

# Olfaction Metal Oxide Semiconductor Gas Sensors and Neural Networks

M. W. Siegel  
Carnegie Mellon University  
The Robotics Institute  
Pittsburgh, PA 15213  
USA

## Introduction

Primitive mobile life forms, and the mobile cells of higher animals, derive their motivation and direct their navigation by *chemical* senses. In the simplest cases these creatures are hard-wired to swim toward nutrient concentration gradients, and to swim against irritant concentration gradients. The vestiges in humans of the sometimes extremely sensitive, selective, and differential chemical senses of primitive forms are taste, our ability to detect and identify four classes of chemicals in water solution on our tongues, and smell or olfaction, our ability to detect and identify many gases, vapors, and complex mixtures in the air passing through our noses.

Primitive "autonomous" mobile robots derive their motivation from human directive, and direct their navigation largely via artificial visual and acoustic senses. Their autonomy is thus in the discretion they are allowed in planning and executing getting from here to there; the *heres* and the *theres* generally are dictated more-or-less explicitly by humans. An evolutionary path for sophistication of the navigational abilities of mobile robots looks self-evident, and in fact such evolution is well underway. On the other hand an evolutionary path along which mobile robots might even begin to progress toward replacing motivation by human fiat with self directed motivation is difficult to conceive, at least if the senses that drive motivation are to remain as complex as vision and hearing.

But self directed motivation of mobile robots based on one or more chemical senses is easy to imagine, and indeed is probably within the capability of existing technology. For example, we could build right now a mobile robot that would meander around a chemical plant, sniffing as it goes for gas leaks (or the vapors of liquid leaks), navigating toward them while avoiding hazards visually, yet always motivated as to overall direction (an undetermined endpoint notwithstanding) by the chemical concentration gradient, with end-point navigation perhaps directed to the offending pipe fissure or open valve by acoustic homing toward the source of the hiss, and effecting simple repairs or summoning human assistance (with appropriate hazard warning) based on the fusion of chemical, visual, and acoustic sensory information in context. Similar scenarios can easily be imagined for firefighting robots, prospecting robots, rescue robots, contraband interception robots, and others.

Toward this end, my colleagues and I have been studying chemical sensing, not with the traditional goal of inventing alternative or improved instruments for doing precise chemical analysis, but rather from the perspective of exploring alternative approaches to an artificial sense of smell to motivate autonomous activity.

First, I describe the characteristics of metal oxide semiconductor (MOS) resistors used as gas sensors, particularly for combustibles. Response, as a change in resistance, depends on combustible gas identity and concentration, sensor temperature, and also on the concentrations of other gases and vapors present, *i.e.*, response to a mixture is *not* a linear superposition of responses to the individual components. These sensors inherently respond to broad classes of compounds, but differential responses conducive to signature identification can be induced by variation of parameters such as temperature and catalyst nature and concentration. I then survey the olfactory system, in contrast with other biological sensory and sensor interpretation systems, in analogy with MOS gas sensors, and in analogy between its neurological organization and the organization of simple artificial neural networks. I then describe the class of large area thick film MOS resistors that we have constructed with a spatial gradient in relative sensitivity induced by differential heating. Multiple resistance measurements along the temperature gradient yield a multi-feature signature that we analyze and interpret using artificial neural networks. In particular, I contrast the approach that emphasizes binary or nearly binary unit activities, and thus forces every mixture into a class distinct from the classes of its single species components, with an approach that is less rigidly binary, encouraging creation of only enough classes to account for the individual species present, and putting mixtures into multiple classes in proportion to the concentration of the mixture components in each class.

## Metal Oxide Semiconductor (MOS) Gas Sensors

The resistance of MOS gas sensors [1] is known empirically to have a power law dependence on oxygen concentration directly and on combustible (reducing) gas concentration inversely. For first order reactions, interpretable as one molecule of combustible species  $X$  reacting with one  $O_2$  molecule



the behavior is well described by

$$R = R_o \left( \frac{[O_2]}{1 + K_1^X [X]} \right)^\beta \quad (2)$$

Clifford and Tuma [2, 3] have observed this power law response for a variety of commercial MOS sensors, and they have found the constant  $\beta$  for Figaro Corporation Taguchi Gas Sensor (TGS) #812 to be between 0.25 and 0.55, depending unpredictably on the particular sensor in hand.

Some insight into this behavior is provided by a seat-of-the-pants model of the interplay between conduction electrons, oxygen, and reducing gases in and on metal oxides such as  $SnO_2$ . The oxides are manufactured with the normal metal-oxygen stoichiometry disturbed in the direction that leaves them, in the absence of atmospheric oxygen, slightly reduced, *i.e.*,  $SnO_{2-\epsilon}$  rather than  $SnO_2$ . The small excess free metal content contributes conduction electrons that make for much lower electrical resistance than is the norm for fully oxidized metals. Atmospheric oxygen captures free electrons or depletes free metal, increasing the electrical resistance. Reducing gases deplete oxygen or restore the free metal, reducing electrical resistance. Thus resistance increases with oxygen partial pressure and decreases with reducing gas partial pressure.

A simple generic equilibrium picture [4] is adequate to transform  $[O_2]_{atm}$  and  $[X]_o$  concentrations outside the sensor and  $[e^-]_{vac}$  electron density inside the sensor under vacuum conditions into  $[O_2]$ ,  $[X]$ , and  $[e^-]$  inside the sensor with the atmosphere admitted. The three reactions



yield, by eliminating  $[O_2^-]$ ,  $[X^-]$ , and  $[XO_2]$  from the three equilibrium equations and three conservation equations implied by reactions 3, 4, and 5:

$$[e^-] = \frac{[e^-]_{vac}}{1 + k_1[O_2] + k_2[X]} \quad (6)$$

$$[O_2] = \frac{[O_2]_{atm}}{1 + k_1[e^-] + k_3[X]} \quad (7)$$

$$[X] = \frac{[X]_o}{1 + k_2[e^-] + k_3[O_2]} \quad (8)$$

There is, of course, no general closed form simultaneous solution to equations 6, 7, and 8; however with the reasonable conjecture

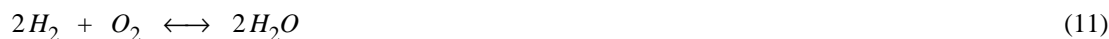
$$\frac{R}{R_o} = \left( \frac{[e^-]_{atm}}{[e^-]} \right)^\beta \quad (9)$$

equation 2 can be convincingly argued to be a plausible solution in physically realistic limits.

In many important cases the response is not first order. For example  $H_2$  is observed to obey

$$R = R_o \left( \frac{[O_2]}{1 + K_2^{H_2} [H_2]^2} \right)^\beta \quad (10)$$

and is thus phenomenologically equivalent (as expected) to



The response to  $CO$  depends on the concentration of water vapor, and thus *vice versa* the response to water vapor depends in part on the concentration of  $CO$ :

$$R = R_o \left( \frac{[O_2]}{1 + K_{11}^{COH_2O} [H_2O][CO]} \right)^\beta \quad (12)$$

The notations  $K_2^{H_2}$  and  $K_{11}^{COH_2O}$  re-iterate that reactions leading to equations 10 and 12 are respectively second order in  $[H_2]$  and first order in each of  $[CO]$  and  $[H_2O]$ .

