Neural Networks

Introduction to Cognitive Science

Representations and Information Processing

- On the view of computationalism, cognition is information-processing of some kind. But what kind?
- In particular, how is information represented, so that information can be:
  - Stored (if important)
  - Recalled (if relevant)
  - Integrated (if helpful)
  - Updated (if needed)
    - And all that quickly and efficiently!

The Jets and the Sharks

Name	Gang	Age	Education	Marital status	Occupation
Art	Jets	40's	J.H.	Sing.	Pusher
Al	Jets	30's	J.H.	Mar.	Burglar
Sam	Jets	20's	COL.	Sing.	Bookie
Clyde	Jets	40's	J.H.	Sing.	Bookie
Mike	Jets	30's	J.H.	Sing.	Bookie
Jim	Jets	20's	J.H.	Div.	Burglar
Greg	Jets	20's	H.S.	Mar.	Pusher
John	Jets	20's	J.H.	Mar.	Burglar
Doug	Jets	30's	H.S.	Sing.	Bookie
Lance	Jets	20's	J.H.	Mar.	Burglar
George	Jets	20's	J.H.	Div.	Burglar
Pete	Jets	20's	H.S.	Sing.	Bookie
Fred	Jets	20's	H.S.	Sing.	Pusher
Gene	Jets	20's	COL.	Sing.	Pusher
Ralph	Jets	30's	J.H.	Sing.	Pusher
Phil	Sharks	30's	COL.	Mar.	Pusher
Ike	Sharks	30's	J.H.	Sing.	Bookie
Nick	Sharks	30's	H.S.	Sing.	Pusher
Don	Sharks	30's	COL.	Mar.	Burglar
Ned	Sharks	30's	COL.	Mar.	Bookie
Karl	Sharks	40's	H.S.	Mar.	Bookie
Ken	Sharks	20's	H.S.	Sing.	Burglar
Earl	Sharks	40's	H.S.	Mar.	Burglar
Rick	Sharks	30's	H.S.	Div.	Burglar
Ol	Sharks	30's	COL.	Mar.	Pusher
Neal	Sharks	30's	H.S.	Sing.	Bookie
Dave	Sharks	30's	H.S.	Div.	Pusher

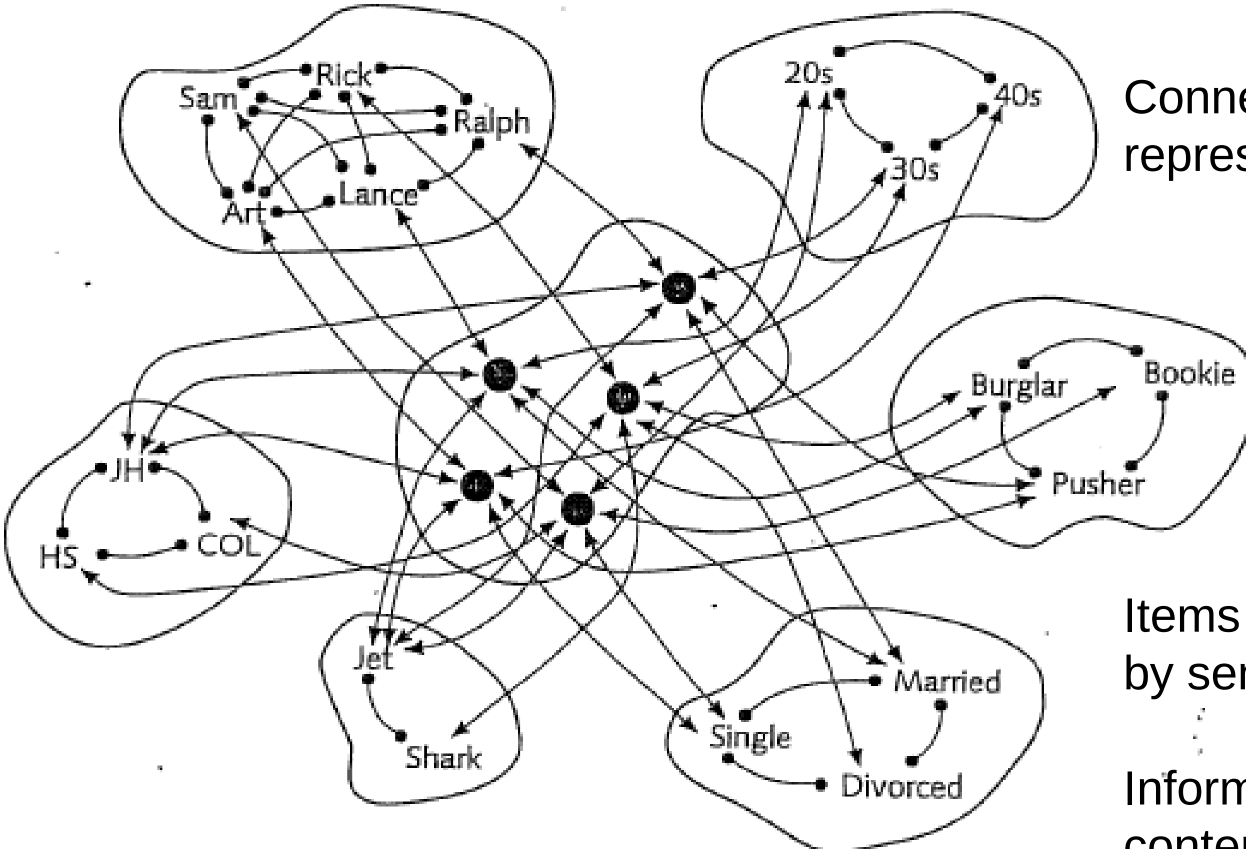
Figure 2.1 Information about individual members of two gangs, which is encoded in McClelland's (1981) Jets and Sharks network. From J. L. McClelland (1981) Retrieving general and specific knowledge from stored knowledge of specifics, *Proceedings of the Third Annual Conference of the Cognitive Science Society*. Copyright 1981 by J. L. McClelland. Reprinted by permission of author.

'Classical' representation: Information is organized like a filing system: item by item.

Ordering and indexing helps to access information

Still, how would you find the "Name of that Shark who's in his 40's, has a H.S. education, is Single, and is a Burglar"?

Connectionist (‘distributed’) representation



Items are grouped (linked) by semantical associations

Information is accessed by content

Figure 2.2 McClelland's (1981) Jets and Sharks network. Each gang member is represented by one person unit (center) that is connected to the appropriate name and property units. Only 5 of the 27 individuals from figure 2.1 are included in this illustration. Adapted (with corrections) from J. L. McClelland (1981) Retrieving general and specific knowledge from stored knowledge of specifics, *Proceedings of the Third Annual Conference of the Cognitive Science Society*. Copyright 1981 by J. L. McClelland.

What are Neural Networks?

- Neural networks are architectures inspired by the brain.
- Neural networks consist of:
  - Nodes (the ‘neurons’)
  - Connections between nodes (the ‘axons’/’dendrites’)
- Moreover:
  - Nodes have a certain ‘activation’
    - In some networks, it is simply ‘on’ or ‘off’
    - In other networks the activation can take on any value, i.e. nodes can be more or less active
  - Connections have a certain ‘weight’:
    - A positive weight means an excitatory connection; negative is inhibitory
    - In some networks, the weight is simply either ‘+1’ or ‘-1’
    - In other networks the weights can take on any value

How do they work?

