing a re-trimming. Driven by tight performance requirements, the mathematical model itself is modular and in the first approach considers one cutting pattern (or trimset) at a time. A re-positioning of the rolls on the set takes into account the exact geometrical position as well as the quality information along the reel width. A second approach looks at the whole jumbo-reel and aims at re-sequencing the fixed cutting patterns in the best possible way. These two approaches can then be arbitrarily combined through an intelligent algorithm that is also able to include patterns from the whole production run. Splicing and reject-zones are also taken into consideration implicitly. The approach results in optimal or close-to-optimal solutions, reducing the quality loss, ie, the economical loss based on degraded quality significantly. This increases both the profitability of production, as well as the reliability towards customers – a more focused quality management also improves customer satisfaction.

During the paper-making process, a large amount of data – on quality relating to numerous criteria – is collected from different on-line measurement and scanning devices.

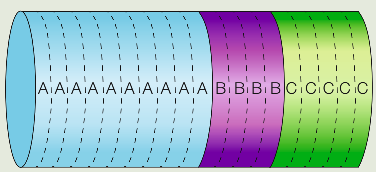
Two steps to success

In order to make this large and very complex problem solvable, the new mathematical approach contains two major steps. In the first step, the jumbo-reel is discretized by dividing it into "slices" 3a. The quality of each slice is projected to the customer requirements. The resulting optimization model builds on a quality classification, eg, A-, B-, or C-quality. The final

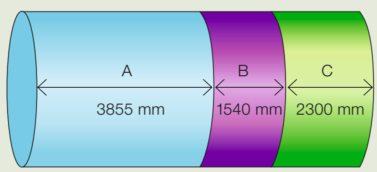
2 Quality analysis of a jumbo-reel (ABB CPM System)

3 Two approaches to finding the optimal way to cut the jumbo reel

a The discrete approach – the algorithm rasterizes the paper, the solution considered can be varied in discrete steps only.



b The continuous approach – dimensions are infinitely variable. This is much more difficult to optimize than the discrete approach of a.



quality for each roll is calculated by combining the quality mapping with key product-roll requirements. This discretization-based method delivers a good approximation to the optimal solution.

The solution is further improved in a second step, which applies a continuous and exact approach, ensuring the feasibility of the final solution. This also allows a smooth adjustment of the edge of each set. The continuous approach builds on dividing the jumbo-reel into continuous quality zones 3b – or sectors, according to the varying quality classifications of every roll. Each sector is associated with a respective quality (again: A, B, C). Similarly to the first step, this calculation combines the quality mapping with some key parameters of the product roll. This step results in an optimal cutting strategy taking into account the quality distribution along the jumbo-reel.

ABB has developed a novel mathematical programming-based approach for automatic computation of an optimized solution of the trim-loss problem.

Whereas neither of these two strategies alone can always handle more complex problems exactly and efficiently, this two-step approach is both robust and effective. It allows the non-convergent nature of the problem to be circumvented and ensures that a close-to-optimal solution is obtained quickly.

Well-hidden mathematics
The user does not have to deal with any of the underlying mathematics or algorithms. The functionality can be fully integrated into the existing environment; the solution will work silently in the background, creating additional benefits for customers. However, for those readers interested in taking a look "under the hood" some of the main features are explored here.

The approach comprises solving mixed integer linear programs (MILP) within specialized algorithms and utilizes robust and well-established technologies. The mathematical models carry certain similarities to production scheduling, as both feature essential logical decisions. In scheduling, the time-horizon is discretized by a fixed number of gridpoints, which are assigned to jobs through binary variables. In the retrim optimization context, the jumbo-reel width or length takes the place of the time variable.

The solution will work silently in the background, creating additional benefits for customers.

Focussing on the re-trimming of one set, it is assumed that a roll is represented by an index r and each discrete "slice" by j. Then, the binary (zero-one) variable, xd_{rj}, equals one at the jumbo-reel position, j, at which the roll, r, begins (left edge). In order to perform an optimization and maximize the total value of a trim set, a cost coefficient, c_{rj}, indicating the value of the roll at the given position is also required. The objective function is very straightforward: Define the location for each roll r that maximizes the total value of a trimset. This is expressed in Equation (1) below.

$$\max \sum_{r,j} c_{rj} \times xd_{rj} \qquad (1)$$

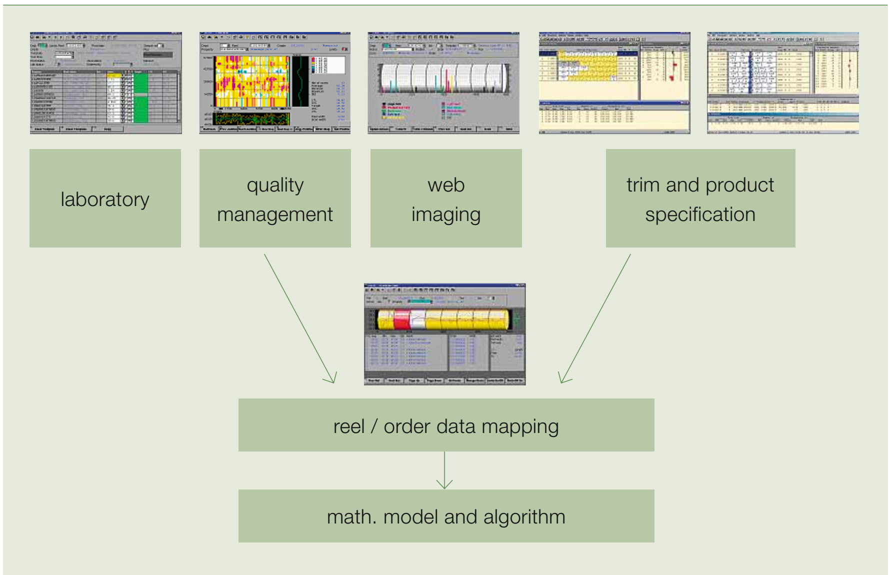
$$xd_{rj} \in \{0,1\}$$

In order to get everything defined correctly mathematically, equations ensuring that no rolls overlap and that each roll must occur exactly once are introduced. Such equations may sound trivial, but sometimes they can become more complex than expected.

The discretized problem results in an optimal cutting plan with respect to a chosen grid density. For typical jumbo-reel widths of up to 10,000 mm, an exact grid (1 mm) would make the problem size intractable. Therefore, a coarser grid (10–20 mm) is selected. In this case, the roll widths have to be rounded down in order to maintain the feasibility of the problem (eg, 578 mm becomes 570 mm when using a 10 mm grid). Such rounding errors are corrected on a consecutive continuous step.

The continuous step also resembles some scheduling strategies, in that it divides the jumbo-reel into flexible slots. These slots are ordered from left to right, and the borders between

4 Cross-application components – many different factors must be considered in finding an overall optimum



them are continuously variable, ie, able to adapt to the respective customer roll widths. Each roll is assigned to exactly one slot and exactly one quality sector through the use of binary variables. In equations (2)–(5) below, the slots are indicated by n and the sectors by s. Thus, a binary variable x_{rn} equals one only if the roll r is assigned to the slot n and similarly, variable x_{rs} indicates that roll r is located at the quality sector s. W_r indicates the roll width, r_r^{start} the position of the left edge of a roll, W_n^B and W_n^E the start and end positions of a slot and finally, S_s^B and S_s^E the start and end positions of the sectors.

$$W_n^E = W_n^B + \sum_r x_{rn} \cdot W_r \qquad (2)$$

$$r_r^{start} = W_n^B \text{ if } x_{rn} = 1 \qquad (3)$$

$$\sum_s x_{rs} = 1 \qquad (4)$$

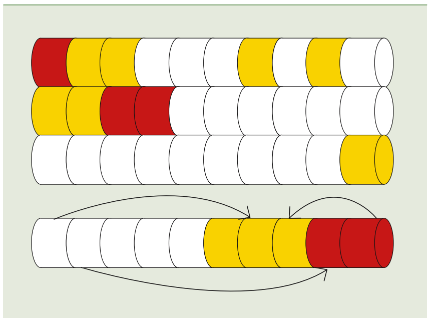
$$S_s^B \le r_r^{start} \le S_s^E \text{ if } x_{rs} = 1 \qquad (5)$$

$$x_{rn}, x_{rs} \in \{0,1\}$$

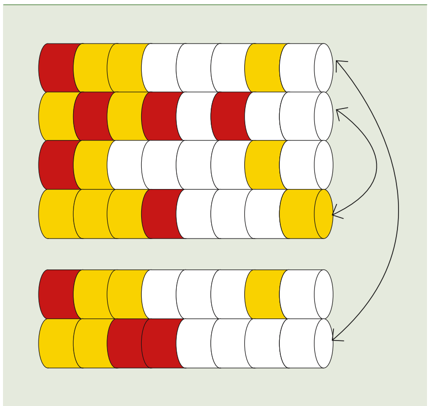
In short, Eq. (2) adjusts the width of a slot according to the roll that has been assigned to it. The slot should start exactly at the left edge of its roll, which is enforced in Eq. (3). The fact that one roll can only belong to one quality sector is expressed in Eq. (4) and finally the roll must be positioned within this sector, as indicated by Eq. (5). These are just some key constraints of the problem illustrating some of the main mathematical and logical dependencies.

5 Optimization appoaches

a Adjusting one set



b Changing the sequence of sets



An intelligent algorithm can reduce the quality losses to a physical minimum. This ensures that the current planning is always up-to-date with respect to the known quality data.