In general, terms of many orders, describing multiple reaction and response modalities, are simultaneously effective. All the parallel channels are summarized by

$$R = R_o \left( \frac{[O_2]}{\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \dots K_{ijk\dots}^{XYZ\dots} [X]^i [Y]^j [Z]^k \dots} \right)^\beta \quad (13)$$

with the understanding that  $K_{000\dots}^{XYZ\dots} \equiv 1$  for all mixtures  $X, Y, Z, \dots$ . Since reaction orders that are ratios of small integers are common, the indices  $i, j, k, \dots$  symbolize the rational numbers, not just the integers. However, in practice, few processes are observed at orders greater than second.

The rate constants  $K_{ijk\dots}^{XYZ\dots}$  depend on the rate of collisions, and on the probability that on any given collision the reaction will actually occur. The rate of collisions depends on the translational speed

which is proportional to the square root of the temperature. The reaction rates then depend on the products of the rate constants and powers of the number densities of the reactants appropriate to the orders of the reactions. The number densities are proportional to the concentrations  $[X], [Y], [Z], \dots$ , *i.e.*, the  $K_{ijk\dots}^{XYZ\dots}$  essentially measure the reaction rates per unit concentrations. The probability that a collision will result in a reaction depends primarily on an activation energy  $E_a$  which is supplied by thermal motion. The reaction probability is then given by the Boltzmann factor  $e^{-E_a/kT}$ . The temperature range of interest is usually rather small, rarely more than a factor of two or three between room temperature ( $\approx 300\text{ K}$ ) and thermal destruction of the sensor. Activation energies are typically  $\approx 1.5\text{ eV}$ , some sixty times the energy of thermal motion at room temperature. Thus between room temperature and thermal destruction the collision frequency changes by less than a factor of two, whereas the Boltzmann factor may change by several orders-of-magnitude. The dependence of the rate constant on temperature is thus well approximated by

$$K_{ijk\dots}^{XYZ\dots}(T) \approx K_{ijk\dots}^{XYZ\dots}(0) e^{-E_a/kT} \quad (14)$$

I have given no indication, so far, of the magnitude of the constants in the model as it might be applied to a real material such as  $\text{SnO}_2$ , *i.e.*, of the *sensitivity* of MOS sensors to gases of potential interest. This is the good news: high sensitivities are easy to obtain, *e.g.*, the resistances of the TGS #812 sensors discussed by Clifford and Tuma typically drop from  $10^4\ \Omega$  to  $10^3\ \Omega$  in response to a change in  $[\text{H}_2]$  from  $100\text{ ppm}$  to  $1000\text{ ppm}$ . The model suggests no reason to expect any intrinsic specificity, other than the low level specificity of response with decreasing resistance to reducing gases, increasing resistance to oxidizing gases, and not at all to more-or-less inert gases. This then is the bad news: MOS sensors are not particularly selective, so while they are good for indicating the presence of *some* combustible gas, they are, one at a time, unsuitable for reporting *which* combustible gas they have detected. On the other hand arrays of sensors with even small differential sensitivities from element to element offer the opportunity both to identify and to quantitate gases and gas mixtures by a variety of signature separation methods.

## Olfaction and Other Senses

In contrast to the visual and auditory senses, the human chemical senses of taste and smell are poorly understood. This has more to do, I think, with the essentially parametric nature of sight and hearing vs the essentially non-parametric nature of taste and smell than it has to do with any particular complexity of biological implementation at the cellular transducer level. In fact the chemical sense organs are structurally simpler, more primitive and apparently less differentiated than are their optical and acoustic counterparts, so the relative levels of our understanding would be reversed were understanding dominated by physiology.

The three types of color vision sensory cells are, via their incorporation of red, green and blue color pigments, sensitive to three overlapping but distinct bands of the optical continuum. Auditory sensory cells cover more but narrower bands, each receptor mechanically tuned to a specific vibrational frequency. Thus color vision and sound are *parametric* sensory modalities: each cellular transducer can be characterized by a single parameter, center frequency of its optical or acoustic response respectively. Taste, with salt, sweet, sour, and bitter constituting four primaries, is often compared

with color vision and its red, green, and blue primaries, but the analogy is superficial. Whereas color is a *measurement* along a continuum spanning the optical frequency range from  $4 \times 10^{14}$  Hz (red) to  $1 \times 10^{15}$  Hz (blue), and pitch is a *measurement* spanning the acoustic frequency range from somewhere above 16 Hz ( $C_0$ ) to somewhere below 32 kHz ( $C_{11}$ ), there is no such continuum along which salt, sweet, sour, and bitter fall.<sup>1</sup> The Doppler shift can transform red into blue, or buzz into beep, but no relativity effect can turn salty (presence of sodium-like metal ions) into sweet (presence of certain carbohydrates with OH appendages). Thus taste is a more like a *classification* than a *measurement*.

Whether smell is more like a *classification* or more like a *measurement* turns out to be unresolved. Amoore [5] correlated elaborate odor classification trials by human subjects with stereochemical data. His results support a geometrical lock-and-key classification model reminiscent of Lucretius: a few tens of primaries (camphor, musky, floral, peppermint, ether, ... ) each corresponding to a molecular shape (sphere, disk, disk-plus-tail, wedge, rod, ... ) from which a set of shape-sensitive receptors can be inferred. Small molecules might mate with different receptors by presenting themselves in different orientations, and huge molecules might fit different side group protuberances on their complex surfaces into a variety of receptors. Thus a single type of intricately shaped molecule might evoke an olfactory symphony, whereas a mixture of chemically different but more-or-less identically shaped molecules might stimulate but a single note.

On the other hand Gesteland [6], with the Lettvin and Pitts of frog's-eye fame [7], studied the frog's olfactory epithelium and concluded that every cell was sensitive (with either an excitatory or inhibitory effect), in varying proportion from cell to cell, to a large subset if not all the chemicals that the frog could smell:

"... if we choose pairs of odours which smell similar to each other or which are sterically similar by Amoore's categories, we can often find a single fibre clearly discriminating between them without too much searching ... it would be possible then, by having very many fibres, to distinguish very many compounds, for we would treat the characteristic of every fibre as a separate dimension."

Although Gesteland *et al* are careful to remind us that frogs are frogs and people are people, the contrast between their conclusions drawn from direct measurements on cells stimulated by odors and Amoore's conclusions drawn from human perception experiments is so stark as apparently to leave the door open for all manner of speculation, which Lettvin and Gesteland [8] do unabashedly in a companion paper to Gesteland, Lettvin, and Pitts [6].

Whatever lack of resolution there is about the olfactory transducers' discriminatory abilities by chemical selectivity (in contrast to computational discrimination), the structure of the wiring diagram of nerve fibers connecting the olfactory sensor layer with the olfactory cognition region of the brain is noncontroversial. With unusual simplicity, presumably due to the antiquity of the machinery, a single layer of interconnections separates the sensory apparatus in the nose from the cognitive apparatus in the brain. The roof of the human nasal cavity contains sensory patches, each about  $5 \text{ cm}^2$  in area, comprised of about  $50 \times 10^6$  ciliated olfactory cells. Each cell sends a single nerve fiber directly up, through the cribriform plate that separates nasal cavity from brain, into the adjacent olfactory bulb on the brain side of the plate. About 1000 primary receptor neurons converge on each secondary mitral

---

1. Not quite: sour being the taste sensation of acids and bitter being the taste sensation of bases, sour and bitter do fall on the *pH* continuum. But the deadband encompassing the neutral region around *pH* 7 (water) is so large that it cannot really be said that the sour-bitter components of taste measure *pH*.

cell. Axons from the mitral cells connect the olfactory bulb to the cognitive areas of the brain. In short, the olfactory apparatus looks a lot like a three-layered perceptron [9, 10] with a layer of sensory units corresponding to the cells of the olfactory patch at the top of the nasal cavity, a layer of associative units corresponding to the mitral cells of the olfactory bulb on the other side of the cribriform plate, and a layer of response units<sup>2</sup> in a cognitive area of the brain found in a fold of the cortex on the under surface, near the temporal lobe [11].