- Through the connections, activations, and weights, nodes will activate or deactivate other nodes:
  - If two nodes are connected by an excitatory connection, then they will try and activate each other
    - E.g. the Ken node is positively connected with the ‘Ken’ node, the Sharks node, the 20’s node, etc.
  - If the connection is inhibitory, the nodes will try and deactivate each other
    - E.g. the 20’s node is negatively connected with the 30’s node.
  - The more active a node is, the more it will activate (or deactivate) the node it is connected to
  - Similarly, the bigger the weight, the greater the effect

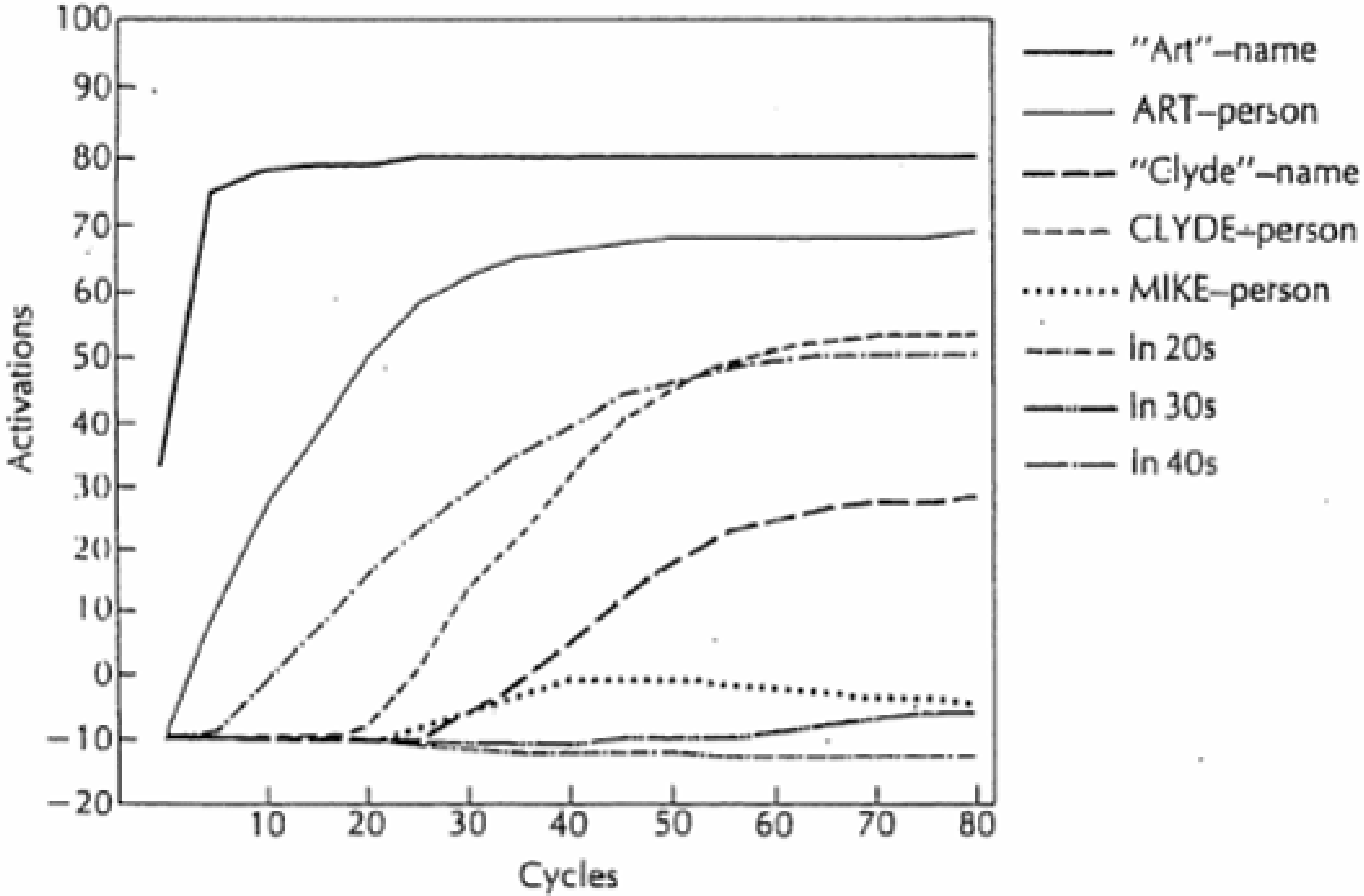


Figure 2.3 The activation values across cycles of some of the units in the Jets and Sharks network after the unit for Art's name is activated by an external input.

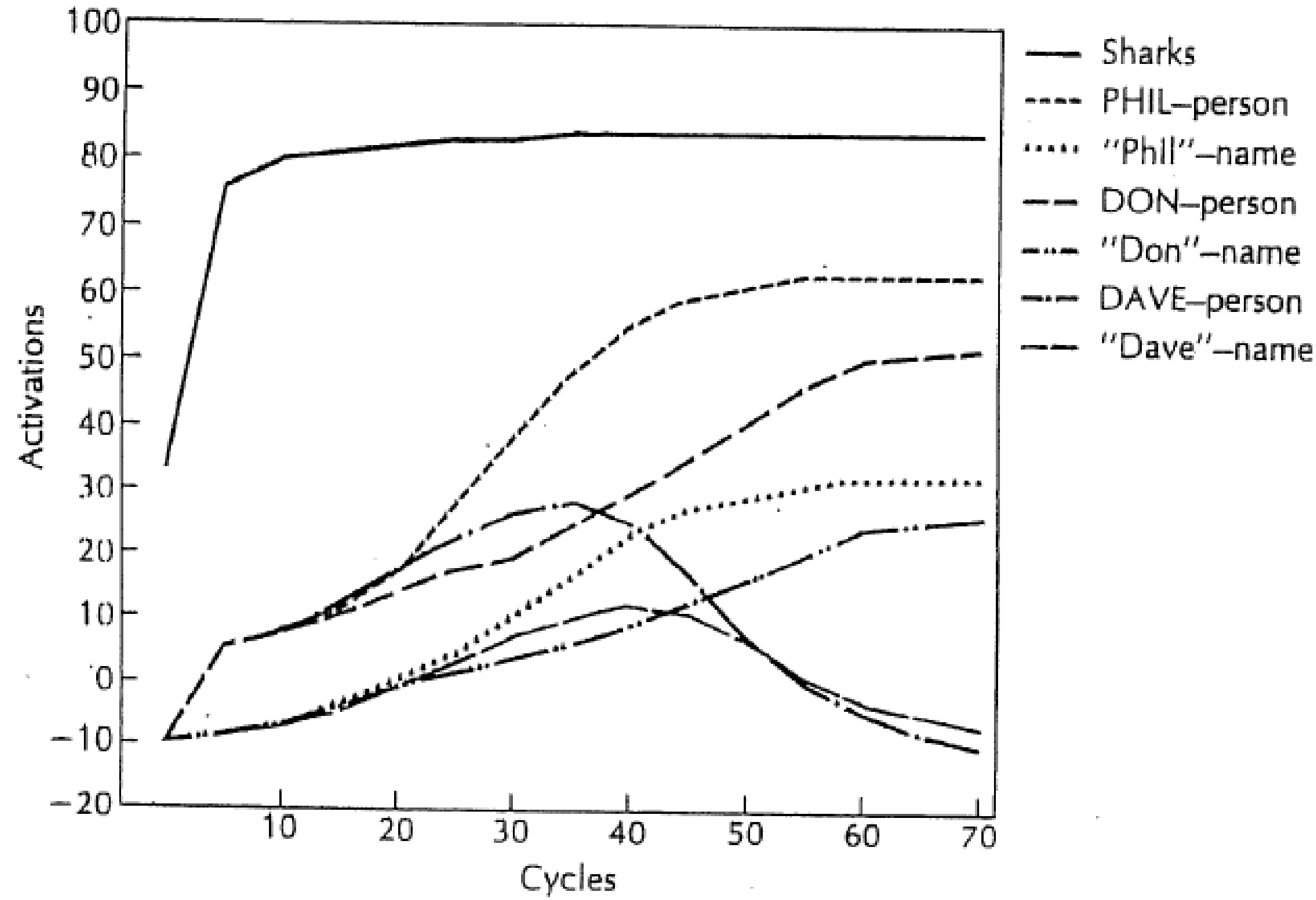
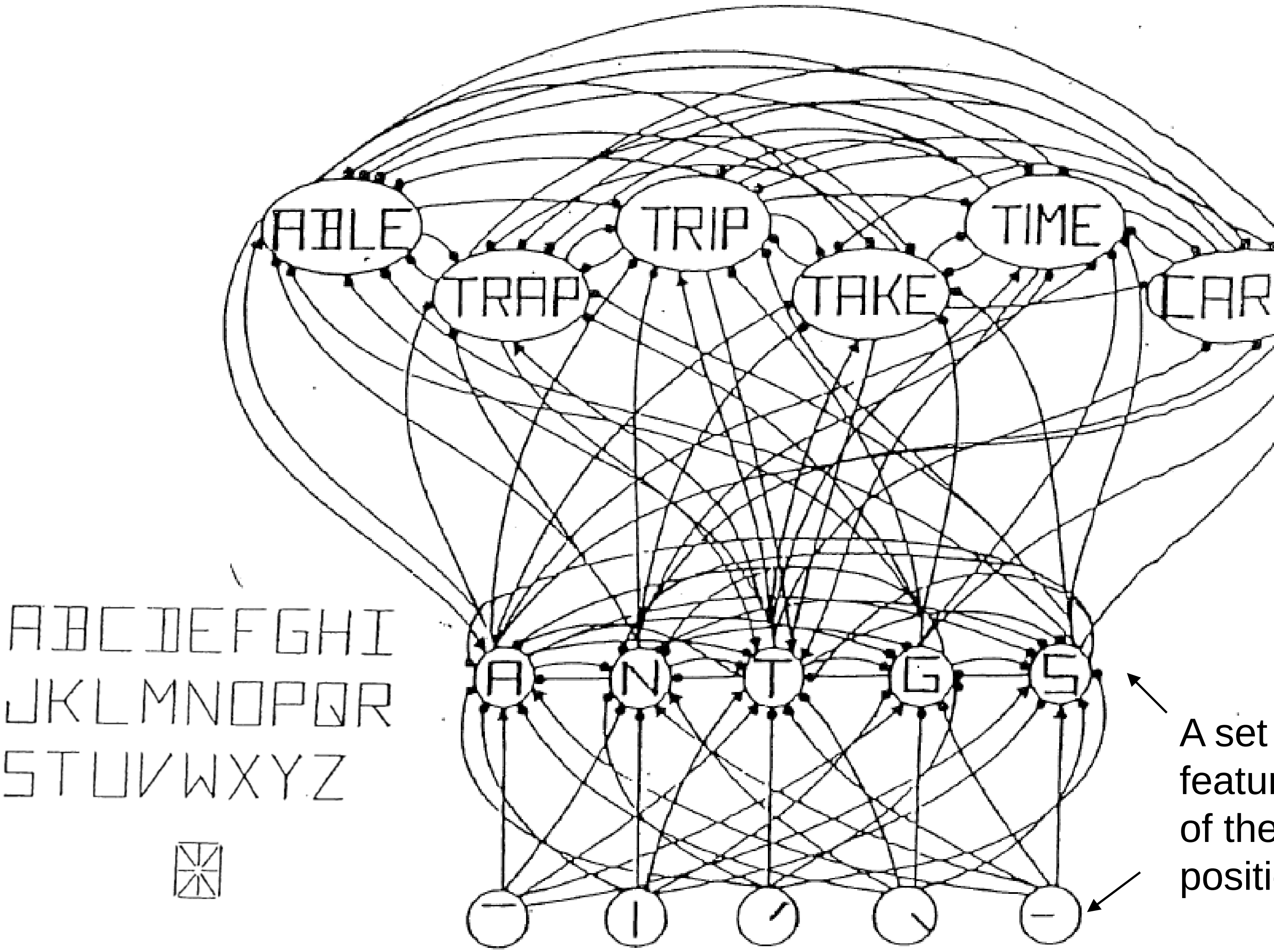
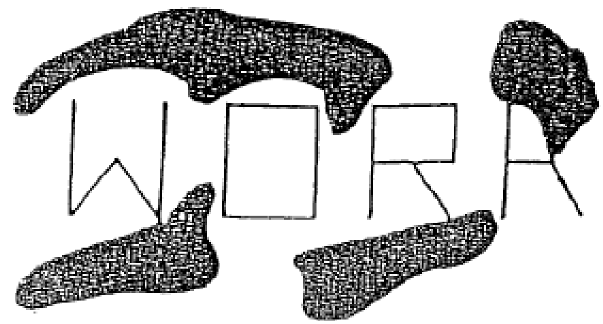