Algorithm for putting it all together

The components discussed in the previous sections are visualized in 4. More important than mastering the mathematical details of the model is understanding how everything can be put together into one robust and uniform concept. To illustrate this, the original problem is reconsidered. The mathematical modeling aspects presented above, combined with the two-step approach allows an efficient solution of the quality-based re-trimming problem. The resulting strategy can be applied in various ways: Re-organization can be executed one set at a time 5a, or by changing the sequence of sets in a jumbo-reel 5b.

In the former case (focus on one planned trimset), the set is adjusted so that the rolls are re-positioned in order to maximize the total value (quality yield), thus minimizing the effect of quality deviations. A simplified example of this would be to place the most valuable roll onto a good region. In 5a, the red color indicates that the roll must be rejected and the yellow color stands for minor quality deviations (B-quality).

The same technique can be also utilized to re-sequence the sets in a jumbo-reel 5b. Following the above principle, by combining the quality distribution information with the jumbo-reel cutting plan, it is possible to improve the quality yield. This is achieved by placing the sets into positions where the total value is maximal.

An intelligent algorithm that solves these two problems in a given sequence can reduce the quality losses to a physical minimum. This ensures that the current planning is always up-to-date with respect to the known quality data. Automatic jumbo-reel splicing and reject zones in cross-direction can also be implemented into the solution, as well as considering patterns from the whole run. To round it off, the cross-application between production planning and quality management offers additional possibilities for making the production both economically and environmentally more attractive.

Illustrative example

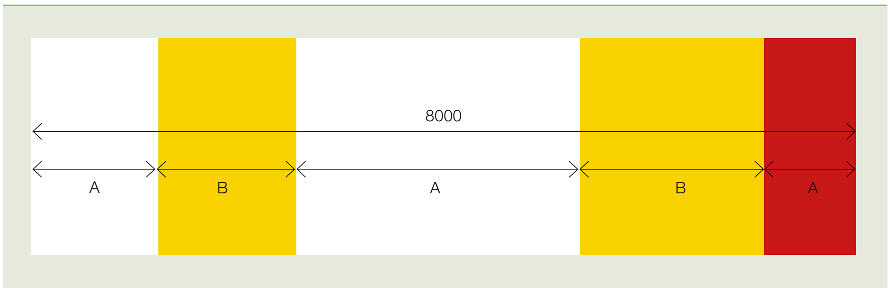
In the following section a simplified example is discussed. A trimset with the roll widths of the Factbox of page 58 is assumed.

The example jumbo-reel has a trim width of 8000 mm. The sum of the roll widths to be trimmed is 7915 mm, which results in a trim loss of 85 mm. Here, each roll is assumed to have exactly the same quality requirements. Therefore, the example can be simplified by directly dividing the jumbo-reel into various quality zones. If a roll spans over several quality zones, it is valued according to the worst quality. For illustrational purposes, a quality distribution is shown in 6a, in which the quality is expressed in terms of A-, B-, and C-quality. For the

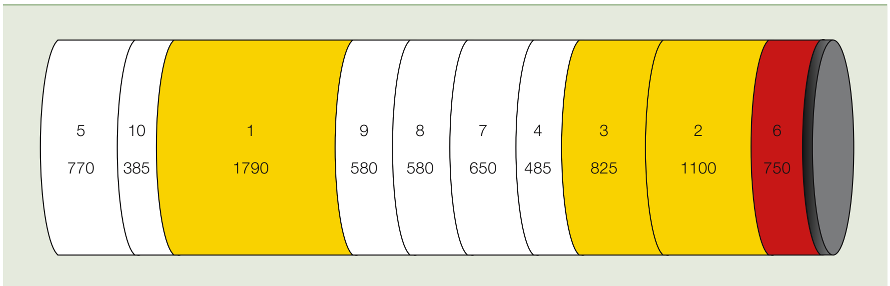
Paper rolls unloaded from the winder ready for packing

6 The reel quality map

a The problem



b The resulting trim set (quality: A=white, B=yellow, C=red)



Factbox Data for example above (roll widths)

Roll	1	2	3	4	5	6	7	8	9	10
Width (mm)	1790	1100	825	485	770	750	650	580	580	385

optimization problems reported here, the value of each roll is calculated based on the following assumptions: Set length = 5896 m, paper weight = 80 g/m², price = 500 €/ton, A-quality (full quality) = 100 percent of price, B-quality (minor defects) = 70 percent of price and C-quality (rejected) = 0 percent of price.

Using no optimization, but just ordering the rolls on the set as given in Factbox, would result in a total profit of 1236 €. The quality-based re-trimming algorithm finds the solution 1427 € 6b, which corresponds to an improvement of around 15 percent. The efficiency of the optimization can be tuned and there is always a tradeoff between solution quality and efficiency. However, the combined strategy gives a good result within a reasonable time.

Nomenclature

- xd_{rj} = 1, if roll r starts at gridpoint j
- c_{rj} = value of roll r at gridpoint j
- x_{rm} = 1, if roll r is assigned to slot n
- x_{rs} = 1, if roll r lies in sector s
- r_r^{start} = left edge of roll r
- W_r = width of roll r
- W_n^B = start (left edge) of slot n
- W_n^E = end (right edge) of slot n
- S_s^B = start (left edge) of quality sector s
- S_s^E = end (right edge) of quality sector s

The set is adjusted so that the rolls are re-positioned in order to maximize the total value (quality yield), thus minimizing the effect of quality deviations.

Cutting out waste

The solution discussed will not eliminate quality problems, but will minimize their effect by ensuring that planning is always geared towards the most profitable option and is able to make most out of the current quality. Furthermore, increasing the awareness of quality may also improve the overall planning culture and thus strengthen the capability to identify and analyze production efficiency with respect to quality.

The primary objective in cutting problems is to minimize the waste that occurs when the full width of the jumbo roll cannot be used for producing the target rolls.

Reduced quality losses mean:
- reduced production times
- less recycling of rejected rolls
- reduced energy and raw-material consumption
- better commitment to customer quality requirements
- more reliable delivery dates
- minimal environmental impact
- lower total production costs
- fewer quality complaints
- increased customer satisfaction

These sound very trivial. In fact, the proposed solution contributes to making these considerations part of everyday standard operation.

Iiro Harjunkoski
ABB Corporate Research
Ladenburg, Germany
iiro.harjunkoski@de.abb.com

Simo Säynevirta
ABB Process Industries
Helsinki, Finland
simo.saynevirta@fi.abb.com

Footnotes

[1] QCS: Quality Control System

[2] The knapsack problem is a combinatorial optimization task, the goal of which is the identification of the subset of a given set of objects so that the sum of their values is as close to a given limit as possible without exceeding it. The name derives from an example in which as many objects as possible are to be fitted into a knapsack of limited size.

Cementing profitability

Towards overall plant optimization in the minerals and cement industry

Dario Castagnoli, Michelle Kiener, Eduardo Gallestey

Rising competition and increasing resource costs are eroding the profit margins of many process industries. For these industries to survive, optimization is essential, but the constant streamlining of such processes brings its own problems for operators. Automation control tools can mitigate these problems and, thanks to today's advanced algorithms and high-power computers, solutions are at hand. ABB has developed Expert Optimizer v5, a computer-based system for controlling, stabilizing and optimizing industrial processes. The system has proved extremely effective and is currently in use in more than 300 reference sites, keeping industrial processes on target.

In response to the increasing demands for efficiency in industrial processes, ABB has built on recent advances in mathematical algorithms and high-power computing to create a comprehensive engineering tool, able to combine the needs of optimal production scheduling with those of classical advanced process control. To ensure the success of this tool in the marketplace and to provide ABB with a competitive advantage, the following characteristics have been incorporated:

- Full compatibility with all state of the art optimization techniques

- Sufficient flexibility to handle applications in different industries with different goals
- Ease of operation, to allow use by non-specialists
- Modularity, reusability and scalability to speed up development

Expert Optimizer v5 with Hybrid Systems Toolbox

Expert Optimizer is a computer-based system for controlling, stabilizing and optimizing industrial processes. It achieves its aims using the most modern control technology, ensuring that the best possible actions are applied accurately, tirelessly and consistently at all times. It has a successful track record in the marketplace, with more than 300 reference sites worldwide.

ABB's vision, to incorporate in one product all the functionalities required to satisfy the range of business needs for optimization solutions, led to the development of Expert Optimizer. The ABB unit, Minerals, worked in collaboration with ABB Corporate Research in Switzerland. In 2004 they began development of new model-based functionality, with a close eye on the market and the technological requirements mentioned above.

Mathematical techniques

The model-based environment of the new Expert Optimizer, known as the Hybrid Systems Toolbox (HST) 1, adopts the Mixed Logical Dynamical (MLD) modeling class [1]. MLD systems were recently developed at the Automatic Control Laboratory at the Swiss Federal Institute of Technology (ETH), Zürich, with whom ABB launched a strategic collaboration in 2000 [2]. MLD systems generalize a wide set of models, including hybrid systems, in which both continuous and discrete processes interact. A simple MLD system in Expert optimizer is shown in 2.

A good example of a hybrid system is an electronic thermostat, where the temperature, a physical quantity, is modeled as a continuous variable and the state of the heater switch (on/off) is modeled as a discrete variable. The ability to model hybrid systems significantly increases Expert Optimizer's range of applicability (hybrid systems are very common in the process industry). Further, unlike standard industrial models, MLD systems are able to model constraints, such as logic relations of the type: "If unit one is ON, then unit two must be OFF", or production constraints such as: "either NO production, or production between MIN and MAX". Last, but not least, MLD systems can handle piece-wise-linear mathematical models. This represents a good compromise between the high complexity of general nonlinear models and the ease of use of the linear theory.

