REDUNDANT DEGREES OF FREEDOM IN SPEECH CONTROL: A PROBLEM OR A VIRTUE?

Peter D. Neilson, Megan D. Neilson, and Nicholas J. O'Dwyer

School of Electrical Engineering and Computer Science,
Clinical Research Unit for Anxiety Disorders, and
Department of Neurology, School of Medicine,
University of New South Wales, Sydney, Australia.

ABSTRACT - It is well-established that rapid, functionally specific compensations for unexpected perturbations occur in speech articulators remote from the site of the disturbance. We interpret this in terms of an adaptive controller which incorporates an inverse internal model of the sensory-motor system involved. By using "compliant" control, in which variables representing redundant degrees of freedom are set equal to the feedback of their actual values, sensory consequences crucial to a task can be protected from external disturbances. This subsumes the notions of coordinative structures and feedforward processes.

INTRODUCTION

In production of speech, the nervous system must coordinate movements of the multiple components of the respiratory, laryngeal and supralaryngeal systems. Very many muscles are involved, leading to the problem of motor control referred to variously as "redundancy", "degrees of freedom problem" or "motor equivalence". Basically, the problem is that the system has more muscles than it requires to produce a particular utterance. Consequently, there is more than one pattern of muscle activity that could be employed. Indeed, in a redundant system, there may be an infinite number of solutions. How does the central nervous system (CNS) choose a particular pattern of muscle activation to achieve its goal?

Kelso and colleagues have focused on this question, pursuing the idea that collective action among neuromuscular components is task-related, and that the units of control are functional groupings of muscles and joints which they call "coordinative structures" or "functional synergies" (see Kelso, 1986). One method of studying complex systems is to perturb them and observe how they recover. Kelso holds that a group of articulators can be said to behave in a unitary fashion if it can be shown that a disturbance to one member of the group is responded to by other members at sites remote from the disturbance. There is now ample evidence from studies of both speech and limb movement that compensation for a perturbation at one site can be achieved by adjusting the activity elsewhere in the system.

Much of the work on speech comes from Abbs and his colleagues (see Abbs et al., 1984; Abbs & Cole, 1986). An early study showed that when the jaw was loaded during a closure movement for a /p/, closure was always attained as a result of compensatory movements of the lips. Likewise, Abbs et al. (1984) observed both lower lip and upper lip responses when a load was applied unexpectedly to the lower lip 30 ms before the onset of EMG activity for lip closure during an utterance. This group also found that for the utterance /aba/, compensation to preserve lip closure occurred in both upper and lower lip, whereas for /afa/, which does not require upper lip movement, only lower lip compensatory responses occurred. Kelso's studies (see Kelso, 1986) include measurement of responses in lip, tongue and jaw muscles across a series of trials in which the jaw was unpredictably perturbed downward while moving upward to achieve lip closure for final /b/ during the utterance /baeb/. They observed short-latency responses (20-30 ms) in upper and lower lips, but no change in tongue muscle activity. However, the same perturbation applied during /baez/, evoked increased tongue muscle activity during /z/ frication, but no active lip response. No distortion of speech occurred during any of the perturbations, even the very first.

Similar results are found in studies of non-speech motor control (see Abbs & Cole, 1986; Kelso, 1986). As with speech, remote responses occurred only when they performed a useful function and could be altered by changing the task. The speed of these task-related responses suggests that compensation does not occur by intentional reaction time processes, but rather by some sort of automatic reflex organization. Evidence that the couplings between the articulators change depending on the utterance, and that presence or absence of the response depends on both the task and on the timing of the perturbation within the task, indicates that the organization cannot be a "hardwired" input-output reflex circuit. Kelso advocates that such phenomena are best represented theoretically in terms of dynamic control structures operating in a self-organizing system. Abbs' interpretation is in terms of predictive feedforward control. As put by Abbs et al. (1984): "The feedforward controller or translator can be conceived of as a neural model of the motor system ..."