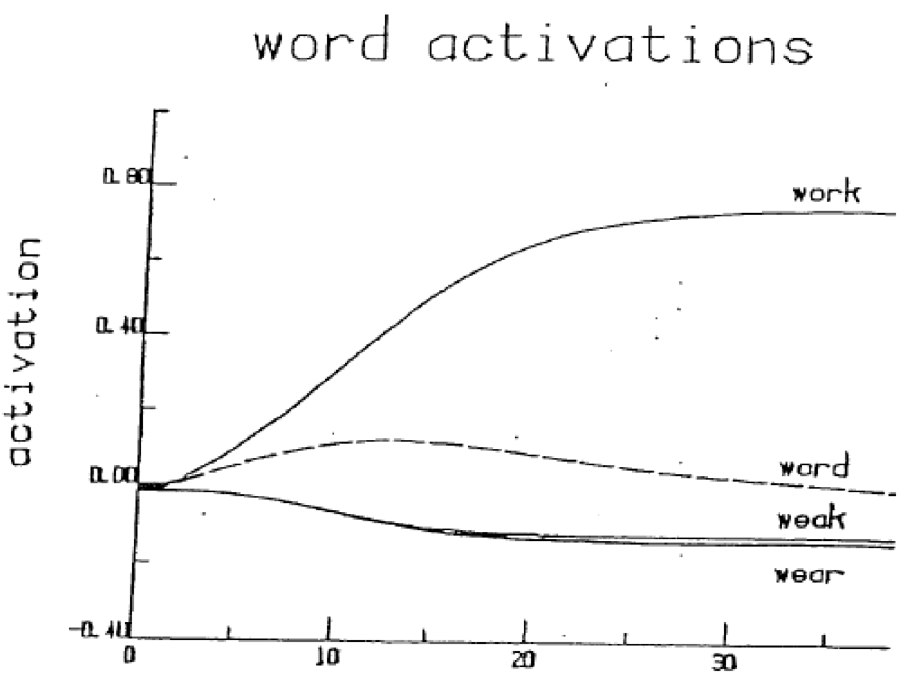
Figure 2.4 The activation values across cycles for name and person units of various members of the Sharks after the property unit for Shark is activated by an external input. The name units become less active than the person units since the name units receive activation only via the person units.

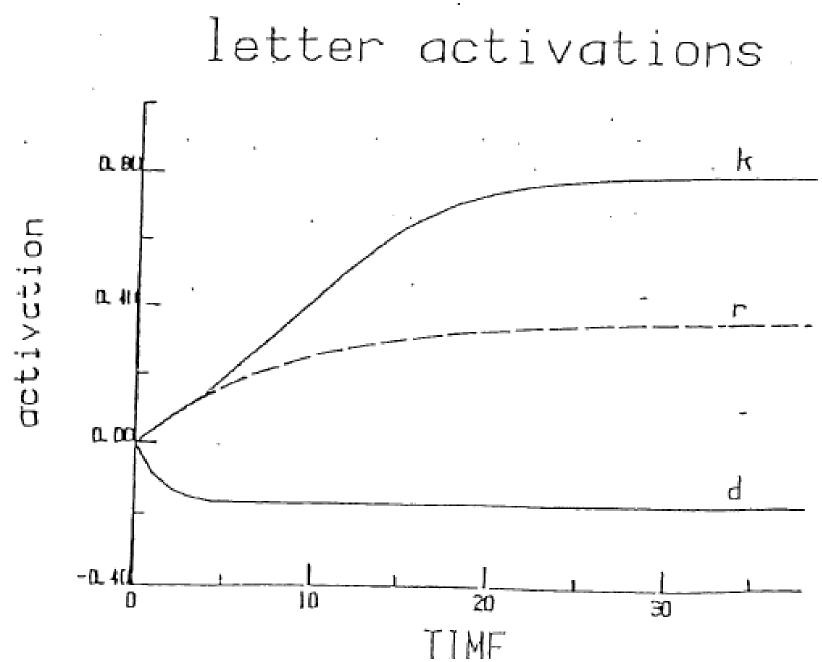
Word/Letter Perception Network, by McClelland and Rumelhart