The correspondence between the physiology of the olfactory machinery and the rubric of artificial neural network models of sensory perception is close, but the match to *simple* models, *e.g.*, layer-to-adjacent-layer connections with a few to a few tens of connections per unit, is arguable. It is certainly obvious that the thousand-fold convergence from sensory to associative layer goes far beyond anything that is typically simulated. Whether the sensory layer is equipped with one class of sensors each of which has a different set of sensitivities to essentially every odorous chemical, or alternatively, equipped with several classes of sensors that pre-classify according to some chemical criteria, *e.g.*, molecular shape, but within each class may have a broad range of sensitivities to a variety of species, is also unclear.<sup>3</sup> The MOS gas sensor that we have built and are about to describe has multiple sensory elements that differ from each other quantitatively rather than qualitatively, and are thus analogous to either the former one class olfactory model, or to one class or channel of the latter multiple class olfactory model.

## Sensor, Apparatus, and Measurements

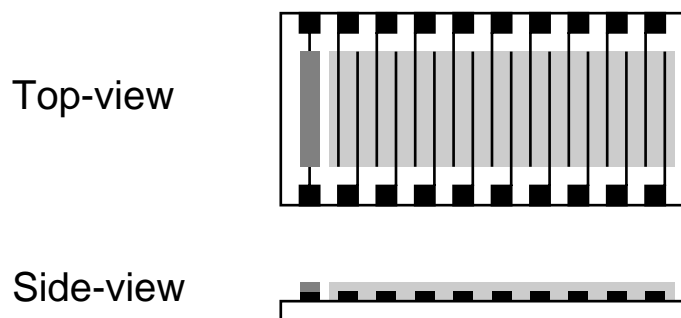
Our sensors are fabricated on alumina ceramic substrates intended for mounting hybrid circuits. The size and shape of the substrate is compatible with a 20-pin dual-in-line integrated circuit. A gold contact pad and electrode pattern is printed, then fired at high temperature onto the substrate, as shown in Figure 1. Next a layer of a modified commercial  $\text{SnO}_2$  semiconductor ink is screen printed over the electrodes. The modified  $\text{SnO}_2$  ink and the screen printing technique provide a cost-effective method to fabricate a high quality sensor array. Catalysts such as fine Pt metallic powder are blended into the commercial inks to enhance sensitivity. The organic binders in the ink are burned away by re-firing. Detailed descriptions of the fabrication procedure, which is actually carried out by our collaborators at Oak Ridge National Laboratories, is given in [12, 13, 14].

For controlled testing the sensor is mounted in a holder that incorporates pressure contact probes, electrical heaters, and temperature sensing devices. A temperature gradient is maintained along the length of the sensor by supplementary heat applied to one end of the substrate with a small, high power ceramic heater. The temperature difference between the hot end and cold end of the sensor surface is typically about  $100\text{ C}$ . The assembly is installed in an air-tight metal test chamber supplied with a constant flow of either pure synthetic air or dilute sample vapor in synthetic air. Vapor samples

---

2. I prefer the graphic *sensory-*, *associative-*, and *response-layer* terminology employed in the early perceptron models of Rosenblatt [9] and Block [10] to the *input-*, *hidden-*, and *output-layer* terminology used in most of the current literature on simulated artificial neural networks.

3. It is clear that the specialized olfactory apparatus of certain invertebrates, *e.g.*, moths that respond to a single molecule of their species sex attractant, illustrates an entirely different approach to chemical sensing, an approach in which astounding sensitivity has been purchased by dedicating enormous resources to apparatus with but a single narrow function.



**Figure 1.** INTEGRATED GAS SENSOR ARCHITECTURE

The gold metallization is shown in black, the  $\text{SnO}_2$  sensing layer shown crosshatched. The disconnected element on the left is an on-chip heater.

are generated by saturating a secondary synthetic air flow in a bubbler, then diluting it in a precision flow controlled mixing system. Binary mixtures are made by turbulent mixing of the outputs of two bubblers. Vapor concentrations are computed from the relative volumetric flow rates, temperature measurements, and vapor pressure tables. A total hydrocarbon analyzer down-stream of the test chamber is used for confirmation.

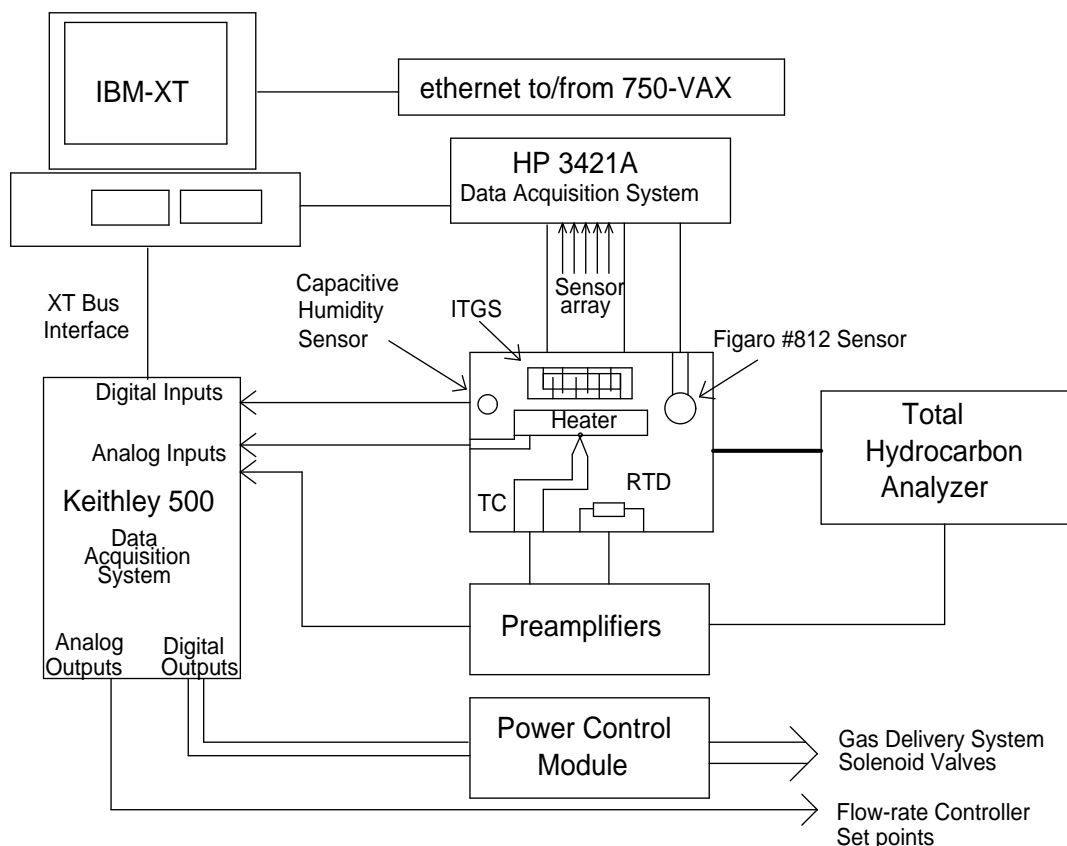
Before taking any measurements, the test chamber is flushed with synthetic air at  $1000 \text{ ml/min}$  flow rate for at least  $5 \text{ hr}$ . Then in a typical experiment the sample gas is turbulently mixed with the main air flow and is introduced into the test chamber for  $50 \text{ min}$ . The response patterns to various gases, vapors, and mixtures are recorded by measuring the resistances cyclically. Figure 3 shows typical data.