> Expert Optimizer is a computer-based system for controlling, stabilizing and optimizing industrial processes.

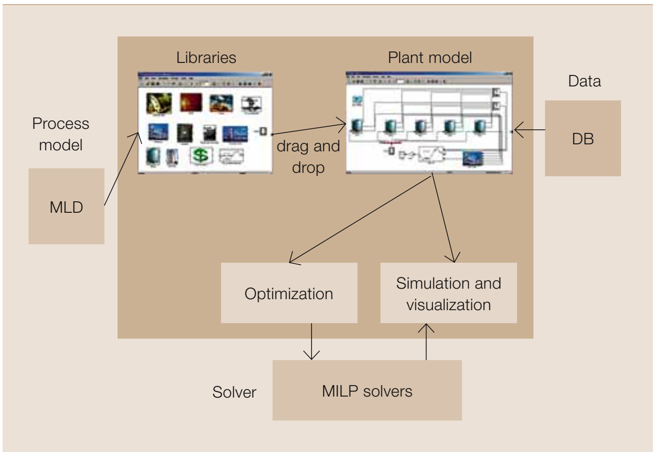
Among the advantages of the MLD framework is the existence of a standardized method to handle model predictive control (MPC), a well-known optimal control technique. The combination of MLD and MPC permits the solution of large-scale optimization problems where a cost-revenues index (known as cost function) is minimized over a given prediction horizon, while fulfilling the operating constraints. Depending on the requirements, the same framework can be applied either as an open loop, decision-making (scheduling) tool or as a closed loop, disturbance-rejection (re-scheduling) tool [3].

Modelling flexibility, acceptable computational times and optimal disturbance rejection are the key advantages of the approach.

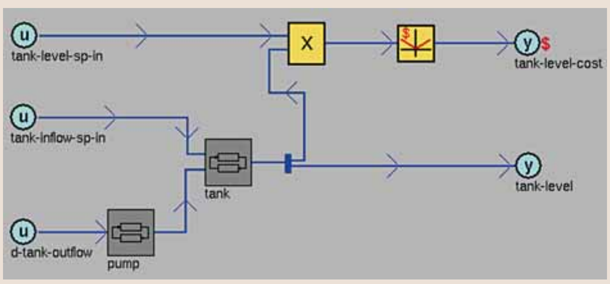
Implementation concept

The drawback of the MLD-MPC approach, as with other model-based techniques, is the relatively high complexity of the theory. This makes the modelling and maintenance of complex industrial systems difficult. The main driver for the development of the new Expert Optimizer was to make the MLD-MPC methodology accessible to non-specialist users, with the vision of making it a standard for model-based optimization projects within ABB and for its customers. Imagine that a certain process could be conceptually decomposed into smaller parts. For example, a hydroelectric power plant could be split into the reservoir, dam, turbine, generator, grid, etc. The idea is to model each part independently from the others in the MLD framework and to represent it graphically with a block. The block stores the model constraints and dynamics; its input and output ports match the inputs and

1 Hybrid Systems Toolbox (HST) outline



2 A simple mixed logic dynamical (MLD) system in Expert Optimizer



outputs of the MLD model. The complete process model is then obtained by graphically connecting the input/output ports of the various blocks. The next innovation is to represent the cost function that defines the optimal control problem as a second graphical MLD block. In the hydroelectric plant example, the goal could be to maximize the profits derived from the sale of electricity. The output of the generator block, representing the power produced, should then be linked to the cost function block, where the time varying tariffs of electricity could be stored. It is worth noting that each block, including the cost function, is a generic MLD system. The result is complete modelling flexibility.

Modularity simplifies the modelling phase and makes it easier to engineer, modify and maintain the models.

The modularity of the approach simplifies the modelling phase and makes it easier to engineer, modify and maintain the models. It also provides the opportunity to create libraries of standard blocks that can be reused in different processes, simply by dragging and dropping them from a library 3.

Once the user has created the plant model, the complex procedures of merging the MLD model blocks and formulating and solving the optimal control problem are automatically run in the background, unseen by the user. The user needs be concerned only with connecting library blocks or, in case of new applications, developing new basic blocks.

These features make the implementation of new robust applications faster than ever. What before might have taken months, can, with Expert Optimizer v5, now be achieved in a matter of a few weeks.

Continuous innovation

Scientists at the ABB Corporate Research Center in Switzerland have actively helped to develop and continue developing new functionalities related to advanced process control and scheduling in the MLD-MPC framework. The involvement of ABB Corporate Research and the contacts with the academic world guarantee that Expert Optimizer keeps pace with the latest innovations.

Typical new applications

Since the new model-based optimization capabilities became available, several Expert Optimizer projects have been installed and are running successfully in different domains. The most significant are in closed-loop process control: raw materials blending, mill and precalciner control; in production planning and scheduling: cement mill scheduling [4], titanium dioxide production scheduling and water distribution; in economic process optimization: alternative fuels management and thermal power plant optimization.

Raw Mix Proportioning

The Raw Mix Proportioning (RMP) module of Expert Optimizer uses new model-based capabilities to help cement manufacturers improve the quality of their raw mix and hence save money downstream. Blending raw mix components involves several non-trivial decisions that have a huge impact on later parts of the process. Deviations in raw mix quality are costly to correct later, with production losses and increased energy consumption being the immediate results.

The proportioning module applies MPC and MLD systems, which enable it to make predictions about the process and its chemistry. To achieve this, the model makes a prediction of the product quality after the mill and/or silos. Using information from the online and/or sampling analyzers, the model compares the prediction to the cost and quality optimization targets. A series of current and future optimal "moves" are then selected and the decision is implemented. Finally, the effects of process unpredictability and dynamics are absorbed and fed into the model to further improve predictions and decisions. See 4 for a screenshot from this model of Expert Optimizer.

The benefits of being able to make process decisions by looking forward instead of back are several:

- optimal trade-off between quality targets and material costs
- early smoothing of long- and medium-term disturbances
- meeting preconditions for maximized use of alternative fuels
- compensation for conveyor belt delays
- complete handling of process dynamics
- minimization of feeder moves

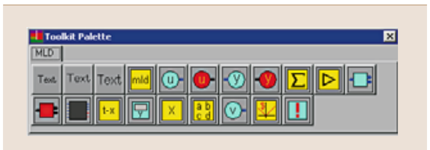
The RMP module is not a "black box" and the customer can change the priorities and weighting of various parameters without input or re-engineering from ABB.

Each block, including the cost function, is a generic MLD system. The result is complete modelling flexibility.

Grinding plant scheduling

The final grinding process of cement production, and the distribution of the various grades of cement to the different silos is dependent on careful scheduling of the mill. Using customer orders and energy-price forecasts, Expert Optimizer is able to help by executing periodic MPC iterations and its output is used as a reference schedule for operation of the mills 5. Here, the cost function represents costs associated with electricity consumption and the amount of low-grade cement produced (cement produced during the switch from one grade to another). Electricity cost reduction is achieved by committing the production to periods of low-tariff electricity, and by managing the mills so that the contracted thresholds of

3 Atomic Mixed Logic Dynamical (MLD) blocks library

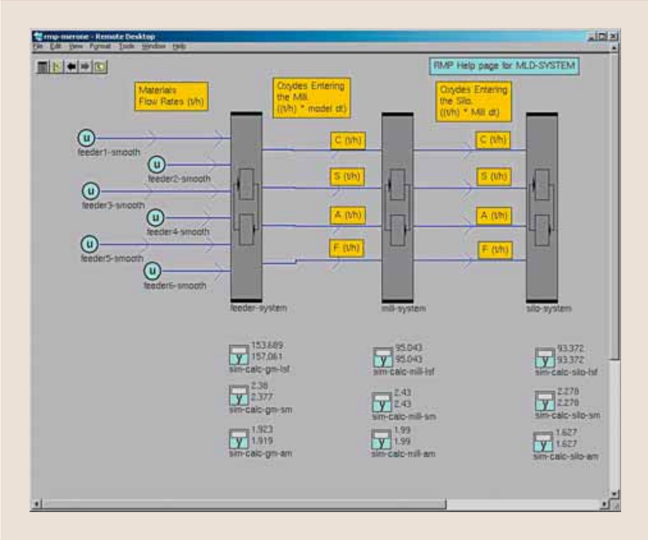


Sensing and controlling

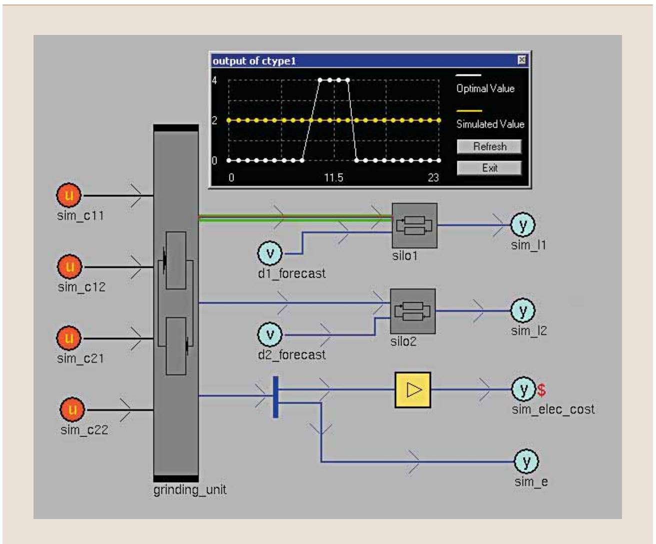
maximum electrical power are not exceeded. The production of low-grade cement is minimized by reducing the number of production switches.