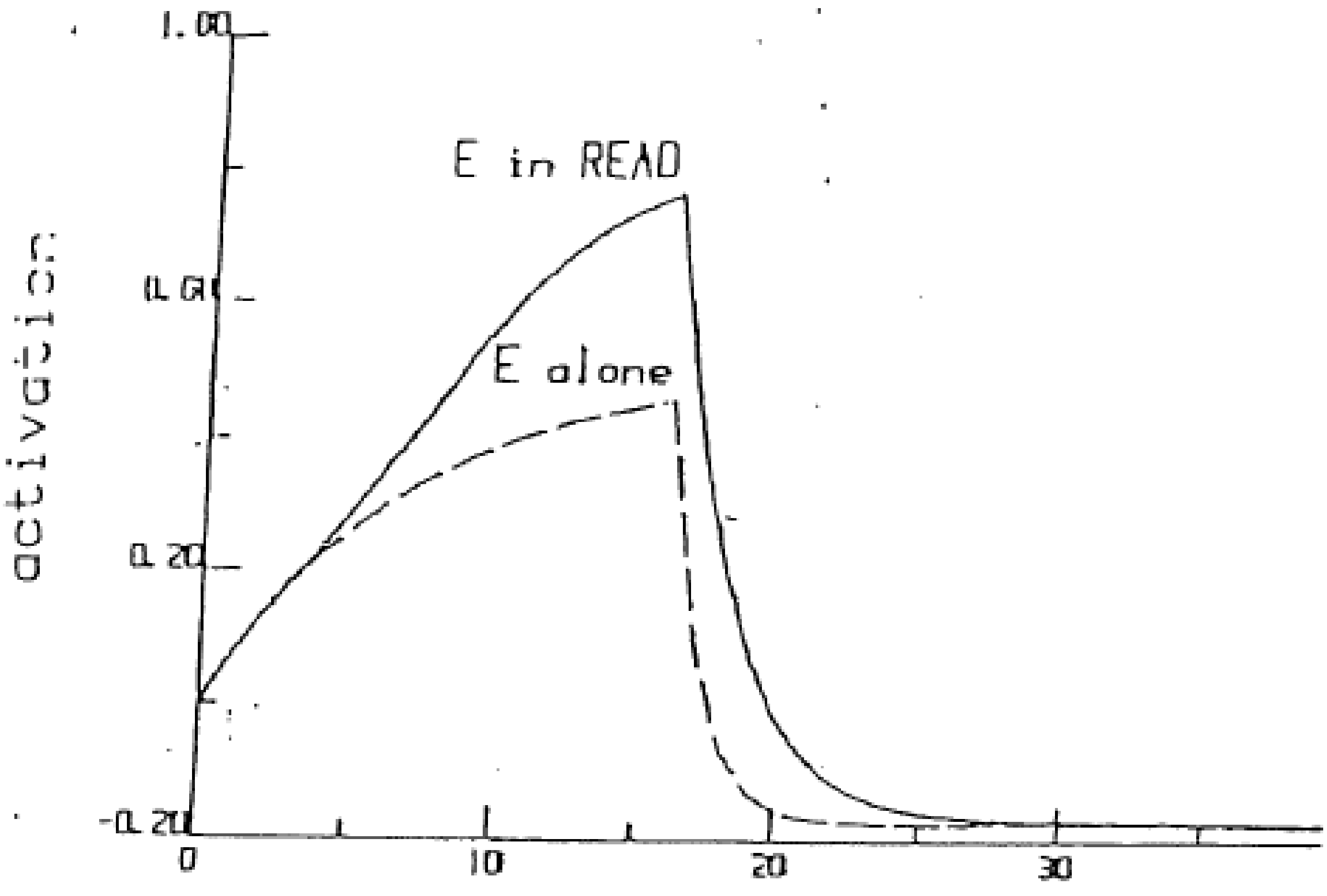


A set of letters and features for each of the four letter positions




word activations
activation
0.80
0.40
0.00
-0.40
0
10
20
30
work
word
weak
wear


letter activations
activation
0.80
0.40
0.00
-0.40
0
10
20
30
TIME
k
r
d



Word Advantage Effect: Letters are more quickly recognized in the context of a word that when presented in isolation

This is a human cognitive phenomenon, and the network can explain it

Neural Networks and Cognitive Science

- Neural Networks have several features that are of interest to cognitive science:
  - Neural plausibility
    - Neural Networks have certain features in common with real brains
  - Decentralized and Parallel processing
    - Like our brain, processing in different parts at the same time
  - Pattern Completion and Correction
    - Networks can handle incomplete or even incorrect input
  - Automatic Categorization and Prototyping
    - Categories ‘emerge’ from dynamics of the network. Moreover, some instances are more prototypical of a category than others
  - Graceful Degradation
    - Changes/damage to the network only gradually effects performance
  - Fuzzy boundary between declarative (‘know-that’) and procedural (‘know-how’) knowledge
    - No clear distinction between data and processing
  - Induction/Learning
    - Networks can learn, and are able to pick up on statistical regularities

Neural Networks and Computation

- Do neural networks fit into the scheme of computationalism?
- Well, it doesn't follow the classical computational scheme/architecture (see next slide).
- However, like other computational systems, features such as physical size, location, or matter are irrelevant to the implementation of a neural network;
  - it is only the abstract causal organization that matters
- Indeed, neural networks can be implemented on your laptop, so they are computational systems.
- Moreover, it can be shown that the computational powers of a neural network equal that of more traditional computational systems.

Neural Networks Computation vs ‘Classical’ Computation

- Some important differences:
  - In classical systems computation is done serially (one instruction at a time) in a central location (CPU) but in neural networks ‘computation’ is done in parallel and is decentralized
  - The representations in classical systems are ‘symbolic’ (natural-language sized, such as ‘cat’, ‘on’, and ‘mat’), whereas in neural networks they are ‘subsymbolic’ (activations/weights represent low-level features that are hard to put into words, if at all)
  - In classical systems there is a clear separation between data/symbols/representations on the one hand, and algorithms/processes defined over those objects on the other hand. But in a neural network this distinction is far less clear, if existent at all.
  - And finally, it can be hard to make sense of the ‘computations’ of a neural network: even if a network successfully accomplishes some task, it can be hard to explain how exactly that happened
    - (this problem obviously mirrors the problem of relating neuroscience to cognitive science)

Creating Neural Networks

- The fact that it is so hard to understand how neural networks accomplish a certain task means that it is hard to create a network that accomplishes some task:
  - How many nodes to use?
  - How to connect them?
  - And what weights should be put on these connections?

Neural Networks
Basic Learning Algorithm

- There are learning algorithms for neural networks that roughly work as follows:
  - Start with random weights
  - And now keep doing the following until you get desired performance:
    - Present an input to the network
    - The difference between the actual output and the desired output is the error
    - Figure out how each weight is contributing to the error (you can’t do this in absolute terms, but you can figure out if you should increase or decrease the weight in order to decrease the error)
    - Adjust the weights accordingly