A simplified block diagram of the measurement and support system is shown in Figure 2. The neural network simulator is generally run off-line on the VAX, although it is PC compatible and we sometimes run it on-line.

The responses plotted in Figure 3 are  $(R_o - R)/R_o$ , the usual feature presented in the solid state gas sensing literature. Each sensitivity to each of the test compounds *ethanol*, *methanol*, and *heptane* increases with operating temperature in the temperature range of these experiments. At the hot end, each absolute sensitivity is three to four times higher than at the cold end, and of a similar magnitude to the TGS #812 sensor. To minimize irreversible changes due to abrupt environmental shock to the sensor, several measurement cycles are made with each sample vapor before changing to another vapor. This method, as opposed to more frequent cycling of sample type, increases vulnerability to slow drift. The drift is compensated by normalization of the data before presentation to the sensory layer of the neural network simulator.

## Classification by Neural Network

A neural network is a multi-input, multi-output machine, usually simulated but sometimes realized physically, with at least two layers of neuron-like nodes designated the "input" or *sensory* layer and the "output" or *response* layer [15, 16]. Neural networks with one or more additional "hidden" or



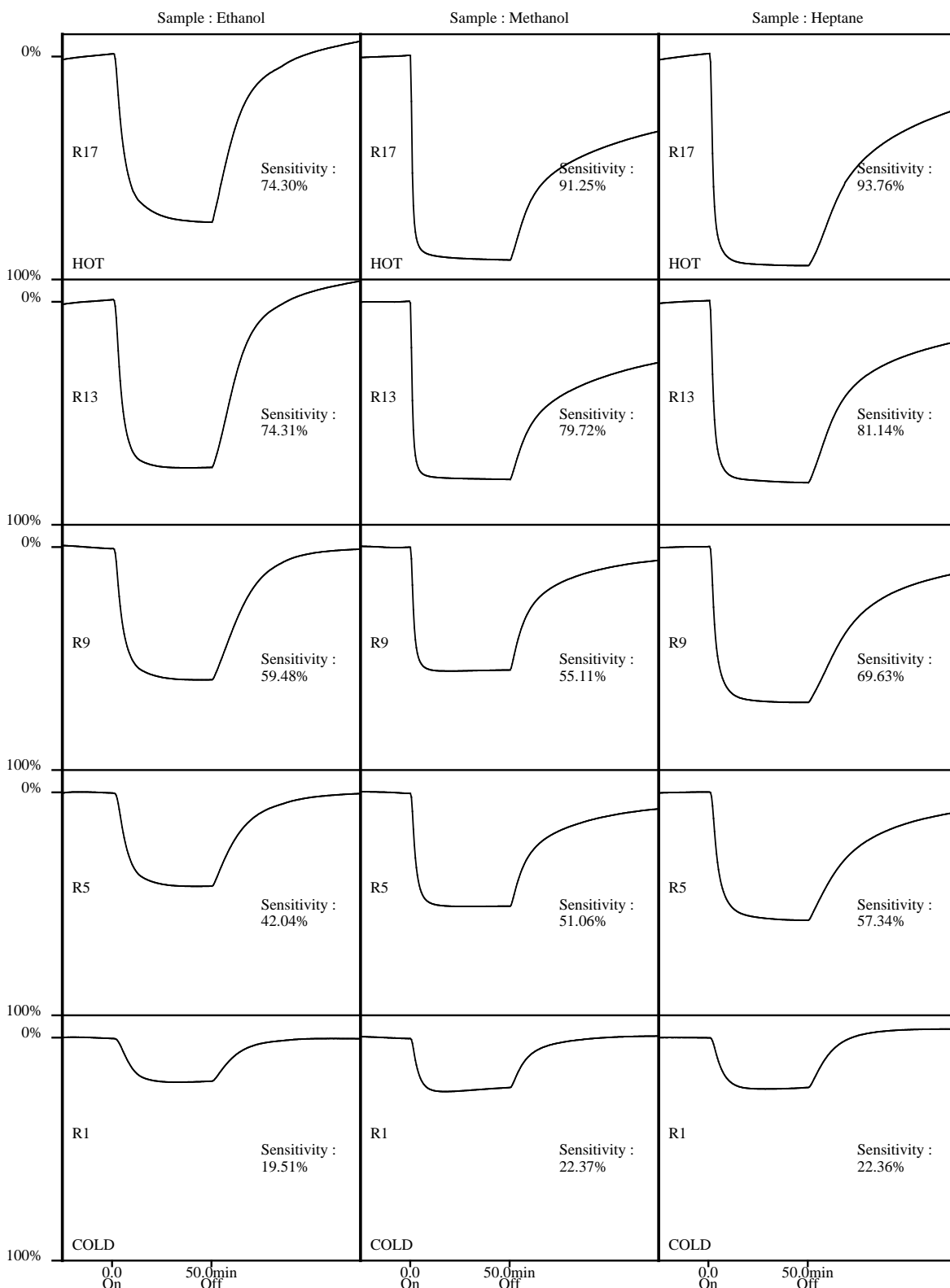
**Figure 2.** DATA ACQUISITION AND CONTROL SYSTEM

High precision measurement of sensor resistances is by an HP 3421A data acquisition system on the IEEE-4888 bus. Moderate and low precision monitoring of housekeeping and support sensors, and control of valves and heaters, is by a Keithley 500 data acquisition system memory mapped on the IBM-XT bus. Preamplifiers for thermocouples and RTDs are Analog Devices modules. The total hydrocarbon analyzer is a Mine Safety Appliances flame ionization detector. The FID and the Figaro TGS #812 sensor are used for checking and comparison. The capacitive humidity sensor is read by the Keithley unit at the capacitance-to-frequency output of a capacitance meter.

*associative* layers are of particular interest because the associative layers are able to hold internal representations that make for a higher level of abstraction than is possible without them.<sup>4</sup> Each layer is assigned a predetermined number of nodes or *units*. Each unit is characterized by a unique threshold (inhibition of stimulation) or bias (propensity toward stimulation), and a set of weights that characterize the strength of connections to higher and lower layers. Each unit has an output or *activity* that is a function of the sum of its inputs, *i.e.*, the activities of the previous layer weighted by the strength of the connections, offset by its threshold or bias. Rather than connecting to lower neural layers, the sensory units couple the network to some information about the external world. Similarly, rather than connecting to higher neural layers, the response units actuate some behavior in the external world. There are many variations on all the themes, including cross-connections within layers, back-connections from higher to lower layers, and connections that jump layers.

4. The classic example is the XOR problem, with two inputs and one output, and the abstraction that what associates inputs into the same class is their *dissimilarity*; an XOR unit cannot be simulated by any two layer network, but it can be simulated by a three layer network with just a single associative unit.





**Figure 3.** SENSITIVITY TO ETHANOL, METHANOL AND HEPTANE VAPORS

Sample vapor from a bubbler is mixed into an artificial air flow and introduced to the test chamber for 50 min. By convention the sensitivity is given as  $(R_o - R)/R_o$ . The time constant of the rapid resistance drop at the hot end of the sensor upon *methanol* and *heptane* introduction is close to the limit set by the ratio of test chamber volume to volumetric flow rate. The longer fall times for *ethanol*, and for *methanol* and *heptane* on the cooler parts of the sensor, and the longer rise times for all samples at all temperatures upon termination of sample introduction, are characteristic of the sensor.