Note that this "control problem" is more closely related to economic process optimization than to regulatory control: it tries to exploit degrees of freedom to increase the plant's financial performance.

4 Raw mix proportioning strategy



5 Grinding plant scheduling



The production of low-grade cement is minimized by reducing the number of production switches.

Conclusions

The results from this research and development work are excellent: twenty years of experience in the process industry with knowledge of established control techniques such as fuzzy logic, rule-based control and Neuro-Fuzzy are now combined with state of the art model-based optimization techniques. Complex real-world applications in the areas of closed-loop process control, open-loop decision support, as well as advanced production planning and scheduling, and economic optimization can now be tackled using a single product. Modeling, optimization and simulation capabilities are accessible within the Expert Optimizer graphic interface. The complexity is kept out of sight of the user, whose only task is process modeling. Flexibility is ensured by a modular structure and by providing libraries of reusable components.

Decisions about what tool functions are necessary to develop a solution are dependent on the amount of process knowledge and understanding available. By making use of the vast experience of its process engineers, ABB can be commissioned to develop the model and control strategy necessary to optimize the dynamic process, yet end users still have the option to capture their own process knowledge. The ease of use of Expert Optimizer v5, combined with comprehensive training and support from ABB, can help end users to develop and implement their own application more quickly.

The methodology developed will enable an owner to operate assets in an economically efficient and environmentally respectful way. With this tool, ABB has put the control engineer in a position to combine strategies of optimal control theory with econometric models of industrial assets.

The results described in this paper were obtained as part of the fruitful, long-term collaboration between the ABB Minerals and ABB Corporate Research in the program, Control & Optimization. This collaboration is an example of the benefits derived from close links between industry and research programs.

Dario Castagnoli
Michelle Kiener
Eduardo Gallestey
ABB Corporate Research,
Baden-Dättwil, Switzerland
dario.castagnoli@ch.abb.com
michelle.kiener@ch.abb.com
eduardo.gallestey@ch.abb.com

References

[1] Bemporad, A. and Morari, M. (1999) Control of Systems Integrating Logic, Dynamics, and Constraints, Automatica 35, no. 3, 407–427.
[2] Poncet, A. and Morari, M. (2005) Value for money, ABB Review 2/2005
[3] Gallestey, E., Stothert, A., Castagnoli, D., Ferrari-Trecate, G. and Morari, M. (2003) Using model predictive control and hybrid systems for optimal scheduling of industrial processes, Automatisierungstechnik 51(6), 285–293.
[4] Castagnoli, D. Gallestey, E. and Frei, C. (2003) Cement mills optimal (re)scheduling via MPC and MLD systems, Proceedings of conference on the Analysis and Design of Hybrid Systems (ADHS 03), 82–87.

Knowing your asset

Condition based maintenance at Outokumpu Steckel Mill

Magnus Tunklev, Per-Olov Gelin, Anders Bohlin

Unplanned downtime can be extremely costly for a production facility that runs 24 hours a day, seven days a week. It goes without saying then that proper maintenance has a very strong influence not only on production output but also on product quality. A clear strategy for maintenance to ensure a reliable process with high availability is therefore of the utmost importance.

The key to improved maintenance planning is knowing the "health" of each asset in real time, as well as the estimated time to failure. Getting this information in a complex system is not straightforward. Several systems that monitor the condition of a system as well as predict remaining lifetime do exist, but what has been missing in today's market is a complete solution package that can be applied to any type of asset and which is based on some kind of industrial software platform.

One such system now exits. It is the result of a close collaboration between ABB and Outokumpu Steckel Mill in Avesta, Sweden. The system provides the necessary information on asset wear as well as the residual lifetime for critical assets, such as bearings in the production line. It can predict the time-to-failure of an asset and significantly improve overall maintenance planning.

Sensing and controlling

For many years, optimization and fine-tuning in process control and automation was the norm. Because of this practice, very low margins now exist for improvement within these areas. Increased OEE[1)] has quickly become a key element for most industrial manufacturing units. As unplanned downtime is extremely costly for any production facility, the focus naturally turns to asset optimization and maintenance planning. A well functioning maintenance organization is vital when it comes to increasing Return on Investment (ROI).

The residual lifetime calculated from the accumulated wear value is only a rough guide as to how the asset is worn on a day-to-day basis.

Condition-based maintenance is an automatic process that strives to identify incipient faults before they become critical. This in turn leads to more accurate planning of preventative maintenance. Several conditioning monitoring systems are available. They differ in that some are based on statistics from historical failures and others on first principles modeling.

For Outokumpu[2)] Steckel Mill in Avesta 1, however, these solutions did not address its need for a complete packaged online condition monitoring system that could be applied to any asset in a production facility to provide information on its status, accumulated wear and residual lifetime. This system should also significantly improve maintenance planning to avoid costly unplanned downtime or unnecessary preventative maintenance. The company teamed-up to ABB and together they developed a system that not only meets all of the above requirements, but which is also fully generic for any industrial plant.

The ABB/Outokumpu solution is based largely on ABB's IndustrialIT Extended Automation System 800xA[3)], which allows an efficient, seamless integration of ABB proprietary solutions and non-ABB solutions like Computerized Maintenance Management Systems (CMMS). The 800xA control platform provides the basic functionality needed for the efficient development and integration of solutions for automatic condition monitoring as a complement to the traditional process control.

The following paragraphs describe this asset optimization and condition monitoring solution in greater detail.

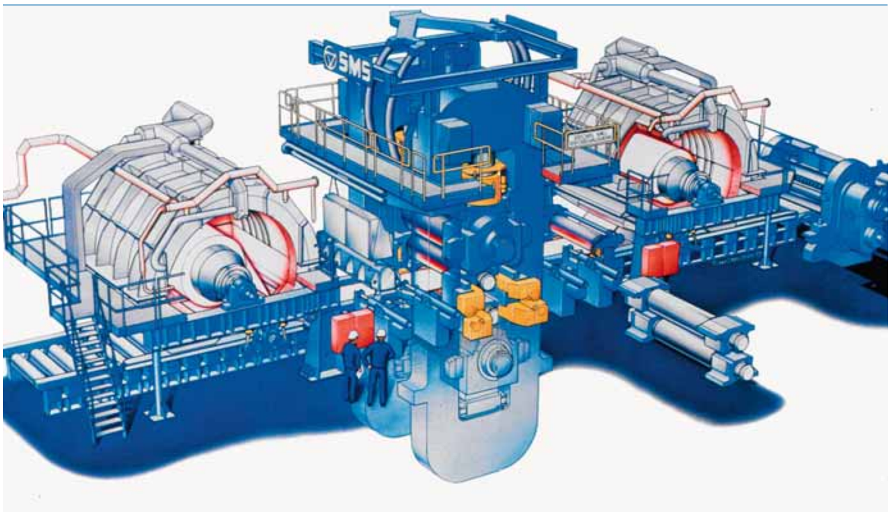
System overview

The overall solution consists of a suite of applications listed in Factbox. As well as including existing ABB software tools, new ones were developed to determine asset wear and lifetime. This solution suite allows: accumulated bearing wear calculation; bearing fault detection; residual lifetime estimation; anomaly detection from normal behavior; sensor diagnostics; and SMS and email messaging.

Condition-based maintenance is an automatic process that strives to identify incipient faults before they become critical. This in turn leads to more accurate planning of preventative maintenance.

A typical asset status screenshot (a sublevel view) is shown in 2. All diagnostics are based on green (status satisfactory), yellow (warning) or red (alarm) signals, and remaining asset lifetime is given in runtime hours. If a problem occurs, the data can be viewed and analyzed by venturing into various sublevels containing more detailed information. One noteworthy feature of this system deals with accumulated wear. This is particular to parts – such as pinch rolls – that are removed for one reason or another, and replaced at different positions. A storage identification (ID) function is manually activated from the wear aspect system by the operator for each asset 3. If an old asset is installed in a new position, the previous wear will be retrieved and the calculation of wear will continue from where it stopped.

1 The Outokumpu Steckel Mill in Avesta



Factbox Outokumpu Steckel Mill installed system components

- 800xA SV 3.1
- Asset Optimizer with Asset Monitors
- InformIT
- Wear Aspect System (New)
- Condition Severity Aspect System (New)
- DriveMonitor for bearing diagnostics
- Argus CC4 for data collection
- Argus OPC Server (New)
- PCA Model Builder Tool (New)

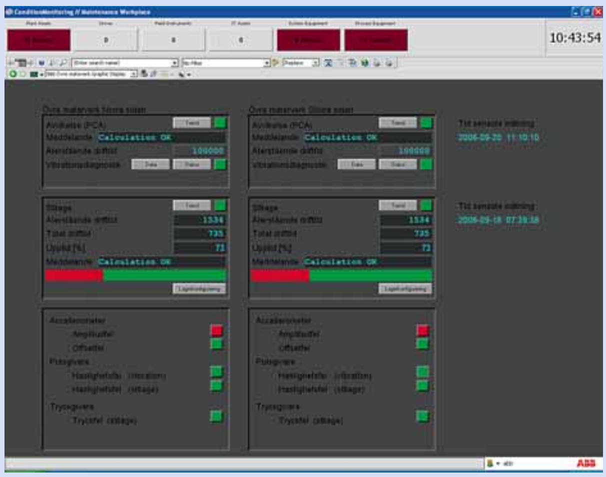
It is important to understand how the lifetime of an asset is determined. There are several different ways of doing this, but a crucial factor is the definition of failure. Failure is defined as the point at which the asset is normally replaced because of noise, vibrations or poor performance, but prior to mechanical breakdown.