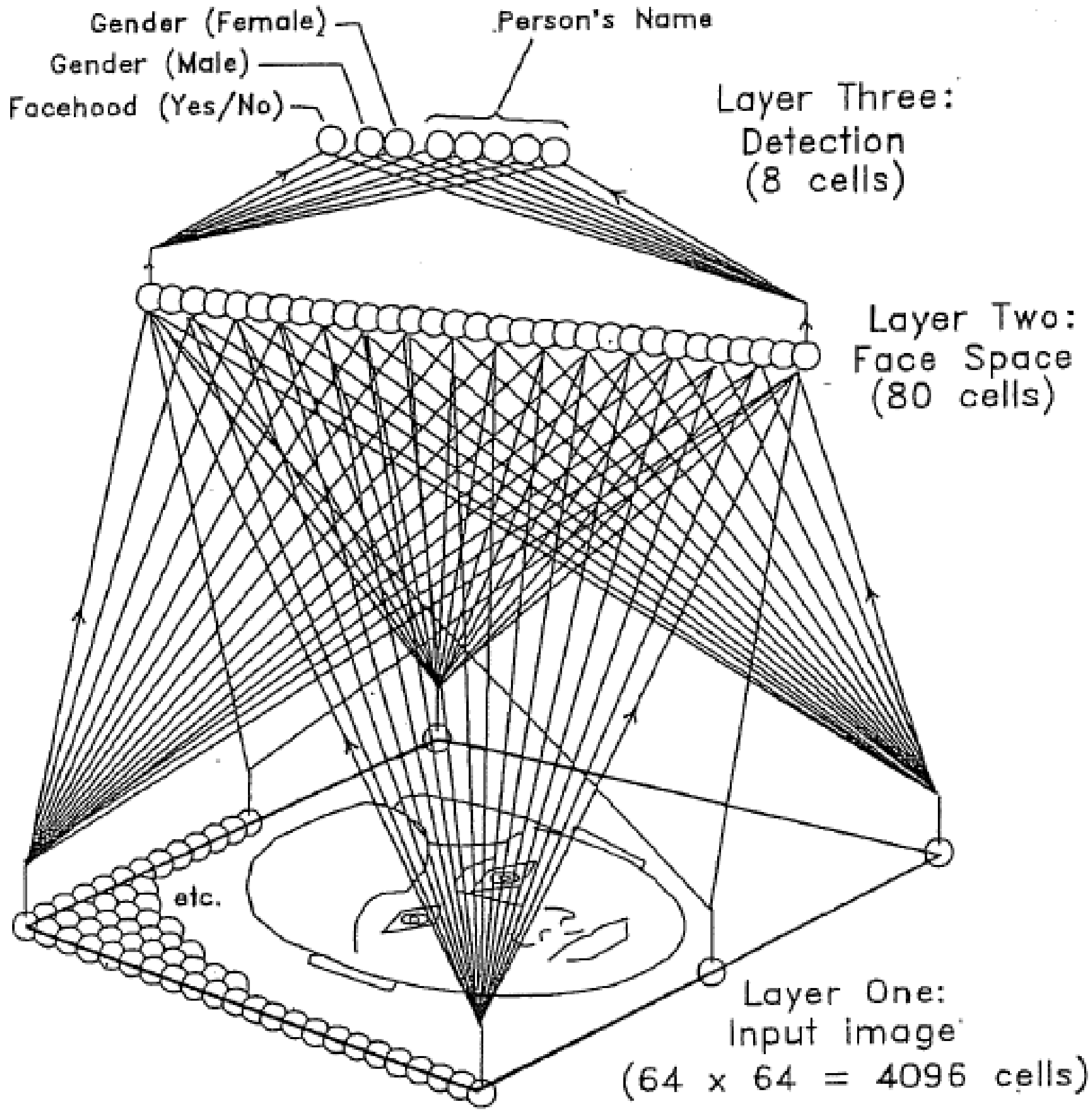
Generalization from Cases

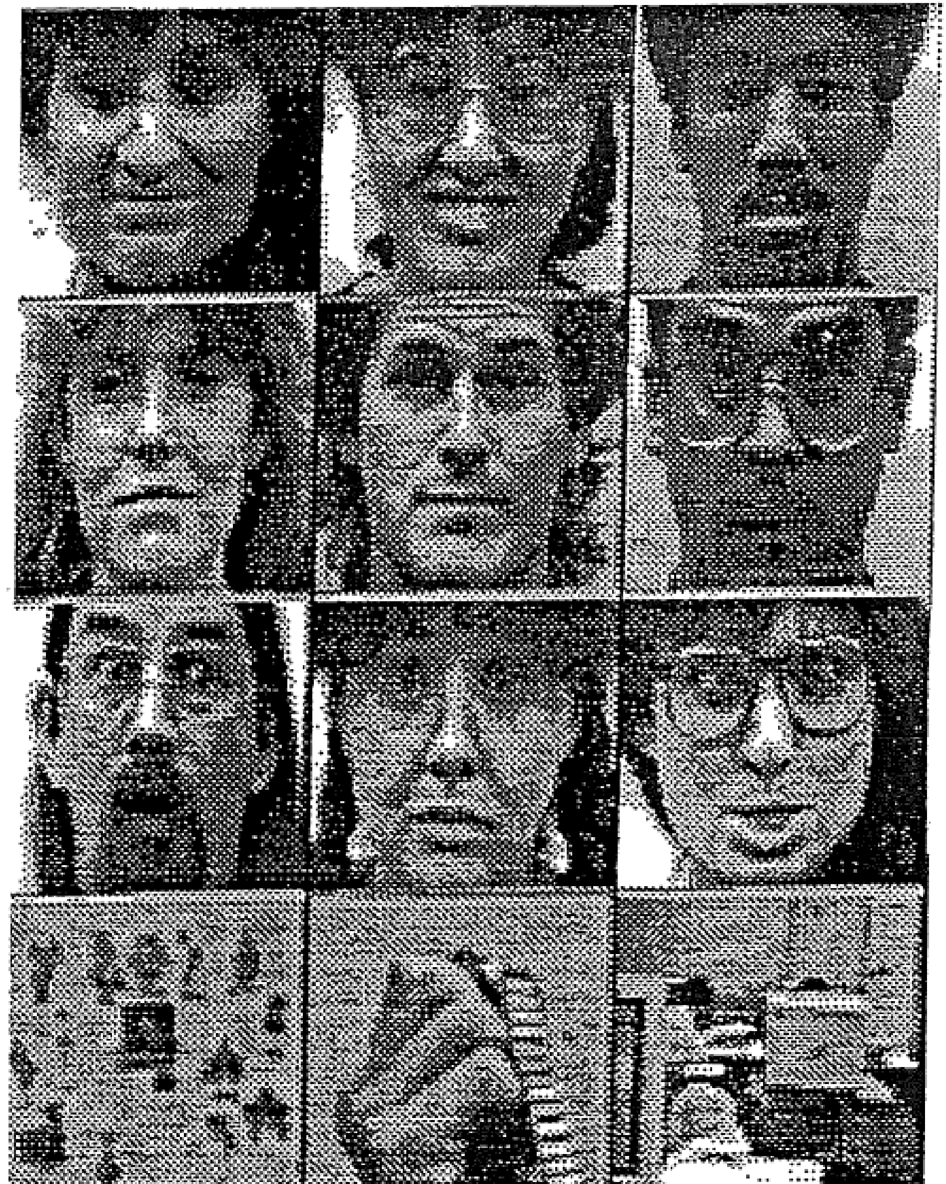
- The interesting thing is that neural networks can generalize from their ‘training set’ to never before seen cases. E.g. you can train it to recognize ‘Bob’ in certain pictures of ‘Bob’, and when you then give it a new picture of ‘Bob’, it can recognize that as ‘Bob’ as well.

Some Neural Network Examples

- Neural Networks have been trained to:
  - Recognize faces
  - Character recognition
    - (the U.S. postal service uses neural network-based software that can read hand-written Zip codes with 99% accuracy)
  - Translate written text into speech (Net Talk)
  - Control power plants
  - Control robotic limbs
  - Predict behavior of aircraft in response to pilot's commands (used in fly-by-wire jets)
  - …

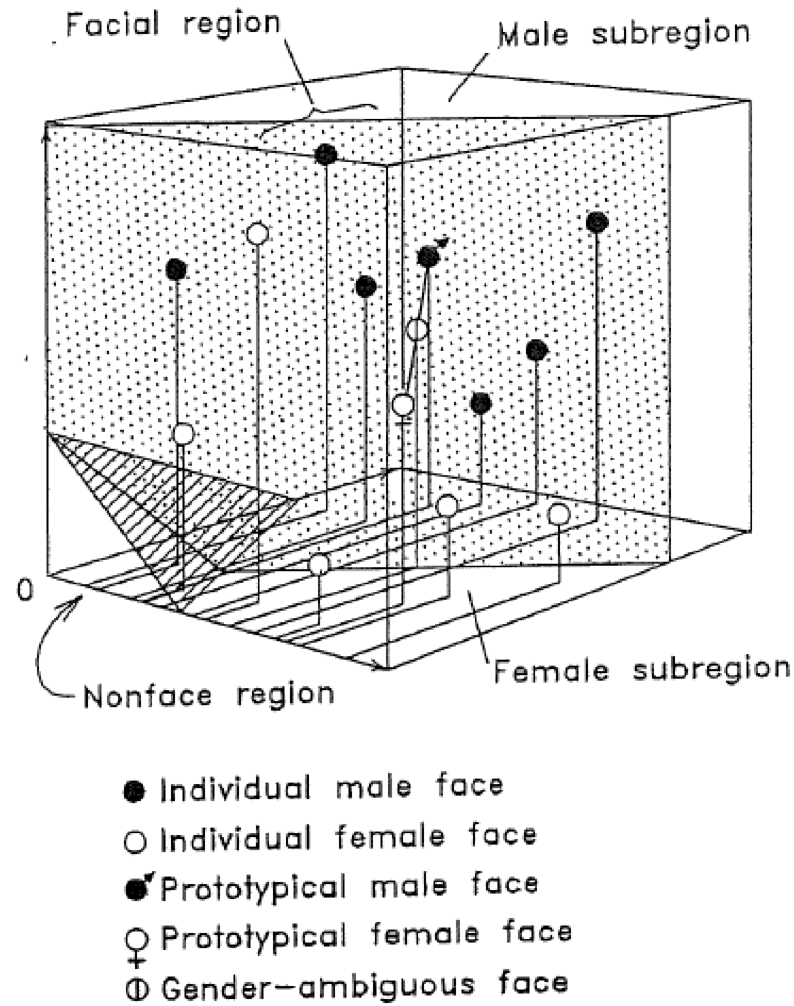
Face Recognition Network by Gary Cottrell

