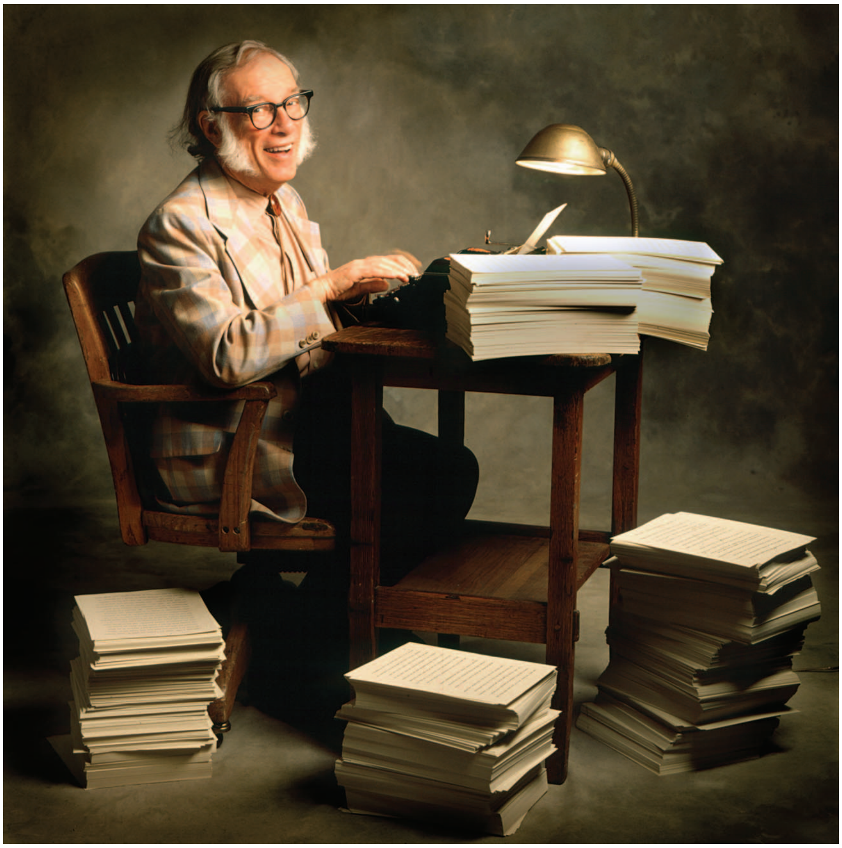
Back to the Future

By AL OPPENHEIM



Isaac Asimov once remarked that "the secret to being a successful prophet is to never try to predict something that will happen in your own lifetime." I hope that at least some of the predictions that I offer in this article are not either too naïve or too fanciful, but in any case, I take some comfort in Asimov's comment.

Looking back at the predictions for 2012 published by the Institute of Radio Engineers (IRE) in 1962, it is interesting to see how accurate many of them were, at least in broad strokes. Major advances in information storage and retrieval, processing of music and sound, the advances in automatic speech recognition were the topics of a number of the articles that seemed to anticipate the future, in many cases, very conservatively. In this article, I offer some speculation on where we might be in 2062, primarily with regard to two topics: the field of signal processing, and university education, building somewhat on the predictions made in 1962. Many of my speculations, I am sure, are obvious and will mature well before 2062. Others would seem to be reasonable speculations at this point but may be totally sidelined by unanticipated paradigm-shifting new technologies.

With regard to signal processing, it is a certainty that "there will always be signals, they will always need processing, and new technologies and mathematics will always emerge for implementing the processing." In 1962, virtually all real-time signal processing systems were analog and any digital signal processing (DSP but it was not called that then) was for purposes of simulation to adjust the parameters for the analog implementation of a system or for offline processing of massive data sets in such application areas as oil exploration, military intelligence, and surveillance. This type of DSP required rooms full of large, expensive, energy-hungry mainframe computers. The "big bang" in signal processing associated with the publication in 1964 of the Cooley–Tukey paper launched the use of the FFT as a major signal processing tool which fundamentally required implementation based on digital platforms. Furthermore, some of the new mathematics for signal processing that was emerging prompted the development of theoretically novel signal processing algorithms that were not realistically implementable in analog hardware. In that era, a fairly small group who was involved in the theoretical development of algorithms was somewhat naively speculating that "someday" integrated

Digital Object Identifier: 10.1109/JPROC.2012.2204109

circuit technology would make these algorithms practical in real-time applications. Often that optimism was not based on a deep understanding of technology but rather on dreams and fantasy, similar to the naivety of a kid who thinks that if he flaps his arms hard enough, he will fly. Comparing the current size, power and cost of DSP chips in 2012 with the DSP technology of 1962, the developments over the past 50 years have significantly exceeded many of the boldest predictions of 1962.