ADAPTIVE MODEL THEORY

In keeping with Abbs, we believe that feedforward control, as well as the concept of coordinative structure, can be incorporated in a broader conceptualization of movement control which we call "Adaptive Model Theory" (Neilson et al., 1988a,b). We postulate that at least three parallel processing stages are involved in movement control: sensory analysis (SA), response planning (RP) and response execution (RE). The RP-stage requires a finite interval of time to preplan a movement as a trajectory of desired sensory consequences to achieve an immediate subgoal. Multiple parameters specify the plan in the same high level sensory feature code in which feedback from the response is experienced. Thus central images of intended and actual responses are represented in the same code and can be compared directly. It is the task of the RE-stage to translate the trajectory of desired sensory consequences into appropriately coordinated and graded motor commands to signal to the muscle control systems. Thus voluntary control proceeds as a concatenated sequence of submovements, each planned in advance by the RP-stage and executed open-loop by the RE-stage. For the actual sensory consequences to match those planned, the transformation of desired sensory consequences into motor commands (by the "controller" in the RE-stage) must be the inverse of the transformation of motor commands to sensory consequences by the "controlled system" (muscle control systems, biomechanics and external systems). In other words, the controller must behave like an INVERSE INTERNAL MODEL of the system being controlled.

The basic hypothesis of Adaptive Model Theory is that the CNS has the required neuronal circuitry to monitor the outgoing motor commands and their ensuing multiple sensory consequences, and to compute, store in memory and adaptively maintain the accuracy of sets of parameters (impulse response function weights and Wiener kernels) describing the multiple input, multiple output, dynamic, nonlinear and varying relationships between them. Once computed, the parameters can be retrieved from memory and used to control the transmission characteristics of adaptive neuronal circuits which then function as an ADAPTIVE INVERSE INTERNAL MODEL of the controlled system. For example, by computing the dynamic relationships between auditory feedback of the speech signal and kinaesthetic feedback of the changing vocal tract shape, the CNS can create an internal model of the acoustic response characteristics of its own vocal tract.

Our recent focus is on the computational processes for real-time adaptive modelling of multiple input, multiple output, dynamic, nonlinear systems with stochastic input and output signals. We have developed a distributed parallel processing circuit based on the neuronal circuitry of the cerebellum and are testing its performance using computer simulation (Neilson et al., 1988a,b). Building on theory from Lee and Schetzen (1965), we employ higher order cross correlations between input and output signals to compute self-kernels and cross-kernels describing the nonlinear transfer characteristics of the system. This method requires that the inputs be independent white noise signals, so in addition to computing higher order cross correlations, the modelling circuitry also has to prewhiten and orthogonalize the input signals. But these are exactly the computations required to generate predictions of future values of sensory signals and to compute and control the formation of functional synergies or coordinative structures!

ADAPTIVE MODELLING AND SYNERGY FORMATION:

The problem of redundancy becomes apparent if one attempts to compute an inverse internal model of a controlled system which has more inputs than it has outputs. The transfer characteristics of a multiple input, multiple output linear dynamic system can be written in the form of a matrix of transfer functions. The characteristics of the inverse model can then be obtained by computing the inverse of the transfer function matrix. since it is only possible to invert a square matrix, the inverse model of a multiple input, multiple output system can only be computed for systems with an equal number of inputs and outputs. In other words, given a system with more inputs than outputs, it is not possible to compute a unique set of input signals to produce a given set of outputs, because this requires solving a set of simultaneous equations with more unknowns than equations. In Adaptive Model Theory there are only two ways to overcome this problem. Either the number of outputs can be increased by specifying additional desired sensory consequences or the effective number of inputs can be decreased by forming motor command synergies. We contend that CNS employs both these methods.