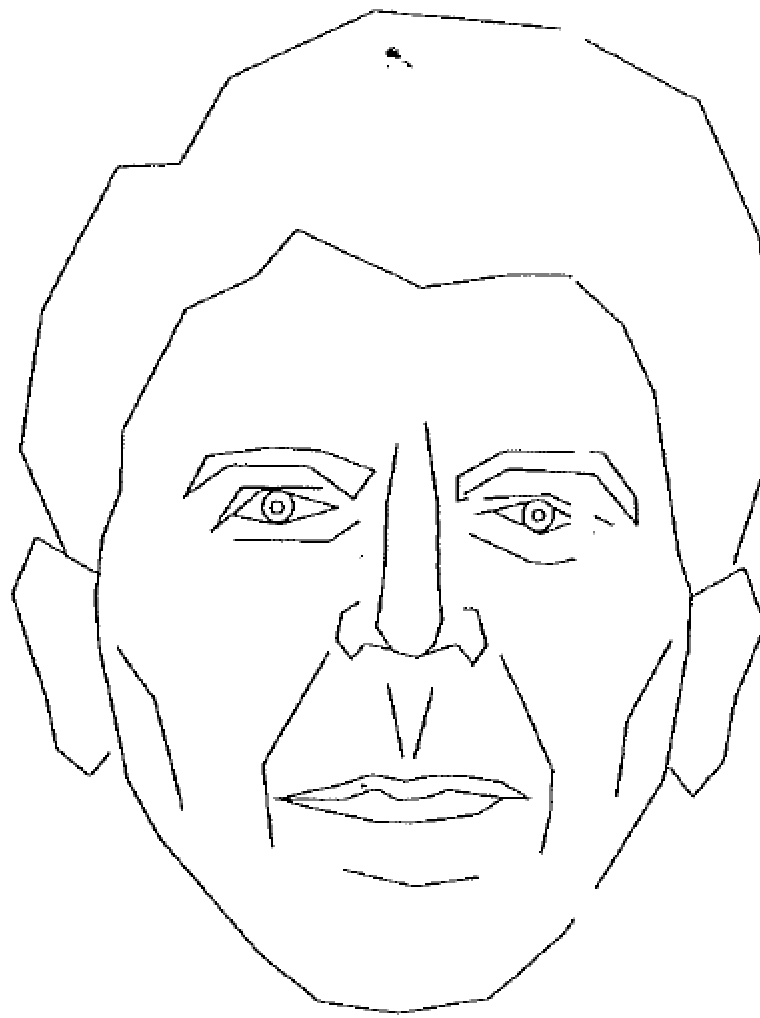


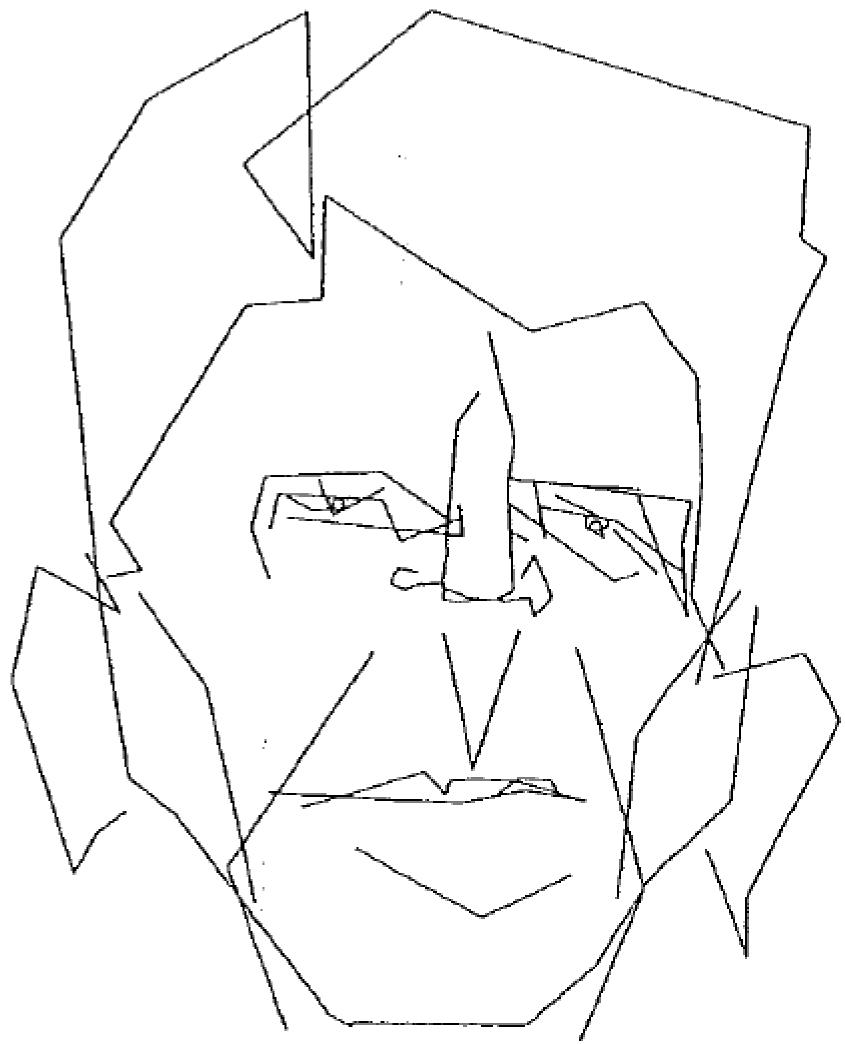


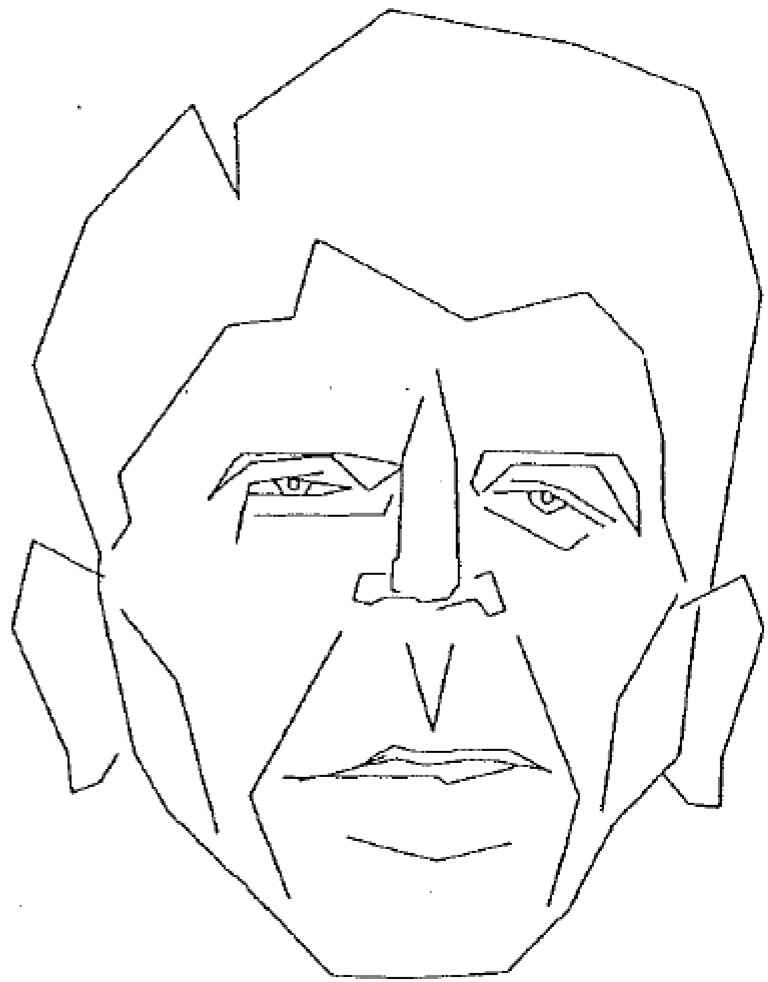
Face Recognition Network Performance

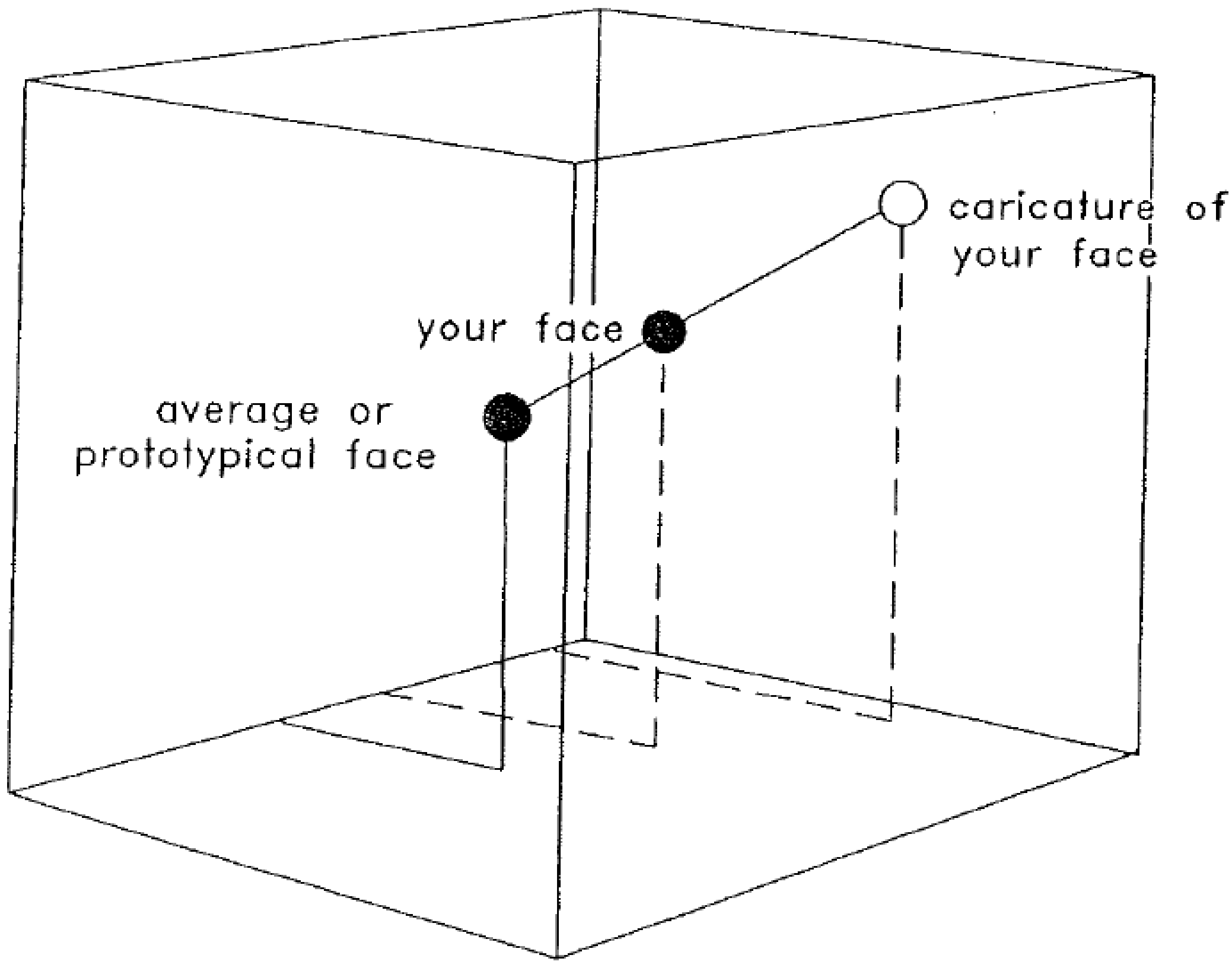
- Achieved 100% accuracy on training set
- Of testing set (set of never before seen of pictures):
  - 100% correct in determining face vs non-face
  - 98% correct in determining name of familiar face (i.e. face of someone who had a different picture in training set)