Looking forward, as technology for implementing signal processing systems continues to advance aggressively in a variety of directions I expect there to be a total blurring of the boundary between continuous-amplitude and discrete (i.e., quantized)-amplitude representations of signals in signal processing systems and also between clocked and unclocked processing. Current technology tends to force distinct boundaries between these signal representations with discrete-amplitude representations processed with digital computation and continuous-amplitude representations processed with "physics" (i.e., analog processing). A blurred interplay between digital and analog signal processing is currently emerging and will certainly accelerate over the next 50 years.

There will also continue to emerge fundamentally new ways of representing signals. The work in wavelets, compressed sensing, and other rich subspace signal representations is a precursor to this more major change. In particular, processing based on "knowledge-constrained" rather than just bandwidth-constrained domains will rapidly evolve. In 2012, there is increasing reference to "analog-to-information" conversion to capture a richer knowledge of what is important in a signal and in effect the information or knowledge subspaces in which it is contained. Preprocessing that is much more sophisticated than anti-aliasing bandwidth reduction will continue to emerge and well before 2062 signal processing will increasingly be carried out in these "knowledge-constrained" domains. Signal processing in knowledge-constrained domains will also naturally lead to "smart signals." Currently, some signal classes are already tagged with information on how the signal should be treated by the user, e.g., the TIFF format. Digital music formats also incorporate headers or tags which can provide the basis for sorting of files and signal equalization. As signal representations continue to advance I envision signals that are coded in knowledge-constrained domains to be able to have attached to them significant information about the processing options and the subjective intent of the processing to be carried out on them. Well before 2062 this information quite likely would be utilized during the storage, transmission, and routing of the signal, not just at the final destination. This will allow some or all of the intended processing to be carried out by smart routers and networks en-route to the destination and offline during storage.

With regard to signal processing platforms, there is every reason to speculate that over the next 50 years there will be even more dramatic developments than in the past 50 years. While over the next 50 years the pace of the development of silicon-based integrated circuit technology might not echo the last 50 years and Moore's law might have flattened out (although the demise of Moore's law has been repeatedly forecast and repeatedly been wrong), it is inevitable that new substrates for implementing signal processing platforms will emerge. One class in particular that I anticipate will have a significant role will be biologically based substrates for computation in general and signal processing in particular. Computational systems based on cell biology offer the potential for robustness, for "self-healing," and for massive parallelism. In another direction, carbon-based technologies (but not the color-coded carbon-based solid-state devices from my undergraduate years) such as graphene and carbon nanotubes might also become a strong competitor or complement to silicon, particularly in very high-speed devices. In integrated circuit fabrication, sophisticated folding of silicon membranes into 3-D and even 4-D (time-varying) structures will also enhance whatever trajectory Moore's law might take. Photonics is also destined to play a significantly expanding role in innovative signal processing systems. Optical systems and lasers (Fourier transforms with lenses, LIDAR, for example) have, of course, played a role in signal processing for many years. What now seems to be emerging and will certainly be a major technology for signal processing is the broad use of photons in place of or along with electrons in signal processing. As a precursor of some of this, see, for example, http://web.mit.edu/newsoffice/2011/lidar-3d-camera-cellphones-0105.html.

It is also quite likely that by 2062 the limiting factors of battery weight, size, and life will have been effectively and creatively resolved. In 2012, there are already significant advances in low-power circuit implementations and energy harvesting from the environment not unlike but obviously much more sophisticated than the self-winding watch and the 1928 atmos clock. By 2062, there will undoubtedly be major advances in energy storage technology, energy harvesting, and energy-efficient devices.

Innovative algorithm development always has and always will play a significant role in signal processing. In the December 2010 report from the President's Council of Advisors on Science and Technology it was comments that:

> "In many areas, performance gains due to improvements in algorithms have vastly exceeded even the dramatic performance gains due to increased processor speed."