The classifying power of a multi-layer neural network depends, fundamentally, on what kinds of input-output functions it is able to represent by virtue of its topology and the definition of the input-output functionality of the individual units. The gap between the discriminant functions that a given network can represent and a particular discriminant function instantiation that successfully segments the members of a large data set into a small number of relevant constituent classes is the existence of a teaching and learning procedure that specifies the desired response to a few pieces of archtypal data, and in a reasonable number of steps each taking a reasonable time incrementally and convergently modifies the weights and thresholds until some acceptable level of performance is demonstrated. The concept long predates the discovery of an efficient teaching and learning procedure. The recently invented *back propagation algorithm* has bridged the gap by providing a practical and effective teach-learn paradigm. The principle is that the output errors are propagated back through the network from output to input, and the resulting "activities" determine the sign and magnitude of the modifications made to each threshold and weight.

My colleagues Lanwai Wong and Toshikazu Takemori have assembled, from bits-and-pieces of approaches and alternatives in the recent literature [17], a three layer neural network simulator with back propagation, with sixteen to thirty-three input units, generally five to ten hidden units, and six output units for classifying the three combustible vapors, *ethanol*, *methanol*, and *heptane*, and their more-or-less equal concentration binary mixtures viewed for this purpose as separate classes [18]. Their implementation<sup>5</sup> of the neural network simulator and back propagation training algorithm is illustrated in Figure 4.

During training the simulator is cyclically presented with pairs of sensor data vectors **s\_data** and the corresponding desired response vectors **r\_class**, and adjustments to weights and thresholds are made according to the back propagation algorithm. When the network is used as a classifier **s\_data** are presented sequentially and for each an approximation to **r\_class** is returned.

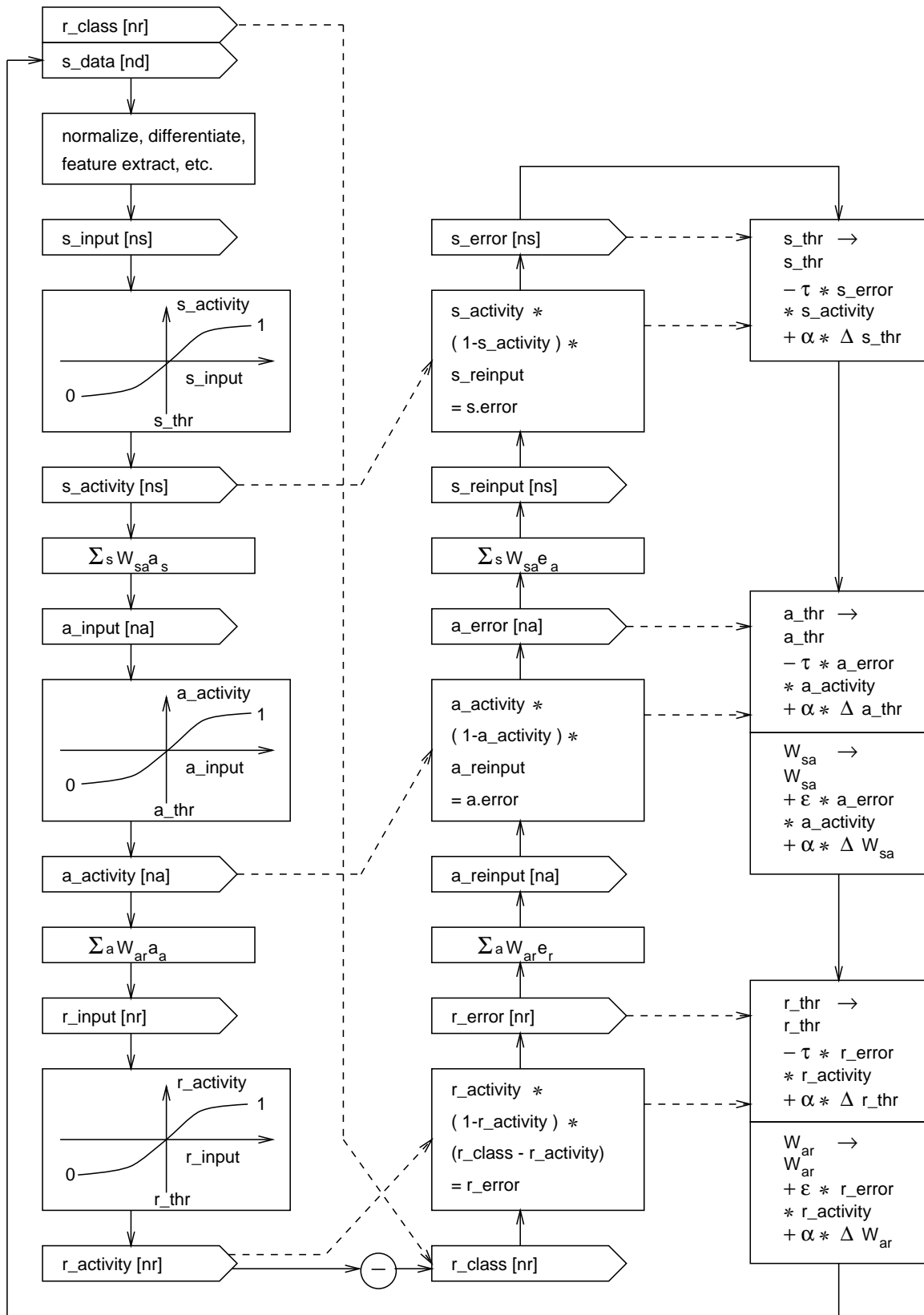
The **s\_data**-components are the sensor responses derived from the seventeen resistance measurements made across segments of a sensor like the one depicted in Figure 1. A linear preprocessing algorithm scales and normalizes **s\_data** into the sensory layer input vector **s\_input**. We regard this as an elementary sort of feature extraction. Our favorite preprocessing algorithms are (1) normalization to the range 0 to 1 by subtraction of the minimum and dividing by the range in each data vector, which yields a seventeen element **s\_input**, (2) differencing adjacent element resistances, or "spatial differentiation," which yields a sixteen element **s\_input**, and (3) the combination of (1) and (2), which yields a thirty-three element **s\_input**. The approach that we took in my work with Wong and Takemori treats each individual gas and binary mixture if it were a separate class,<sup>6</sup> so **r\_class** is a 6-dimensional vector. Each **r\_class** used in training has five zero components and one unity component that corresponds to the class specified.

Each sensor unit's **s\_input**-component  $s_{input_{s\_unit}}$  is passed through a sigmoidal thresholding and activation function of the form

---

5. They did not actually employ thresholding and an activation function at the sensory level, but for generality I have indicated the possibility.

6. In the following section I will discuss the alternative (and I think preferable) viewpoint that this is a three class problem that is best described by structuring the network to report partial membership in multiple classes, weighted by relative concentration.



**Figure 4.** THREE LAYERED NEURAL NETWORK WITH BACK PROPAGATION.

$s\_data$ , an  $nd$ -dimensional vector, is the input. During training  $r\_class$ , an  $nr$ -dimensional vector, is also input for each  $s\_data$  vector. Output is the  $nr$ -dimensional vector  $r\_activity$ . During training the difference between  $r\_class$  and  $r\_activity$  is back propagated through the network, and the errors used to "adapt" the weights and thresholds.

$$s\_activity_{s\_unit} = \frac{1}{1 + \exp(-(s\_input_{s\_unit} - s\_thr_{s\_unit}))} \quad (15)$$

Each associative unit's **a\_input**-component is computed by summing the weighted inputs to that unit from all the sensory layer units:

$$a\_input_{a\_unit} = \sum_{s\_units} w_{s\_unit,a\_unit} s\_activity_{s\_unit} \quad (16)$$

The associative unit activities **a\_activity** are computed in a manner analogous to equation 15, and the response unit activities **r\_activity** are computed by similar forward propagation as illustrated in Figure 4. **r\_activity**, interpreted as an indicator of the class of the corresponding **s\_data**, is the answer. Operated as a classifier the algorithm stops here. During training the answer is compared with the known correct answer, and the back propagated error generates the modifications to the weights and thresholds as depicted.