Formation of functional synergies or coordinative structures is automatically taken care of by the modelling circuitry of Adaptive Model Theory. To illustrate, consider a system with two input signals, u1(t) and u2(t), and two output signals, y1(t) and y2(t). If we assume that the two inputs are perfectly correlated and attempt to model this system, it virtually collapses into a single input system. It is not possible to compute separately the relationship of each input with each output. Now suppose that the two inputs are not perfectly correlated, so that u2(t) contains some variation not related to u1(t). By cross correlation we can compute the dynamic relationship between u1(t) and u2(t). We can then transform u1(t) through this model to obtain the part of u2(t) correlated with u1(t). Subtracting this from u2(t) gives u2'(t), the part of u2(t) not correlated with u1(t). In essence, we have ORTHOGONALIZED the two input signals by a process similar to Gramm-Schmidt orthonormalization. The orthogonalized inputs, u1(t) and u2'(t), can now be cross correlated with each output to identify a two-input, two-output system. This procedure generalizes to an n-input system where the n inputs are reduced to m orthogonal inputs equalling the number of degrees of freedom. Each of the m orthogonal input signals can then be cross correlated with each output signal to identify a virtual m-input system. The networks which orthogonalize the input signals form an essential part of the modelling circuitry. Once established, they can be used in reverse to transform the orthogonalized signals back to n interrelated signals. This is exactly the operation required of a SYNERGY GENERATOR.

Reducing the number of input signals to the number of degrees of freedom in those signals may or may not remove the redundancy from the system. If the reduced number of inputs still exceeds the number of outputs, the effective number of output signals must be increased by specifying additional desired sensory consequences for the task. For example, if the desired consequence involves a speech sound requiring linguo-palatal contact, there are infinite combinations of tongue height and jaw opening positions which could achieve this goal. But if a certain jaw opening is also specified, the required articulator positions are uniquely determined. In a sense, the additional desired sensory consequence can be regarded as redundant in that it is not crucial to the primary goal of achieving linguo-palatal contact. Nevertheless, using computer simulation, we can show that such redundant degrees of freedom are a virtue rather than a problem because they serve to defend a primary sensory consequence against external disturbance.

COMPUTER SIMULATION

While simulation cannot verify a theory of CNS function it can show that a proposal is technically possible. The following is not an attempt to emulate actual neural circuitry but uses the PC-MATLAB program package to simulate a two-input, two-output linear dynamic system with interactions between the inputs and outputs, as shown in Figure 1. Each transfer function, Hij, was arbitrarily selected as a first order system with time constants between 2 and 10 sec. We used determinants to solve the simultaneous equations and derive the transfer functions Gij for each block in an INVERSE MODEL of the system, also shown in Figure 1. Stochastic signals, u1(t), u2(t) and u3(t) were generated by filtering random numbers through second order filters. The operation of the system was then simulated by applying u1(t) and u2(t) at the inputs and examining the outputs y1(t) and y2(t). Results are shown in Figure 2 and demonstrate that the inverse model compensates for the dynamics and decouples the system. Despite the long time constants and interactions within the plant, output y1(t) tracks input u1(t) and output y2(t) tracks input u2(t), almost exactly. The very small lags between inputs and outputs in Figure 2 can be attributed to the influence of the additional high frequency time constants added to the inverse model. Such lags are inevitable in any real system. The simulation shows that an inverse internal model of a multiple input, multiple output controlled system can be used in series with that system to produce a set of output signals which recreate the inputs. In the context of Adaptive Model Theory this shows the technical feasibility of employing an inverse model to transform preplanned trajectories of desired sensory consequences into appropriately graded and coordinated motor commands to drive the controlled system to produce sensory consequences which match the desired ones.

Our final step is to simulate what we call a "compliant controller". Consider the configuration shown in Figure 1 and suppose that the primary goal is for y1(t) to follow u1(t), while y2(t) is a redundant sensory consequence signal of no special interest in the primary task. Further, suppose that y2(t) is fedback and connected as an input to the inverse model, as illustrated in Figure 3. This creates a positive feedback loop on to y2(t) which can be opened by disconnecting point X, as shown in Figure 3. The resulting configuration can be interpreted as follows: The desired value of y2(t) (applied at the input) is set, by feedback action, to equal the value it actually has at the output. We call this "compliant" because the desired value of y2(t) is simply required to track its actual value. This can be very fast because the RP-stage, which requires a finite interval of time to preplan, does not have to compute a desired trajectory for y2(t) but simply sets it equal to its feedback value. The advantage of the configuration is that any disturbance applied to the plant will be fedback via y2(t) to the inverse internal model which will apply a fast correction, via its cross coupling pathways, to cancel the influence of the disturbance on y1(t). Thus the sensory consequence y1(t) being controlled will be protected against external disturbances by the compliant redundant y2(t). The operation of the compliant controller was simulated by applying u1(t) as before and using u3(t) as a disturbance signal applied to the input of the plant as shown in Figure 3. The results of the simulation are presented in Figure 4 which shows that the output y1(t) accurately tracks input u1(t) and the influence of the disturbance u3(t) on y1(t) has been cancelled by the feedforward action of the compliant controller.