There are many examples of this in signal processing such as the FFT algorithm and multirate algorithms and many more to come between now and 2062. Many widely used algorithms in 2012 were developed to be efficient in terms of metrics that are rapidly becoming less relevant, such as efficient use of memory and numerical computation. Increasingly relevant are efficient use of power and related to that, minimal on-chip communication. In some contexts, it is now often more efficient to use direct computation of the Fourier transform rather than the FFT to reduce on-chip communication cycles at the cost of more multiplications. This change in metrics is leading to a significant reinvention and restructuring of current algorithms to exploit more highly parallel architectures and platforms. The mathematical and software tools for organizing parallel and distributed processing of signals in an efficient and sophisticated way for the most part are not available in 2012. However, I am confident that these will be in place well before 2062. In 2012, quantum computing is being explored with intense positive and negative speculation on its promise. Quite possibly by 2062 we will have figured out practical ways of exploiting the immense computational parallelism that quantum computation seems to promise.

In addition to relying heavily on advances in implementation technology, signal processing has always benefited from innovations in mathematics for developing and describing signal processing algorithms. There are a number of aspects of signal processing that are likely over the next 50 years to either motivate or benefit from new mathematical developments. By 2062, signal processing is likely to have developed new mathematical formalisms for describing, analyzing, and synthesizing nonlinear systems and nonlinear signal transformations. Currently, signal processing systems often rely heavily on linear transformations. While some current signal processing systems exploit nonlinearities of various types, there is a general lack of formalism for designing nonlinear signal processing systems. Very powerful mathematical formalisms for synthesizing important subclasses of nonlinear transformations and signal representations will likely emerge by 2062.

In 2012, the mathematics for signal processing also tends to be heavily oriented either toward precise descriptions or toward capturing uncertainty and imprecision through the mathematics of probability. As the platforms for signal processing become massively complex and increasingly susceptible to error and faults during fabrication or over time, signal processing algorithms need to be developed and organized in ways that are fault tolerant, i.e., such that their performance degrades gracefully in the presence of implementation failures and in ways that are at least somewhat self-correcting. Furthermore, in many signal processing contexts, the required processing can often afford to be approximate, in the spirit that "good enough is good enough." While algorithms for "approximate processing" exist in 2012, I anticipate that a mathematical structure will have emerged by 2062 that will provide a strong basis for developing algorithms that are inherently fault tolerant and that can be as approximate as desired. Partially related to this, I also anticipate that new mathematical formalisms will emerge for describing and manipulating uncertainty. Uncertainty and imprecision in signal processing are most typically captured through the mathematics of probability which carries with it an underlying notion of an ensemble of signals. Often in signal processing, the ensemble is a fabrication and it is each individual signal that is important. Fuzzy logic has tried in some ways to capture this concept but it is my believe that much richer formalisms are yet to be developed. Related, in part, also is a need for a descriptive formalism for the "information content" of a signal that is not dependent on a probabilistic viewpoint. Such a description, when developed, will undoubtedly couple with the information-tagging of smart signals.

As the technology for implementing signal processing becomes smaller, cheaper, more energy efficient, its incorporation into devices to enhance our lives is inevitable. Energy-efficient and self-powered devices will clearly lead to the use of DSP algorithm in smaller, personal devices, including the ones like those mentioned for health monitoring (see "The doctor will see you always," IEEE SPECTRUM, October 2011, pp. 57–62). Over the next 50 years, there are likely to be innovative ingestible and implantable microscopic signal processing devices both for sensing and regulation of various body signals.

A close colleague and visionary at a major semiconductor company made the observation not too long ago that in Ben Franklin's day, eyeglasses were an indication of aging and impairment while in 2012 they are a fashion statement. Hearing aids in 2012 tend to be avoided if possible because they tend to be associated with hearing impairment and aging. By 2062, quite likely, sensory enhancers of various types will incorporate very sophisticated signal processing. Eyeglasses will evolve into vision enhancers with heads up displays and wireless connectivity. They will become essential not just for normal reading and distance but also for low light vision, and for feature enhancement and other image processing. In 2012, we see a clear trend in computational photography for the camera device to be the photon collector and for the role of the lens to be enhanced and even replaced by signal processing applied to the received photons. By 2062, quite likely the implementation technology and algorithms will result in wearable vision enhancers some aspects of which will have evolved from earlier work on computational photography. Hearing enhancers will undergo similar dramatic development. In 2012, we routinely have "things" (e.g., earbuds) in our ears to enjoy music, podcasts, etc. By 2062, these earbuds will be wireless, essentially invisible, and either incorporating or attached to very powerful signal processing. Effective enhancement will be provided not only for the hearing impaired but also for those with normal hearing. For example, active cancellation of unwanted sound sources will be routine to allow for quiet listening and conversation in noisy environments. Quite likely speech understanding, text-to-speech, and automatic language translation will also be incorporated in the hearing enhancers along with audio communication and entertainment sources.