Back propagation differs from forward propagation only in the direction that the network is traversed, and in the replacement of the sigmoidal activation and thresholding functions that transform **u\_inputs** into **u\_activities** (where **u** denotes **s**, **a**, or **r**) with an error propagation function defined by

$$u\_error_{u\_unit} = u\_activity_{u\_unit} \cdot (1.0 - u\_activity_{u\_unit}) \cdot u\_re-input_{u\_unit} \quad (17)$$

These parabolic functions of the activities are zero valued at zero and unity activities, and they peak in the middle. Like positive feedback on an amplifier, they force each unit's activity toward an extremum.

The corrections applied to the weights and thresholds include both a part proportional to the product of the unit's activity and back propagated error, and a "momentum" part that damps oscillation by incorporating into each modification a term that is a fraction of the previous modification to that weight or threshold. Training until each **r\_activity**-component differs by no more than a few percent from the corresponding **r\_class**-component takes a few hundred cycles through a training set of two or three **s\_data** vectors from the set of each of the six sample types.

Some formal mechanism is needed to decide what class a given **r\_activity** is reporting, since it is never the case that an **r\_activity** has exactly one unit component and five zero components. The filter that we use to automate the decision is

$$P_{r\_unit} = r\_activity_{r\_unit} \prod_{all\ r\_units \neq r\_unit} (1 - r\_activity_{r\_unit}) \quad (18)$$

so that **P** is essentially the vector of joint probabilities  $P_{r\_unit}$  that the sample is in class  $r\_unit$  and not in any other class.

Using a data set of 1900 **s\_data** vectors approximately equally divided among the six classes, where the data were collected in nineteen separate cycles of one hundred measurements each on seven separate days during a two week period, 1608 or 85% were correct when the seventeen normalized values were used as input. When the sixteen normalized differences between adjacent sensors elements were used as the input, 1814 or 96% were correct. When the thirty-three values corresponding to the union of the normalized and normalized difference values were used, 1747 or 92% were correct, essentially the average of the two separate values. Understanding this somewhat surprising effect awaits further work. Adding additional training data, particularly data selected from cycles on which performance is poor, improves performance statistics substantially. This is of course

in keeping with the well known result that perceptron-like neural networks are guaranteed to learn their training data perfectly. However we observe that it is not only the learning of the new training data that contributes to improved overall performance: there is general improvement in performance on all data in the cycles from which the new training data are taken. This illustrates the ability of machines of this type in some situations to generalize, and to aggregate classes from disjoint sets. Experimental and performance details are elaborated in Wong, Takemori, and Siegel [18], from which parts of this and the previous section are abstracted.

## Quantitation by Neural Network

When a sour tasting *HCl* solution (hydrochloric acid) is mixed with a bitter tasting *NaOH* solution (sodium hydroxide, an alkali or "base"), the *HCl* and *NaOH* react chemically, yielding a solution of the salty tasting compound *NaCl* (sodium chloride, the archetype of the "salt" class). We would thus be delighted were a taste-mimicking four response unit (*sour*, *bitter*, *salty*, and *sweet*) neural network to classify uniquely *HCl* solution in one class, *NaOH* solution in another class, and *NaCl* solution in a third class. In contrast, when the pleasantly smelling  $C_2H_5OH$  vapor (*ethanol*, the active ingredient in alcoholic beverages) is mixed with the unpleasantly smelling  $C_7H_{14}$  vapor (*heptane*, a component of gasoline), under ordinary conditions of temperature and pressure they do not react chemically, they just mix. Thus between the alternatives of a three response unit neural network with *binary* output units for the three classes *ethanol*, *heptane*, and *ethanol-heptane mixture* and a two response unit neural network with *continuous* output units for the two classes *ethanol* and *heptane* the latter would seem preferable, particularly if it were quantitatively accurate to interpret an output of  $x_e$  by the *ethanol* response unit and an output of  $x_h$  by the *heptane* response unit as meaning that the mixture is in the proportion  $x_e : x_h$  of *ethanol* vapor to *heptane* vapor. In this section I will contrast the *neural network as classifier* approach that we took to the system of *ethanol*, *methanol*, and *heptane* and their binary mixtures as a six class, binary output problem, with an alternative (and I think more realistic and more useful) *neural network as quantitator* approach wherein the aim is to use the neural network method to compute the concentrations of the components of a mixture.

First note, however, that response units that are constrained to binary output values are by no means useless for quantitation. For example, we might imagine that multiple repetitive classifications could be made by a network having a probabilistic output behavior, such that the responses of its *ethanol*-response unit and its *heptane*-response unit to a mixture of relative concentration  $x_e$  of *ethanol* and  $x_h$  of *heptane* would be 1 in fraction  $x_e$  of the trials and 0 in fraction  $1-x_e$  of the trials, and (independently) 1 in fraction  $x_h$  of the trials and 0 in fraction  $1-x_h$  of the trials respectively. Such behavior would have analogies to measurements on quantum mechanical systems, where the result of each trial is constrained to be drawn from the set of system eigenvalues [19], other analogies to "flip-flop" sensors [20],<sup>7</sup> and further analogies in to biological neural systems with characteristic amplitude-to-frequency response. But it is interesting also to consider the alternative of neural networks with multiple continuum outputs.

---

7. A "flip-flop" sensor is one in which the transducer output creates a bias that upsets the otherwise *a priori* equal probabilities that a bistable circuit will power-up in the 0-state or 1-state.

In the early days of neural networks the nodes or units were usually regarded as strictly binary, *i.e.*,

$$u_{activity_{u_{unit}}} = \begin{cases} 0 & \text{if } u_{input_{u_{unit}}} - u_{thr_{u_{unit}}} \leq 0 \\ 1 & \text{if } u_{input_{u_{unit}}} - u_{thr_{u_{unit}}} > 0 \end{cases} \quad (19)$$

in contrast to the continuous sigmoidal functionality (equation 15) that is currently preferred. The change was driven primarily by the desire to approximate the discontinuous Heaviside Function (equation 19) with a continuous function like the sigmoid to facilitate analyzing certain training and behavior models of neural networks, *e.g.*, models that employ analogies with collections of interacting atoms undergoing thermal relaxation from strained to equilibrium states. Their predilection toward pseudo-binary output is then driven by the slope-at-threshold of the sigmoid, scaled by a "sharpness" multiplier that can be inserted in the exponential, and also by the error weighting, in back propagation, by  $u_{activity} * (1 - u_{activity})$ . I have recently begun to investigate the effectiveness with which the current generation of neural networks can quantitate the response of our integrated gas sensor to mixtures, with an eye towards improving that effectiveness by the adjustment of the following options: pre-processing the input data, including an input thresholding and activation function, sharpening the sigmoidal threshold function, and changing the error weighting scheme used during back propagation. This is obviously a program with many dimensions; there is only time and space here to report the approach and preliminary findings.