  - 81% correct in determining gender of unfamiliar face
  - Over 70% correct in determining name of familiar face when 1/5 of picture was blackened

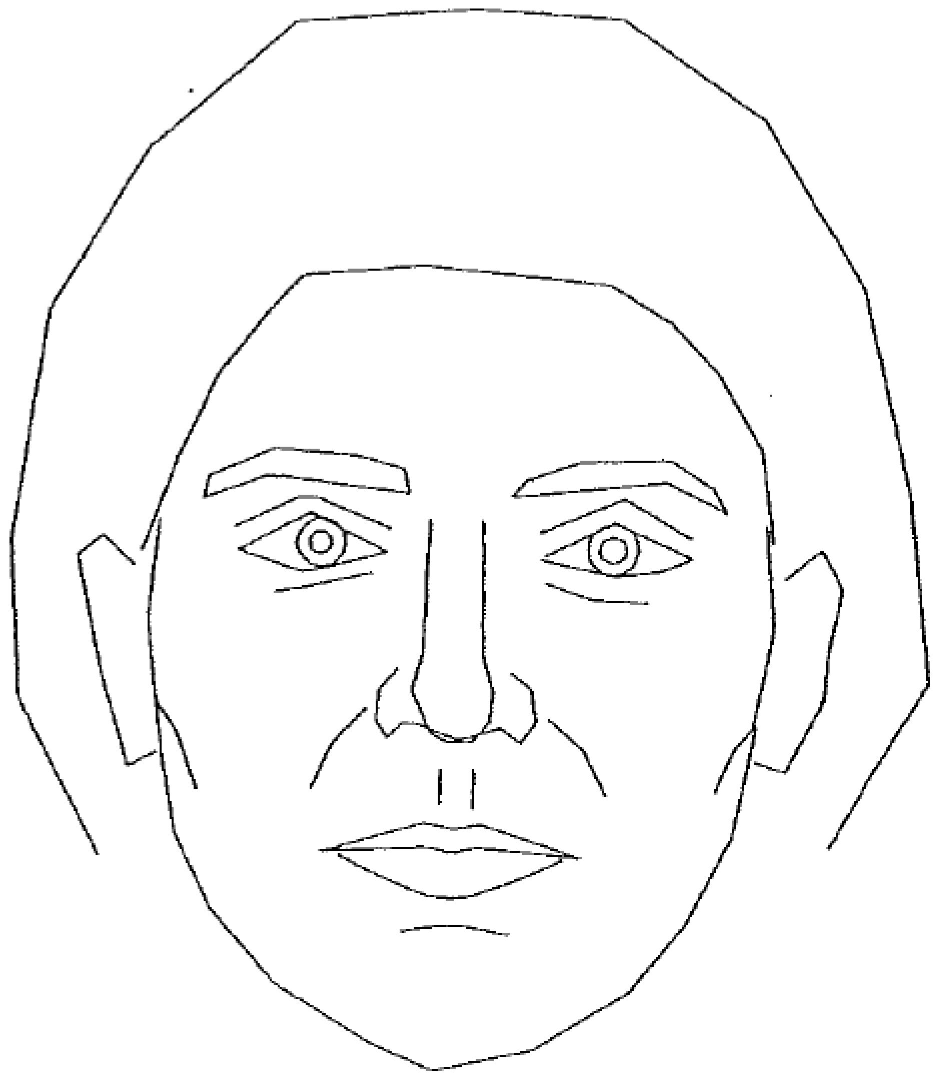

Facial region
Male subregion
0
Female subregion
Nonface region
Individual male face
Individual female face
Prototypical male face
Prototypical female face
Gender-ambiguous face