Residual lifetime can be calculated if the rate of wear per run-time is known. When estimating the lifetime of a bearing with a specific load, the well established L10 theory [2] of bearing wear from SKF is used. The tricky part is to keep track of the variations in load and rotation speed and integrate the totally accumulated wear over time.

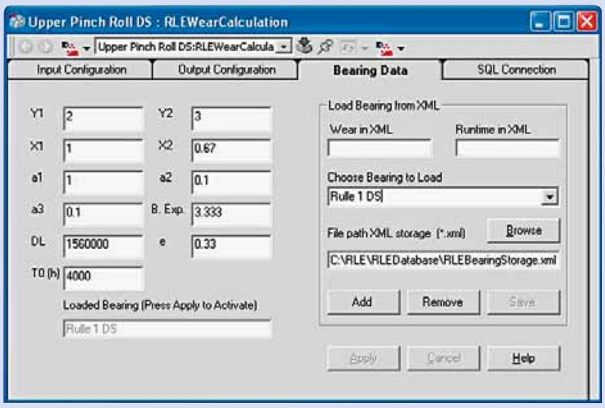
Firstly, load and rotational speed samples are continuously collected by a data logging system called Argus[4]. These data are then delivered as OPC values to ABB's System 800xA. The residual lifetime calculated from the accumulated wear value is only a rough guide as to how the asset is worn on a day-to-day basis. This is because the absolute value of the bearing wear may not be very accurate as it depends on environmental variables such as misalignment, bearing currents, cracks and looseness.

For practical reasons, the bearing wear calculation needs to be split into four intervals: prehistoric, old, new and prediction 4. The prehistoric interval is defined as the time before logging starts, and is particular to old bearings thar were in use long before the system was installed. If a new bearing is installed while the logging system is active, the prehistoric time is then zero. The old interval starts at the moment of logging right up to the second last batch, the new interval is the latest batch and finally, the prediction interval is used to estimate the future wear from moving averages of the current wear.

2 A sublevel view of the upper pinch roll with simulated bearing fault detection, anomaly detection, wear calculation and sensor monitoring



3 A screenshot showing bearing configuration and ID selection



Prehistoric wear

It is highly likely that many bearings will have been in operation for several months before the data logging system for wear calculations is implemented and this must be considered in the wear calculations. In the ABB/Outokumpu system, the time in operation prior to logging is represented by the parameter T0 in 3. Future and prehistoric wear is then estimated using averages from the old interval. The reason for this is that wear can vary significantly from batch to batch and it needs to converge towards a reasonable average rate. The averages used are: run-time per total time, wear per revolution, and revolution per runtime. As the data are collected in the old interval, the averages are continuously updated online and will converge after a few weeks[5]. The total accumulated wear is calculated as the sum of all the different intervals. The residual runtime and the residual total time can now be calculated using the converging averages from the old interval.

Anomaly detection

Sensors are the eyes and ears of process control and modern manufacturing plants are alive with them, observing every aspect of plant activity. However, faults do occur that are very rare and problems arise when these new or extremely rare faults are not spotted in time and to sufficient accuracy. This has to be considered in any new system design.

Footnotes

1) OEE – or Overall Equipment Effectivesness – is the industry accepted tool to measure and monitor production performance.

2) Outokumpu is an international stainless steel and technology company. Its headquarters are in Espoo, Finland.

3) System 800xA is an automation platform that integrates the core automation system in a plant (the process control system) with all the other applications essential to plant productivity and efficiency – such as engineering, documentation, quality control,safety, smart instrumentation, asset optimization and maintenance management. For more information, see http://www.abb.com. Select Product Guide. Under Control Systems select 800xA

4) Argus was developed by ABB Service.

5) A short stop at the beginning of production will significantly lower the runtime per total time average and effect the calculations. After a few weeks, however, this effect will not be seen.

Sensing and controlling

The preferred method of detecting deviations is to train the system using normal data. The chosen method for the modeling of normal behavior is a linear variable reduction method called Principal Component Analysis, (PCA)[6]. The PCA model tool developed within the project is generic and can be applied to any kind of process data. For the Outokumpu mill it has been applied to vibration data. As a fault evolves over time, the deviation from normal behavior – in PCA space – is seen in the residual Q of the new data projected onto the PCA model. As the Q value increases, the rate of change can be used to predict the time before a preset alarm limit is reached.

This system should also significantly improve maintenance planning to avoid costly unplanned downtime or unnecessary preventative maintenance.

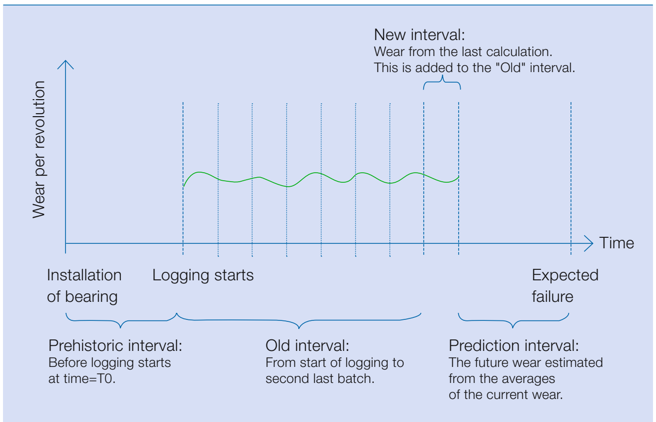
If the PCA model is applied to a new bearing with no defects, the residual lifetime will be determined as infinite. Therefore the residual lifetime calculated using the accumulated wear value is taken as the guideline. The PCA model will determine a realistic value only when a fault has been detected. When this does happen, the PCA residual lifetime value is considered more reliable because the PCA model is deemed a better estimator of fault evolution.

Asset state detection

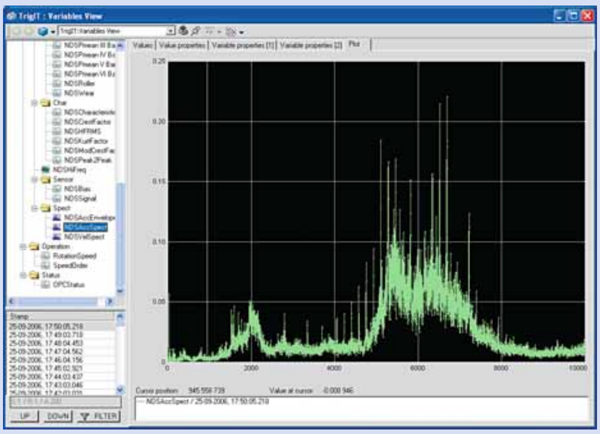
Each bearing must be configured, and the specific bearing fault frequencies are calculated using material from the manufacturer. The current and actual state of the bearings (healthy or broken) is detected by a diagnostic module called DriveMonitor[1]. The DriveMonitor solution detects outer race, inner race and roller faults online. Fault detection algorithms for other assets can also be configured using this tool. For example, to detect sensor fault in the accelerometers 5, the bias and standard deviation of the signal are calculated. Depending on the result, an alarm is issued if a preset threshold level is exceeded. Asset Monitors are applied directly to some signals outside the scope of the DriveMonitor tool (such as load and speed sensors signals used for wear calculation) to ensure that alarms are triggered if threshold levels are exceeded.

Finished roll of metal sheet.

4 The different wear intervals



5 DriveMonitor displaying an Acceleration FFT spectrum from the exhaust fan with a likely outer race bearing fault in the blade side bearing



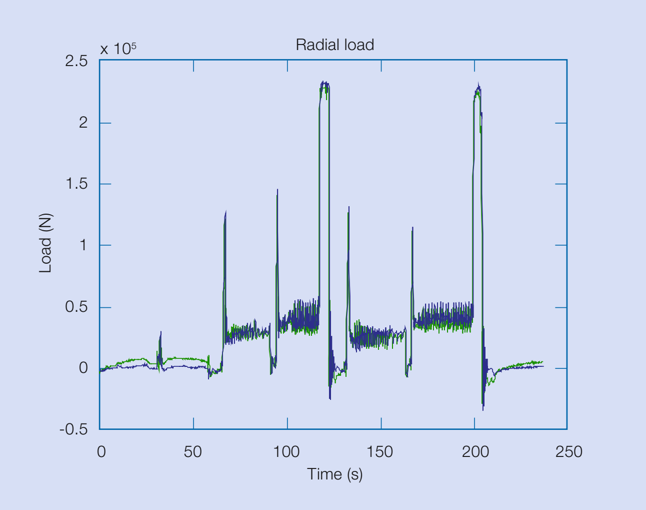
Faults do occur that are very rare and problems arise when these new or extremely rare faults are not spotted in time and to sufficient accuracy.