In the 1962 predictions, a number of the articles offered speculations about what engineering education would look like in 2012. Much of the commentary focused on the likely development of teaching machines. In his article, Everitt focused particularly on education versus training and the likely development and use of "teaching machines" for the training component. In thinking about training versus education an analogy that often comes to mind is the development of a tennis player. A live coach who can motivate, adjust, and assess is irreplaceable. However, the training can also benefit significantly from the use of a nonjudgemental and tireless ball serving machine for repetitive drill. The same can be said for engineering education. In 2012, we have not come close to utilizing teaching machines in engineering education at the level achievable with current technology, perhaps because of the significant financial investment and associated man-hours required to develop the teaching machine content, interface, etc. In 2012, the Internet is clearly having a dramatic effect on education and in some sense is the teaching machine of today. It has opened the door significantly to "distance learning" and to on-demand instruction, but in 2012, much of its potential is so far largely untapped.

Correspondingly, in 50 years, the style of residential university education will undoubtedly change significantly. Ponte in his 1962 article speculated about the "HyperTexas University" which capitalized dramatically on significantly faster and broader band communication channels than were available in 1962. Much of what he envisioned in detail has not yet come to pass and the form that much of that will take by 2062 will likely be significantly different than he envisioned because of the unforeseen (in 1962) development and impact of the Internet. It has long been clear to many of us in academia that the basic residential university model in which students live on or near the campus and come to a specific place at a specific time for delivery of class material in segments of one or multiple hours (lectures and smaller class gatherings) is not ideal. Typically low attendance at these scheduled events strongly suggests that that model is not the most desirable for content delivery or desired by the students. It is widely recognized in 2012 that the Internet is a key engine of change for that model. It seems to be without question that long before 2062, the university structure and mission as we know them in 2012 (and more or less what it has been for many decades prior) will change significantly in large part because of what the Internet promises in terms of the delivery of both training and education. While the primary mission of a university is education, a strong research program at research universities is nevertheless an important component of that mission. At the advanced level, research and teaching are symbiotic, with the most advanced results quickly finding their way into the syllabus and a strong research environment affording students the opportunity to be closely mentored by experts who are on the cutting edge of their field. The development of innovative and forward-looking syllabi depends on a deep awareness of where a field is heading. Consequently, the role of research at universities will almost certainly be enhanced, providing increased interplay between the research and the development of course content, and with enhancement of the opportunity for students to work closely with and be mentored by world-class research staff, faculty, and more advanced students. In his 1962 article, Everitt quoted Booker T. Washington: "I am convinced that there is no education that one can get from books and costly apparatus that is equal to that which can be gotten from contact with great men and women." That quote will be as relevant in 2062 as it was in 1962.

What seems clearly to be on the verge of change in 2012 and will fully impact the structure and role of research universities well before 2062 is how, when, and where course content is delivered. Truly gifted people for delivery of well-developed course content at an introductory level are often more likely to be found at colleges and universities that are primarily focused on teaching, than at top research universities. Truly gifted people for innovative and visionary development of course content and for more advanced mentoring are more likely to be found at cutting edge research universities. As the use of the Internet for delivery of both education and training advances, I envision an increasing separation in the roles of the research universities and the teaching colleges and universities with strong partnerships developing. For example, an undergraduate degree quite likely will be awarded to a student and/or accredited by a top research university utilizing largely online lectures delivered remotely and in an on-demand format by a gifted teacher from any one of a number of other institutions or even from freelance or commercially oriented teaching organizations. To a limited extent that has always been done in the sense that courses at all universities take advantage of outside resources such as textbooks, supplementary materials, laboratory platforms, etc., developed at other universities or companies. The access provided by the Internet to gifted teachers and tutors worldwide can be viewed as just a significant expansion and enrichment of the reservoir of resources. While both the evaluation of students' understanding and the course syllabus will be designed and carried out by the accrediting institution, the delivery of the content will be remote and on-demand for the students. This has many advantages to both the students and the universities. This will likely mean that faculty at the top research universities will be less involved in day to day classroom teaching, particularly at the foundational level, and more involved in course content and syllabus development and in mentoring at the more advanced levels. It will also almost certainly mean that the requirement for students to reside at or near the accrediting university will be significantly diminished. Various models of this type are already being developed and experimented with in 2012, just one example being the program announced in 2011 by the Massachusetts Institute of Technology http://web.mit.edu/newsoffice/2011/mitx-education-initiative-1219.html. Many other similar and alternative models are certain to follow and well before 2062. I anticipate that the style, character, and "business model" of university education, certification, course content delivery, and residential requirements will have changed dramatically from what they are in 2012. ■