CONCLUSION

We interpret the phenomenon of compliant control simulated above as subsuming the notions of coordinative structures and feedforward processes in sensory-motor systems with redundant degrees of freedom. The speech production system contains such redundancy and the task-specific compensations observed in that system can be accounted for by compliant control operating via adaptive inverse models.

REFERENCES

Abbs, J.H., Cole, K.J. (1986) "Neural Mechanisms of Motor Equivalence and Goal Achievement", SMCL Preprints (University of Wisconsin: Madison).

Abbs, J.H., Gracco, V.L., Cole, K.J. (1984) "Control of Multimovement Coordination: Sensorimotor Mechanisms in Speech Motor Programming", J. Motor Behav. 16, 195-231.

Kelso, J.A.S. (1986) "Pattern Formation in Speech and Limb Movements Involving Many Degrees of Freedom", Exp. Brain Research Series 15 (Springer-Verlag: Berlin)

Lee, Y.W., Schetzen, M. (1965) "Measurement of the Wiener Kernels of a Nonlinear System by Cross-correlation", Int. J. Control 2, 237-254.

Neilson, P.D., O'Dwyer, N.J., Neilson, M.D. (1988a) "Central Processes Underlying the Movement Disorders of Cerebral Palsy: A Computational Model of Brain Function", J. Theor. Biol. (in press).

Neilson, P.D., Neilson, M.D., O'Dwyer, N.J. (1988b) "Intermittency and Internal Models: A Theoretical Account of Human Tracking Behavior", Biol. Cybern. 58, 101-112.

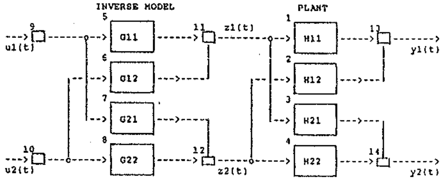

Figure 1.
Two-input two-output controlled system decoupled by an inverse model controller.

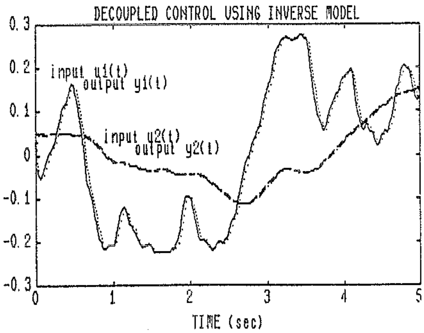

Figure 2.
Simulation of decoupled control using inverse model.

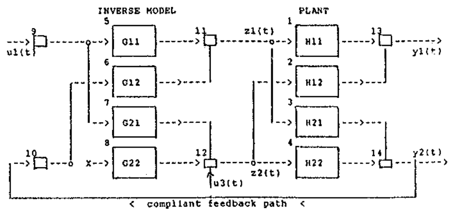

Figure 3.
Compliant controller.

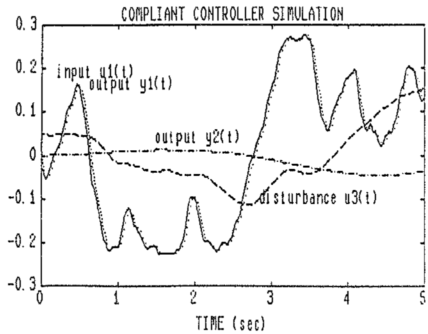

Figure 4.
Simulation of compliant controller.