Experimental results

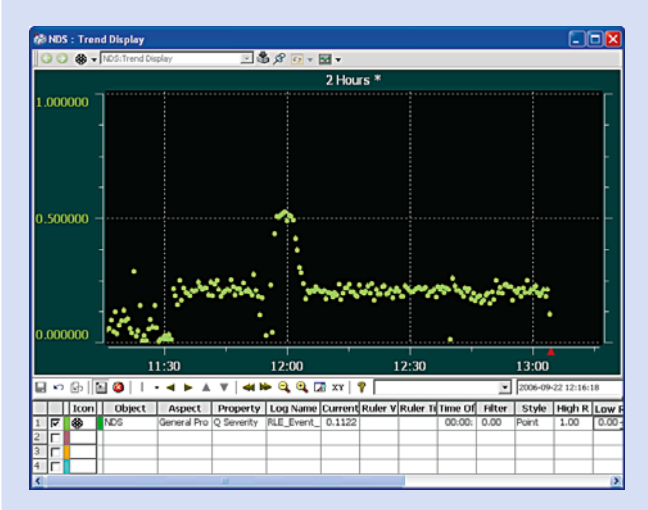
The asset optimization and condition monitoring system has been installed on the following four rolling mill assets: upper pinch roll; lower pinch roll; three-roller table; and the exhaust fan[7]. The following paragraph looks at data acquired (using the Argus PC) from the upper pinch roller bearings.

The process duration for each slab of material is normally of the order five to seven min-

6 Radial load on the upper pinch roll bearings for driving side (—— DS) and non-driving side (—— NDS).



7 A Q trend – using the PCA model tool – on the vibration data from the fan.



utes. For the slabs in this example, the bearing loads, as shown in 6, and rotational speeds were measured by the Argus unit and stored in a file. The file data were then converted to an OPC array by the Argus OPC Server. The sharp load spikes seen in 6 occur because each slab end hits the pinch roll when the side that is hooked has been released from the coiler. Engineers calculated that if these spikes were halved to 1.2×10^5N, the bearing lifetime could be extended by a factor of five! The accumulated wear calculation shows the difference in wear for different slabs. This is most likely due to the slab thickness, the time it takes to roll it, and the specific material used. The more detailed influence of different variables remains to be studied.

From the results obtained so far, it would be reasonable to extend the time between maintenance stops and have more up-time for the production. The algorithm was tested on the upper pinch roll driving side (DS). The wear per batch is significantly different, and can vary by as much as a factor of five. This linearly affects the remaining runtime. 7 shows the Q trend, using the PCA model tool, on the vibration data from the fan.

This asset optimization and conditioning monitoring system is fully generic for any industrial plant.

Everyone's a winner

Asset optimization and condition monitoring solutions are providing ABB with new and exciting business opportunities. However if the system is marketed and sold only within the rolling mill segment, the development payback time for ABB is estimated at six years. Anders Bohlin, senior project manager at Outokumpu, reckons that if the system performs to expectations the pay back time for Outokumpu could be very short.

As was previously stated, this system is extraordinary in that it is fully generic for any industrial plant, not just rolling mills. This means the true business potential will be significantly greater if the system is marketed within many other industries such as pulp and paper, petrochemicals, mining, cement, food and beverage, and pharmaceuticals.

Magnus Tunklev
ABB Corporate Research
Västerås, Sweden
magnus.tunklev@se.abb.com

Per-Olov Gelin
ABB Industry Solutions
Västerås, Sweden
per-olov.gelin@se.abb.com

Anders Bohlin
Outokumpu Stainless
Avesta Works Steckel Mill
Avesta, Sweden
anders.bohlin@outokumpu.com

Footnotes

[6] Principal Components Analysis (PCA) is a powerful way of identifying patterns in data and expressing the data in such a way as to highlight their similarities and differences. Once a pattern is found, the data can be compressed without much loss of information.

[7] With driving side and non-driving side on all.

References

[1] Wnek, Maciej; Orkisz, Michal; Nowak, Jaroslaw; Legnani, Stefano; DriveMonitor: Embedded product intelligence that enhances lifecycle management and performance in drive systems; http://www.abb.com

[2] SKF product documentation "SKF spherical roller bearings – setting a new standard for performance and reliability"

Microsystems at work

A fast oxygen sensor for continuous gas analysis

Peter Krippner, Berthold Andres, Paul Szasz, Thomas Bauer, Manfred Wetzko

Drawing on 40 years of experience in the design and development of continuous gas analyzers, ABB has devised a more effective, harder-wearing, lower-cost solution from scratch. By combining new materials with cutting-edge micro-electromechanical system technology, the company has come up with a radically new sensor system that competes well with its rivals in terms of performance, durability and cost.

These lower costs, gleaned largely from the ability to mass-produce the new instrument, along with a faster response time and an ability to cope with variable gas flow rates, will significantly influence the future of oxygen concentration measurement.

ABB's continuous gas analyzers, the Advance Optima 1 and EasyLine series, use high performance technology to measure the concentration of gases in various aspects of processing industries. Such measurements are vital to environmental protection, process optimization, quality assurance and cost reduction. The analyzers are used for continuous emission-monitoring of power plants and waste incinerators 2, and, among other applications, to guarantee the purity of gases in air separation plants, to measure flammable gases in hazardous locations, to protect electrostatic dust filters, and to optimize combustion processes in harsh environments such as cement plants. Both the Advance Optima and EasyLine gas analyzers are based on modern modular analyzer technology that uses state-of-the-art electronics, intelligent algorithms and improved continuous measuring technology.

Advance Optima and EasyLine gas analyzers are based on modern modular analyzer technology that uses state-of-the-art electronics, intelligent algorithms and improved continuous measuring technology.

A gas that is of interest to nearly all industrial processes is oxygen. Its concentration is usually measured using electrochemical cells or paramagnetic sensors. Electrochemical cells offer a significant cost advantage over paramagnetic sensors, but have longer response times (in excess of 20 seconds) and reduced lifetimes when exposed to dry gases. Their high cross-sensitivity (sensitivity to gases other than oxygen) and low compatibility with corrosive gases (eg, sulfur, chlorine and fluorine compounds) generally limit their use to applications in which the components of the gas mixtures are known. Paramagnetic sensors, despite their cost, offer a number of advantages over electrochemical cells. In these sensors, the interaction of the gas with the sensor is physical rather than chemical, so the poisoning that is typically observed in electrochemical sensors does not occur. In non-corrosive atmospheres, paramagnetic sensors have, in principle, an endless lifetime. And if the correct materials are chosen, they can achieve extended life times, even in dedicated corrosive atmospheres. In addition, the cross sensitivity of paramagnetic sensors is negligible. Overcoming the present production drawbacks and further enhancing the performance of paramagnetic oxygen sensors will open up new market segments and allow production volume to be increased.

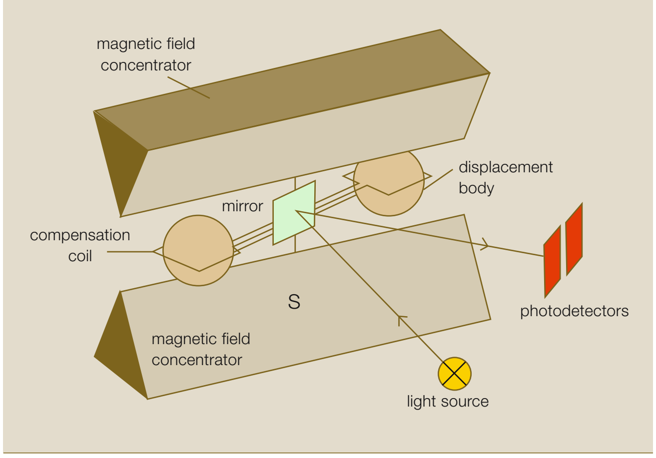
Paramagnetic oxygen measurement, the basics

Paramagnetic oxygen sensors are used to evaluate the magnetic properties of

2 Typical application area for oxygen sensors: emission monitoring in power plants

1 The Advance Optima range of continuous gas analyzers

3 Arrangement of the displacement body in a paramagnetic oxygen sensor



gas mixtures. Unlike most other gases, oxygen molecules exhibit a relatively strong, positive susceptibility and are therefore paramagnetic. Since most other gases exhibit a low diamagnetic susceptibility, the magnetic susceptibility of a gas mixture depends strongly on its oxygen content. This property is exploited by so-called magneto-mechanical oxygen sensors, which measure the magnetic susceptibility of a mixture and thereby determine the portion of oxygen present. Measurements are taken using a displacement body located in a strong magnetic field gradient and a torque measurement setup 3. Briefly, the torque generated on the probe body in the presence of oxygen is compensated by a small current in a coil that is fixed to the same probe body. To keep this compensation in balance, an optical detection unit measures the position of the probe body. The sensor driver electronics chooses the compensation current such that the position of the probe body is kept constant at all times. The current needed to compensate for the torque generated by the oxygen is a direct measure of the oxygen concentration in the sensor cell.

Conventional paramagnetic oxygen sensors

The geometry of the sensor elements used in conventional paramagnetic oxygen sensors is as shown in 4. The sensor itself consists of classically machined parts that are assembled manually. The core of the sensor, with a dumbbell-shaped probe body (with a diameter of approximately 2 mm and a length of 20 mm), is depicted in 5. To manufacture this element with sufficient reproducibility, highly skilled staff are needed. The construction of the small glass bulbs and their installation is time-consuming and the materials used for the mounting brackets are not compatible with highly corrosive gases. The performance of today's sensors is, however, excellent for many applications, especially for emission monitoring in fossil-fired power plants. Based on this paramagnetic principle, ABB offers a module for the measurement of oxygen concentration that can be used with both the Advance Optima and EasyLine series of continuous gas analyzers.

The main characteristics of the sensors are:

- minimal measuring range of 0 .. 1 vol% O_2
- detection limit of 50 ppm O_2
- response time of 3 seconds

The cross sensitivities to other gases are listed in the Factbox. The table shows the zero-point shift of the output signal when undiluted gases are present. The values indicate that this zero-point shift is caused by the magnetic properties of the gases and are the unavoidable remaining – and therefore accepted – zero point errors for paramagnetic oxygen sensors.