A measurement of the concentrations of *ethanol*, *methanol* and *heptane* in an arbitrary mixture can be represented as a vector in a three dimensional space:  $\mathbf{c} = x_e \mathbf{e} + x_m \mathbf{m} + x_h \mathbf{h}$  where  $\mathbf{e}$ ,  $\mathbf{m}$ ,  $\mathbf{h}$  are unit vectors in the *ethanol*, *methanol*, and *heptane* directions. The data on which the work described in the previous section is based represent measurements taken in severely constrained portions of this space: single species measurements are taken in the unit vector directions  $\mathbf{e}$ ,  $\mathbf{m}$ , and  $\mathbf{h}$ , and binary mixture measurements are taken along the lines  $0.5\mathbf{e} + 0.5\mathbf{m}$ ,  $0.5\mathbf{e} + 0.5\mathbf{h}$ , and  $0.5\mathbf{m} + 0.5\mathbf{h}$ . For the numerical experiments I have also generated a simulated data set<sup>8</sup> of random binary mixtures  $x_e \mathbf{e} + (1 - x_e)\mathbf{m}$  for a simulated *ethanol-methanol* mixture, etc. The simulation populates the  $\mathbf{em}$ -,  $\mathbf{eh}$ -, and  $\mathbf{mh}$ -planes with "data."

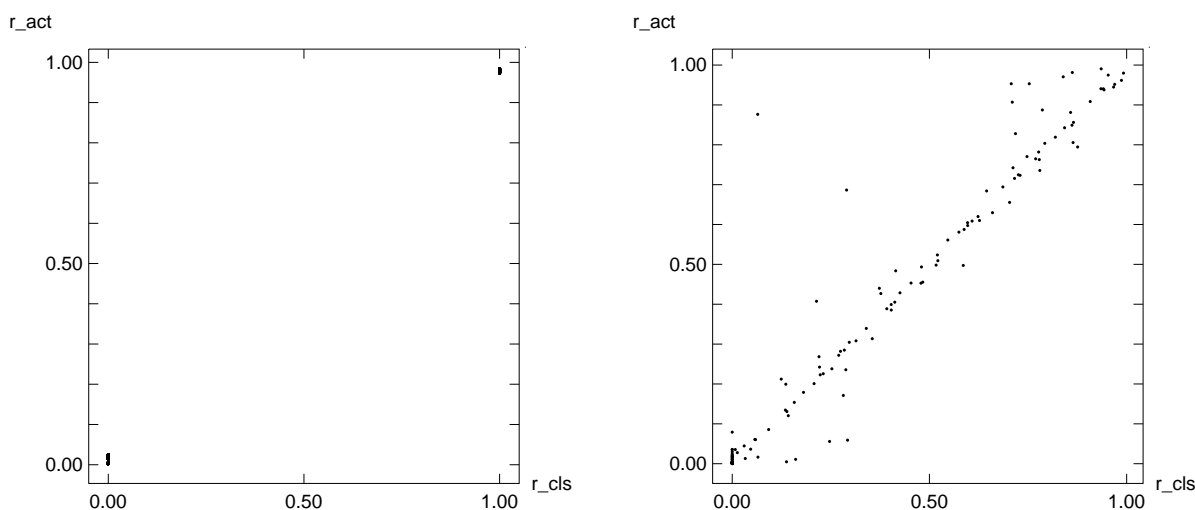
Having thus assembled a data set composed of measurements on single vapors, measurements on binary mixtures of vapors in fixed (nominally equal) proportions, and simulated binary mixtures of vapors in random proportions, I then train neural networks on subsets of the three data sets and observe the performance of these networks in classifying and quantifying the data sets from which their training data was drawn, and the other data sets too. I represent the results graphically as plots of  $r_{activity_{r_{unit}}}$  vs  $r_{class_{r_{unit}}}$  points, without distinguishing (during the present discussion) as to which (*ethanol*, *methanol*, *heptane*) unit is responsible for which point. I then investigate classification and quantitation behavior over a 3 data-set  $\times$  3 neural-network matrix.

I will first discuss the *first* matrix element, behavior of the network trained on a subset of the single vapor samples and exercised on the full data set of single vapor samples, and (at the risk of telegraphing my punchline), the *last* matrix element, behavior of the network trained on a subset of the simulated data and exercised on the full set of simulated data. The former is shown on the left side of

---

8. The combining rule I used to create simulated resistances was to add the randomly weighted *reciprocals* of the measured resistances and take the reciprocal of the result. This approach models (naively, but it is a beginning) linear superposition of response to multiple gases the by addition of *conductance* changes.

Figure 5 and the latter on the right side of Figure 5. These illustrate very well the meanings of *classification* and *quantitation* by a pair of neural networks of exactly the same topology, started in exactly the same initial state, and trained by exactly the same back propagation algorithm, the difference being that in the former case the **r\_class** vectors were limited to the unit vectors, and in the latter case the **r\_class** vectors were random two unit vector admixtures. The interesting thing about the former case is how *uninteresting* it is: 450 data points are so tightly clustered in the lower left and upper right corners that the plot looks almost empty: this net is an effective classifier. The latter case illustrates that the same paradigm can also produce an effective quantifier: 150 data points lie for the most part along the line of unit slope and zero intercept, which is the performance desired of a quantifier. The short vertical cluster in the lower left corner shows that the zero concentration components, one in each data vector, are being correctly reported as near zero. But there is still a hint of predilection toward binary behavior: when *r\_class* is less than about 0.25 or more than about 0.75 there is a tendency for *r\_activity* to be near 0 and 1 respectively.

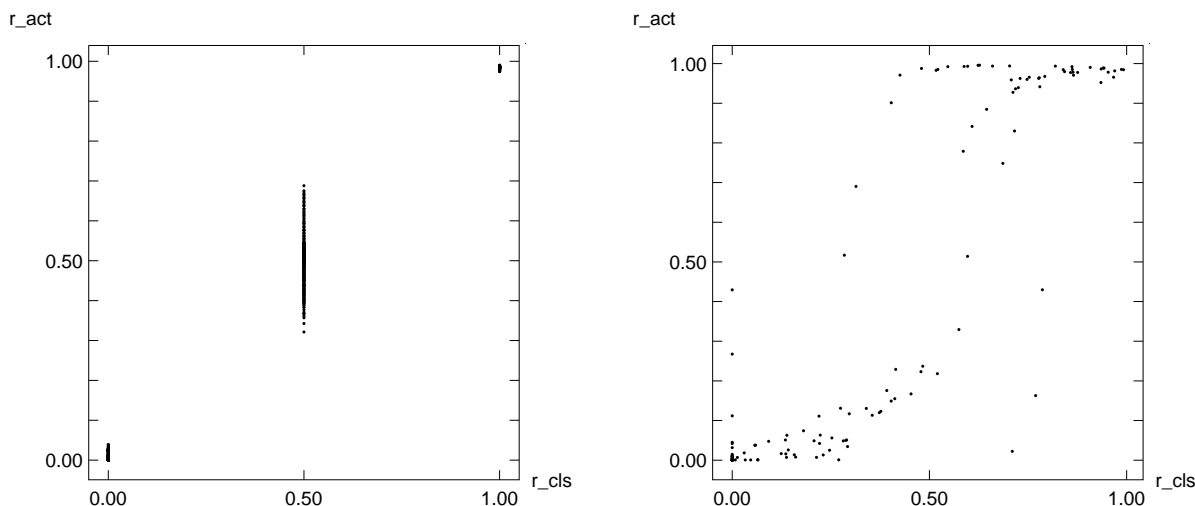


**Figure 5.** ACTIVITY vs CLASS VECTOR FOR VAPORS AND VAPOR MIXTURES

*Left:* trained on single components, exercised on single components.