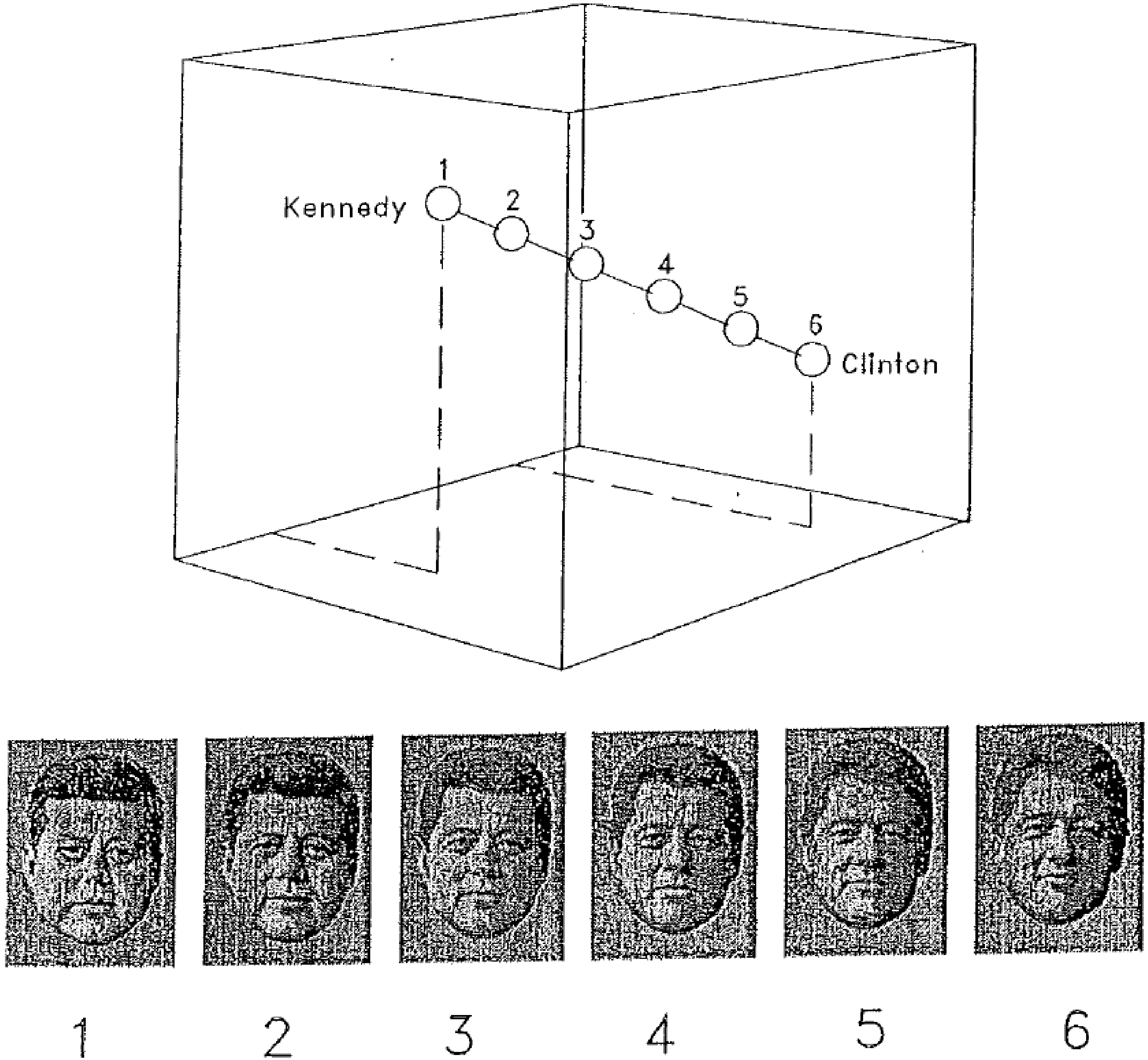


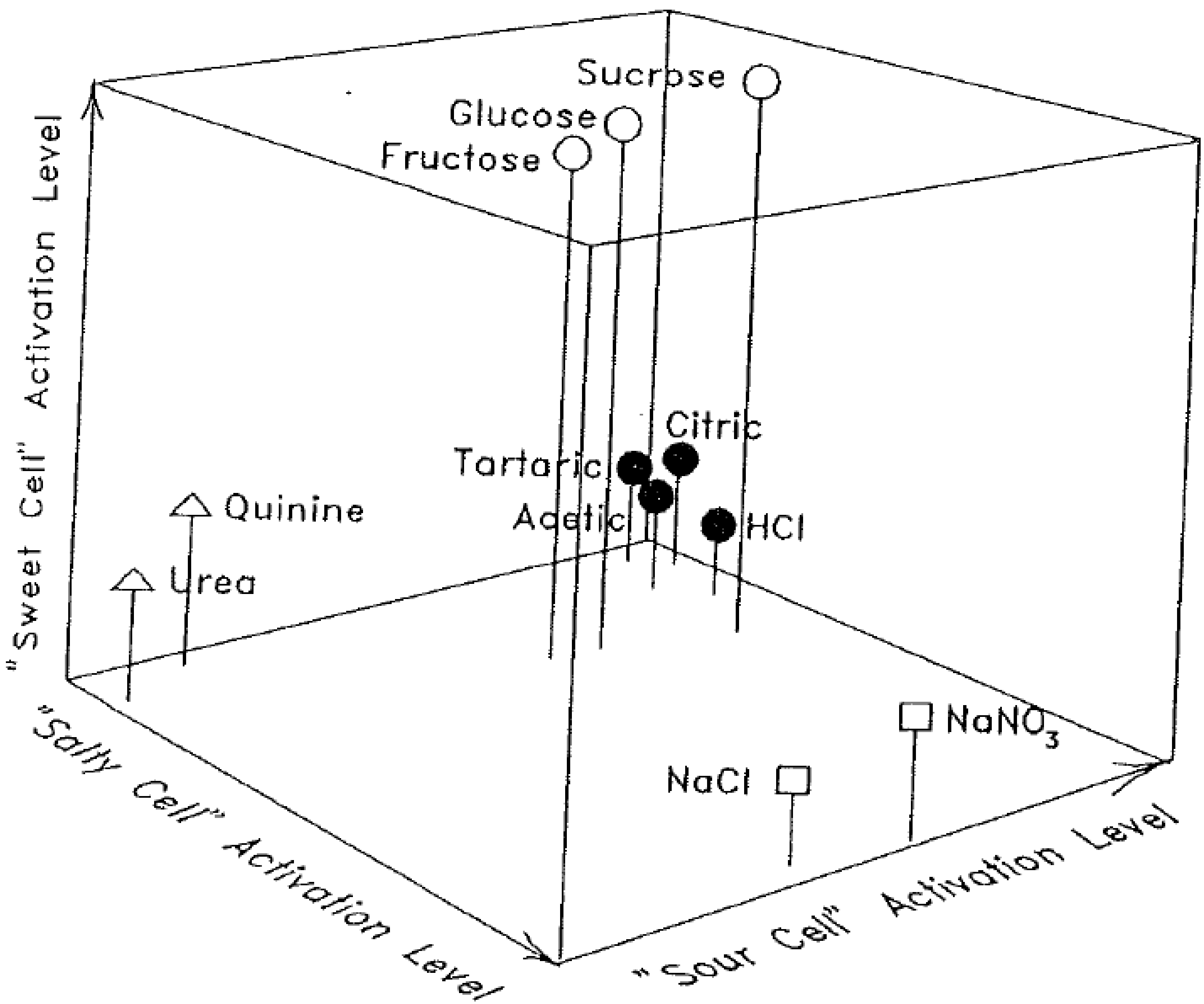





caricature of
your face
your face
average or
prototypical face




Kennedy
Clinton
1
2
3
4
5
6
1
2
3
4
5
6



Sucrose
Glucose
Fructose
Citric
Tartaric
Acetic
HCl
Quinine
Urea
NaNO3
NaCl
"Sweet Cell" Activation Level
"Salty Cell" Activation Level
"Sour Cell" Activation Level

Some Issues with Backpropagation Learning

- Cottrell had to present has training set 1000’s of times to the network to reach the performance it did.
- Cottrell ‘told’ the network what was right or wrong, and ‘reached in’ to change the weights (supervised learning)
- So how do we humans engage in ‘one-shot’ learning, usually unsupervised?
- Do we need to tweak some parameters, or is a completely different kind of learning called for?

Neural Networks: More General (current) Shortcomings

- Neural Networks are good for perception and control, with possibly some very short-term prediction: all pretty ‘low’-level cognitive tasks.
- Neural networks do not seem to be very good at more ‘high-level’ cognitive tasks such as complex reasoning or problem solving (things that more traditional, symbolic, AI is good at).
- Indeed, while neural networks seem to have the potential to explain low-level cognition that characterizes much of ‘animal’ cognition, it is not clear how it can be used to explain humans’ ‘higher-order’ cognitive faculties such as logical reasoning or long-term planning and prediction.

Situated Cognition to the Rescue?

- Suppose we situate the neural network in a symbolic environment (and embody the network with perceptual apparatus and motor control, so it can interact with its environment)
- Maybe the network can learn the kinds of perception-action sequences (!) that end up manipulating symbols in fruitful ways, and thus engage in higher-order cognition.
- An interesting ‘blend’ of Connectionism and Situated Cognition resulting in something that can be described as a Classical symbol system!