Performance drawbacks of the classical paramagnetic sensors include an insufficient compatibility with certain media (eg, chlorine or inorganic acids such as hydrochloric acid), and the slow response time caused by the large inner sensor volume.

The evaluation of magnetic properties of gas mixtures is used in paramagnetic oxygen sensors.

Complete change: materials, geometry and manufacturing

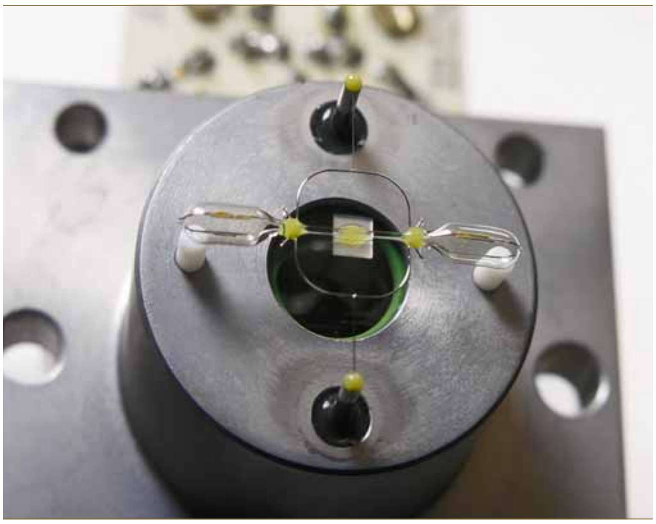
ABB has been pushing conventional paramagnetic sensors to the limit in terms of cost and performance for more than 40 years. To achieve yet more improvements, the company has gone back to the drawing board, initi-

Factbox Zero-point shift (in vol% 0_2) of the new sensor's output signal in the presence of undiluted (100 vol%) gases

Argon	-0.26
Carbon monoxide	-0.01
Hydrogen	+0.28
Hydrogen sulfide	-0.45

4 A conventional paramagnetic oxygen sensor: parts are assembled manually

5 The core of the conventional sensor: highly trained staff are required to manufacture this element



ating a new high-risk development from scratch. The aim of the exercise was to achieve a significant reduction in cost, while at the same time improving the performance of the sensors. With this in mind, the developers changed both the sensor element and its surrounding sensor cell, maintaining only the sensing principle. Changes included:

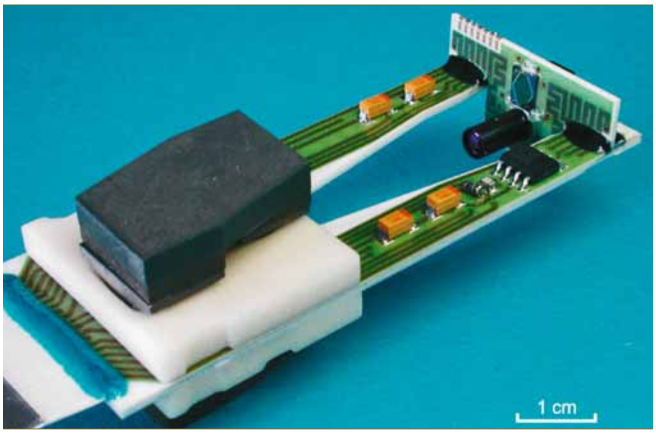
- Materials used (silicon and ceramics are used instead of glass and steel)
- Geometry (a planar setup replaces a complex three-dimensional arrangement)
- Manufacturing (silicon etching and automated batch processing is used instead of metal machining and manual assembly)

The result of this effort is shown in 6. Comparing this new sensor with the current sensor shown in 4, it is clear that the new design and technology used embody a paradigm shift in paramagnetic oxygen sensing.

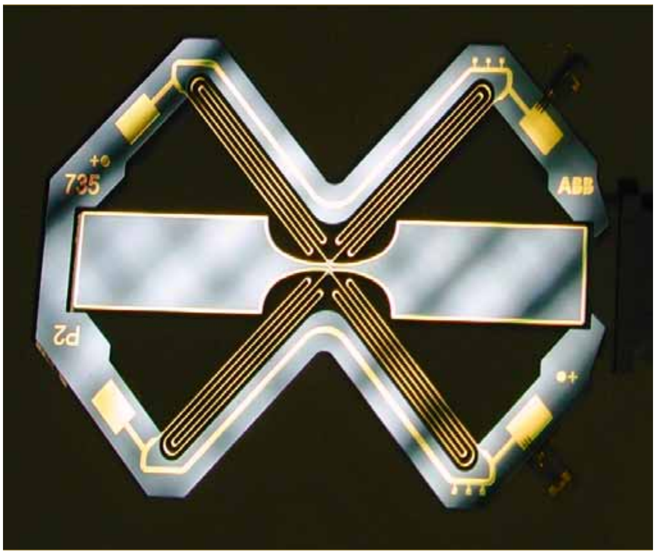
The "heart" of this new sensor is a micro-machined silicon chip 7 that provides fast response times and reproducible fabrication in large production volumes. The advantages over the old sensor are obvious: There is no 3-dimensional, manually-assembled setup. Instead there is a single planar chip. Nevertheless, this chip uses the same measurement principle: transforming oxygen concentration into rotational movement that is read out by a light beam.

The new setup follows a layered approach in a highly compact arrangement. The sensor basically comprises flat structures, bonded together to form the sensor cell 8. Ceramics are used to achieve high media compatibility and thermal stability. Electrical functions such as preamplification are integrated into some of the ceramic parts.

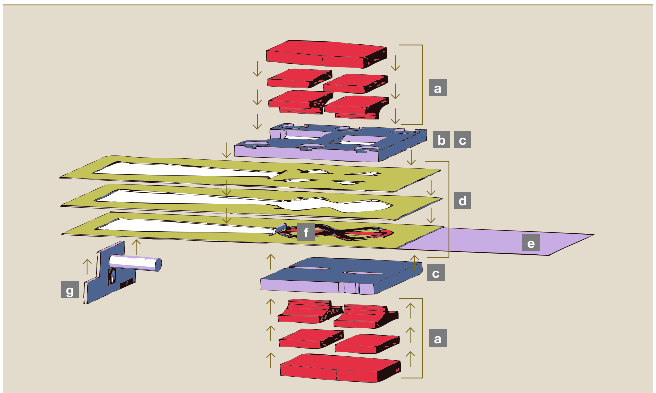
6 The new, highly integrated sensor module embodies a paradigm shift in paramagnetic oxygen sensing



7 New, planar micro-electromechanical system (MEMS) sensor chip (inner volume approximately 100 mm^3)



8 The layer concept to create a three-dimensional structure using "two-dimensional" parts



a magnetic circuit
b gas inlet system
c magnet carrier
d sensor cell
e electrical interface
f sensor chip
g opto-electrics

Breaking the rules: silicon chip oxygen measurement

The key decision made by the ABB developers was to combine a proven and well-known measurement principle with cutting-edge micro-electromechanical system (MEMS) technology and put MEMS technology to an entirely new use. The main performance advantage of the chip solution compared to the classical paramagnetic solution is the extreme reduction of the gas volume inside the sensor. The result is a much faster sensor, with response times about one second, a vast improvement on the three-second responses of conventional sensors. This opens up additional market segments where speed is essential (eg, the monitoring of combustion engines).

The planar sensor chip can be integrated easily into the layered structure of the sensor, allowing simple, automated assembly, as required for large production volumes at affordable cost.

Transferring the conventional, three-dimensional setup into a planar chip was made possible by a smart idea and a massive effort in the development of chip manufacturing technology.

This effort resulted in some outstanding features, including:

- Durability in aggressive gases (due to a bulk silicon structure with corrosion-resistant metal tracks)

- High reproducibility (due to the highly accurate etching process used)
- Fast response-time (about one second, due to thinness of the chip)
- Significant cost advantages in mass production

For ABB customers, the development will provide a new paramagnetic sensor with outstanding media compatibility and a range of other advantages. Long service intervals, together with low cost spare-parts, will significantly decrease life cycle costs. Increased reliability of process parameter measurements in highly corrosive media will improve product quality and safety.

The sensor basically comprises flat structures, bonded together to form the sensor cell.

Development tools

Turning away from 40 years of experience and initiating a new development from scratch would have been impossible without the use of powerful simulation tools to speed up development. During the development process, numerical simulation tools were used to optimize the mechanical, fluidic, magnetic, thermal, electrical and optical properties of the sensor.

The use of planar silicon-chip technology was made possible by the development of a new suspension system for the displacement body. There is no obvious way of creating a planar structure that will rotate easily and that is also relatively resistant to linear acceleration. 9 shows the simulated mechanical behavior of the newly developed suspension system that is included in the silicon sensor chip.

An ABB-proprietary magnetic field simulation tool was improved expressly to simulate not only the field distributions, but also their influence on the displacement body, taking into account the magnetic properties of the gas. The magnetic circuit is optimized to generate a maximum magnetic field gradient in the region of the displacement body.

Fluidic simulations helped to achieve the fast response time of the new sensors. The challenge was to bring a large volume flow (up to 100 l/h) close to the micro-sensor and exchange the gas inside the sensor cell as quickly as possible. At the same time, the influence of the gas flow on the displacement body had to be minimized. 10 shows the result of a gas flow velocity distribution simulation in the sensor cell.