*Right:* trained on random binary mixtures, exercised on random binary mixtures.

The left half of Figure 6 illustrates the behavior of the network trained on a subset of the unary and binary mixture data and exercised on a larger set of unary and binary mixture data, and the right half of Figure 6 illustrates the behavior of the same network when exercised on the simulated data set of binary mixtures of random composition. The former case shows that the network is able to classify both the unary and binary mixtures, although the distribution is much tighter for the unary than for the binary cases; nevertheless, the error rate for this network as a classifier on this data set is close to zero. The latter case is actually a very difficult test addressing the question of the extent to which reference data taken on lines in a three dimensional space can calibrate measurements made on planes containing those lines: not yet the general question of the extent to which reference data taken on lines can calibrate the volume of the space, but moving in that direction. Although the three types of points corresponding to the three binary mixture types are not distinguished in this plot, it seems clear that there are three sigmoidal curves here, each representing a gentle tendency to binarize one of the simulated mixtures according to whether its proportions are greater or smaller than the actual proportions in the real binary mixture.



**Figure 6.** ACTIVITY vs CLASS VECTOR FOR VAPORS AND VAPOR MIXTURES

*Left:* trained on unary and binary mixtures, exercised on unary and binary mixtures.

*Right:* trained on simulated binary mixtures, exercised on real unary and binary mixtures.

How to minimize the tendency toward binary response when continuum response is what is appropriate is a goal of future work. I expect it can be done by selecting network and training algorithm parameters properly. The results to date, with entirely conventional approaches to network and training algorithm parameters, are encouraging. The remaining question is the generality with which the resulting networks can represent physically appropriate nonlinear response functions, and the extent to which they can abstract those response functions with only the shadows of physical models of sensors available and only the calibration ("training") data sets which are restricted by combinatorial enormity to data located on limited subspaces of lower dimensionality.

## Symbols Used

$[X]$	Density, concentration, or partial pressure of species $X$ . It is often useful to interpret the units as being relative to $[O_2]_{air}$ , the density, concentration, or partial pressure of $O_2$ in nominal air.
$R$	Sensor's electrical resistance. [ohms]
$R_o$	Sensor's electrical resistance in pure air. [ohms]
$K_n^X$	Rate constant for $n$ -th order reaction, <i>i.e.</i> , one that depletes $[O_2]$ in proportion to $[X]^n$ . [concentration-units <sup>-n</sup> x sec <sup>-1</sup> ]
$K_{ijk\dots}^{XYZ\dots}$	Rate constant for multi-order reaction, <i>i.e.</i> , one that depletes $[O_2]$ in proportion to $[X]^i[Y]^j[Z]^k\dots$ . [concentration-units <sup>-(i+j+k+\dots)</sup> ]
$\beta$	Power law exponent describing dependence of $\frac{R}{R_o}$ on gas concentrations.
$k_i$	Equilibrium constant.
$kT$	Product of Boltzmann's constant and the absolute temperature.
$E_a$	Activation energy.



## Acknowledgements

Paul Clifford, under the mentorship of David Tuma and Angel Jordan, brought a wealth of experience and perspective on MOS sensors to my lab during its formative years. The apparatus and code were implemented in stages by Mark Licata, David Wagner, Tamal Mukherjee, Lanwai Wong, and Toshikazu Takemori. Alan Guisewite provided us all with craftsmanship and continuity. The sensors were fabricated at Oak Ridge National Laboratories by Robert Lauf and Barbara Hoffheins. Support for developing the concept, building the apparatus, and initiating the measurements at CMU, and for the fabrication program at Oak Ridge, was provided by Cabot Corporation (Boston), championed by Law McCabe. Toshikazu Takemori's participation in our research was supported by Osaka Gas Company (Japan). An unrestricted grant provided by Mettler Instrumente AG (Switzerland) on the initiative of Rudolf E. Kubli enabled me to pull it all together for the NATO workshop.

## References

1. D. Schuetzle and R. Hammerle, *Fundamentals and Applications of Chemical Sensors*, American Chemical Society, Washington, DC, 1986.
2. P. K. Clifford and D. T. Tuma, "Characteristic of Semiconductor Gas Sensors Part I: Study State Gas Response", *Sensors and Actuators*, Vol. 3, 1982-1983, pp. 233-254.
3. P. K. Clifford and D. T. Tuma, "Characteristic of Semiconductor Gas Sensors Part II: Transient Response to Temperature Change", *Sensors and Actuators*, Vol. 3, 1982-1983, pp. 255-281.
4. A. W. Laubengayer, *Reaction Rates and Chemical Equilibrium*, Rinehart & Company, New York, 1957, pp. 239-41, ch. 15.
5. I. Newton Kugelmass, editor, *Molecular Basis of Odor*, Charles C. Thomas, Springfield, IL, American Lecture Series, Vol. 773, No. 1, 1970.
6. R. C. Gesteland, J. Y. Lettvin, W. H. Pitts, "Chemical Transmission in the Nose of the Frog", *J. Physiol.*, Vol. 181, 1965, pp. 525-59.
7. J. Y. Lettvin, H. R. Maturana, W. S. McCulloch, and W. H. Pitts, "What the Frog's Eye Tells the Frog's Brain", *Proceedings of the IRE*, Vol. 47, November 1959, pp. 1940-51".
8. Jerome Y. Lettvin and Robert C. Gesteland, "Speculations on Smell", *Cold Spring Harbor Symposia on Quantitative Biology*, Vol. 30, 1965, pp. 217-25.
9. Frank Rosenblatt, *Principles of Neurodynamics*, Spartan Books, Washington DC, 1962.
10. H. D. Block, "The Perceptron: A Model for Brain Functioning", *Reviews of Modern Physics*, Vol. 34, No. 1, January 1962, pp. 123-35.
11. Anton J. Carlson and Victor Johnson, *The Spinal Cord and Brain*, University of Chicago Press, Chicago, Illinois, 1941, ch. 10.
12. B. S. Hoffheins, R. J. Lauf, M. W. Siegel, "Intelligent Thick Film Gas Sensor", *Hybrid Circuits*, Vol. 14, September 1987, pp. 8-12.
13. M. W. Siegel, R. J. Lauf, and B. S. Hoffheins, "Dual Gradient Thick-Film Metal Oxide Gas Sensors", *Proceedings of the Tokyo meeting*, Transducers '87, June 1987, pp. 599-604.
14. B. S. Hoffheins, R. J. Lauf, M. W. Siegel, "Intelligent Thick Film Gas Sensor", *Proceedings of the Atlanta Meeting*, International Society for Hybrid Microelectronics, October 1986.
15. Marvin Minsky and Seymour Pappert, *Perceptrons*, The MIT Press, Cambridge, MA, 1969.
16. David E. Rumelhard, James L. McClelland, et al, *Parallel Distributed Processing*, The MIT Press, Cambridge MA, Vol. 1, 1988.

17. David C. Plaut with Geoffrey Hinton and Steven Nowlan, "CMU Back-propagation Simulator User's Guide", Tech. report, Carnegie Mellon University, October 1987.
18. Toshikazu Takemori, Lanwai Wong, and M. W. Siegel, "Gas Identification System using Graded Temperature Sensor and Neural Net Interpretation", Technical report CMU-RI-TR-20-89, The Robotics Institute, Carnegie Mellon University, July 1989.
19. Donald H. Menzel, editor, *Quantum Theory*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1951.
20. S. Middelhoek, P. J. French, J. H. Huijsing, and W. J. Lian, "Sensors with Digital or Frequency Output", *Proceedings of the Tokyo meeting*, Transducers '87, June 1987, pp. 17-24.