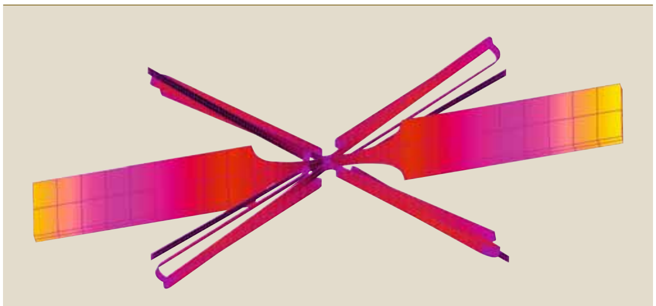
Chip manufacturing technology

The silicon chip in the new sensors is fabricated in two basic steps. First, the coil that compensates for the force generated by oxygen molecules is formed on the surface of the wafer by metal deposition and structuring. In the second step, deep reactive ion etching (DRIE) is used to etch fully through the wafer. The result of this process is shown in 11. Compared to the standard DRIE processes that are widely used in the production of micromechanical gyroscopes, the requirements of this application are rather more specialized.

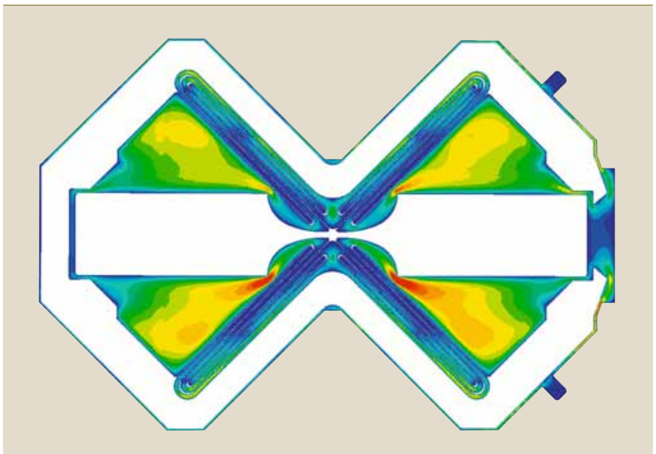
The main performance advantage of the chip solution is the extreme reduction of the gas volume inside the sensor.

The etched structure forms the suspension system for the displacement body with a minimum width of approximately 20 µm, over the total wafer thickness of several hundred microns. As the etched sidewalls are also used as a mirror to reflect the

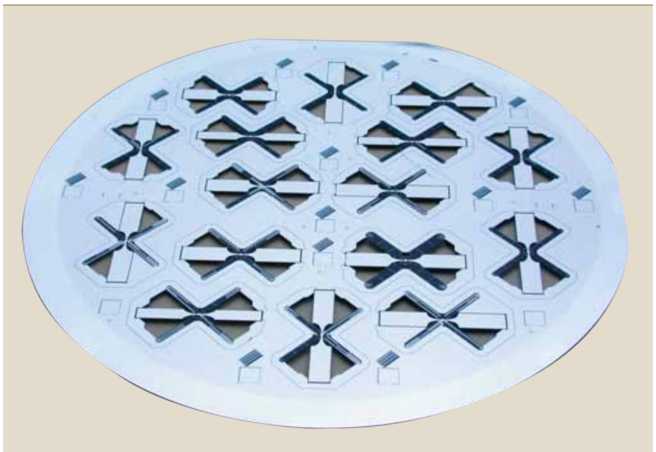
9 Simulation of mechanical behavior of the newly developed suspension system in the silicon sensor



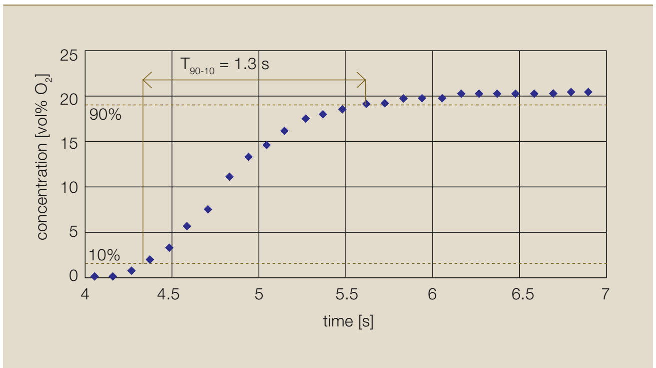
10 Simulation of gas velocity distribution in the sensor cell



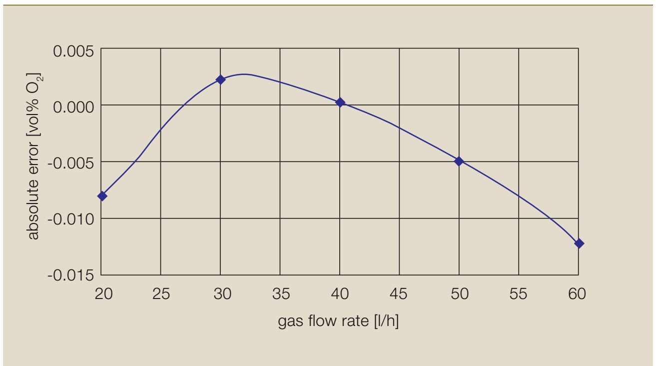
11 Silicon wafer with oxygen sensors



12 Response time of the new sensor



13 Absolute measurement error as a function of gas flow for air



light beam, there are additional extreme requirements with respect to roughness. Process development and design optimizations produced a sufficiently homogeneous and perpendicular side wall to fabricate the fine suspension system and the displacement body with its reflective side wall, thereby allowing a cost-effective, one-step etching process.

Fluidic simulations helped to achieve the fast response time of the new sensors.

The media compatibility of silicon and silicon oxide is sufficient for most gases, even those that are highly corrosive to other materials. The corrosion-sensitive points of these new sensors are the metal tracks. Using corrosion-proof rather than non-corrosion-proof metals that are commonly used in MEMS technology (eg, aluminum, gold) and specially qualified protective coatings on bond wires results in a sensor that can withstand wet chlorine atmospheres over months. This compares to a lifetime of less than a day in other "state-of-the-art" sensors.

Characteristics and performance

The new sensor system offers significant performance improvements with respect to response time and corrosion resistance. The response time (gas change from nitrogen to air at a volume flow of 60 l/h) of the new sensor is reduced to 1.3 seconds (see 12). Even more important than a fast response time is the stability of the sensor signal in the case of a variable gas flow rate 13. A flow rate varying between 20 l/h and 60 l/h causes a zero-point drift of only 150 ppm O_2, without any correction of the raw signal.

The development of this novel paramagnetic sensor for the measurement of oxygen concentration is a major achievement in the field of continuous gas analysis.

Concluding remarks

The development of this novel paramagnetic sensor for the measurement of oxygen concentration is a major achievement in the field of continuous gas analysis. The new sensor combines the high performance of the classic, paramagnetic sensors with the cost advantage of electrochemical sensors. It will provide ABB and its customers with a competitive advantage and will be a major influence on future trends in the measurement of oxygen concentration.

Factbox Patents arising from this work

- US patent 20020075007
  Device for measuring the oxygen concentration in gases.
- European patent 1 202 051
  Vorrichtung zum Messen der Sauerstoffkonzentration in Gasen unter der Verwendung eines inhomogenen magnetischen Feldes
- US patent 20040108442
  Device for suspension of a sample body.
- European patent 1 424 553
  Einrichtung mit Aufhängung eines Probenkörpers.

Peter Krippner
Paul Szasz
Manfred Wetzko
ABB AG
Ladenburg, Germany
peter.krippner@de.abb.com
paul.szasz@de.abb.com
manfred.wetzko@de.abb.com

Berthold Andres
Thomas Bauer
ABB Automation GmbH
Frankfurt, Germany
berthold.andres@de.abb.com
thomas.l.bauer@de.abb.com

1/2006:
Contemporary manufacturing challenges

3/2006:
Ingenuity at large

2/2006:
Embedded system technologies

4/2006:
Innovation highlights

Publisher's office
ABB Schweiz AG
Corporate Research
ABB Review/REV
CH-5405 Baden-Dättwil
Switzerland

The ABB Review is published four times a year in English, French, German, Spanish, Chinese and Russian.

Printers
Vorarlberger Verlagsanstalt GmbH.
AT-6850 Dornbirn/Austria

Layout
DAVILLA Werbeagentur GmbH
AT-6900 Bregenz/Austria

ISSN: 1013-3119

www.abb.com/abbreview

Preview 1/2007

The human factor as part of larger system configurations is a topic of significant interest both academically and practically. In some complex system applications, the human in the loop might be the weakest link and the greatest liability. In many systems, a system failure as a result of human error cannot be tolerated – production output, and more importantly, human lives may be at risk. The former category includes industrial manufacturing and power generation and distribution. The latter includes airplane cockpits, some industrial processes such as chemicals and nuclear, emergency systems of different types and traffic handling systems. Many methods have been adopted to minimize the risk associated with the human-in-the-loop issue, and indeed safe and secure operations of such systems are essential for a modern society. The more such applications are automated, the more the issue of the human factor becomes crucial. Both very fast processes such as electrical networks and much slower processes have their own complexities that need to be handled.

Areas of concern are decision support (including smart alarming), information visualization and ergonomic data presentation and ease of use from design to implementation. In the next issue of ABB Review with the theme "Human in the loop", we explore both academic research and practical implementations of recent innovations in these fields. For ABB, these topics are highly relevant to its businesses. As the world's leading automation supplier, the company needs to lead from the front.