W9278091

1494680

# DESIGN OF A MEMORY BASED EXPERT SYSTEM FOR INTERPRETING FACIAL EXPRESSIONS IN TERMS OF SIGNALLED EMOTIONS

GARRETT DONOUGH ANTHONY KEARNEY

A thesis submitted in partial fulfilment of the requirements of the Council for National Academic Awards for the degree of Doctor of Philosophy

THESES
THAMES POLYTECHNIC LIBRARY
006.3302 8553 69 JAN KEA

March 1991

Thames Polytechnic
London

## Acknowledgements:

My thanks and admiration go to the many members of the Staff and students of the College, and also many others, including my family who found the time amid their other commitments to judge the photographs used in this thesis. Without their generous contribution the project would not have got off the ground.

I wish to thank my supervisors, Dr. Sati McKenzie and Professor Max Bramer for their guidance and encouragement throughout the period of research; and acknowledge, with thanks, the helpful criticism from Dr. Geoffrey Sullivan of Reading University in his capacity as external advisor.

Lastly, I enjoyed the benefit of sustained support from the research and programming group at Thames who provided the help and encouragement to carry me over the 'bad' times to which all long-term endeavour is prone. I wish to thank, in particular, Liz Bacon, James Amuah, Ian Lee, Jem Pearce, John Platts and Doug Muirden.

## Abstract:

A memory based expert system(JANUS) has been designed to interpret facial expression in terms of the signalled emotion. Janus accepts a geometric description of the face obtained from measurements on a digitised full face photograph and returns the appropriate emotion label. An intermediate representation in terms of verbal face actions(e.g. mouth open, eyes closed) is also used. A production rule system converts the geometric description to verbal form, while a dynamic memory interprets the face actions in terms of emotions. Following the work of Schank(1982) and Kolodner (1984), the dynamic memory is structured as a tree of packets, storing, in Janus, typical facial expressions connected by links to atypical but related face expressions previously encountered. This enables new input to be channelled along the appropriate path to an interpretation based on previous experience. The system is capable of learning new emotion labels and associated face actions for use in subsequent interpretations.

A prototype system has been developed on a SUN 2/MJN/120 system using POPLOG. Validation studies suggest that the interpretations offered by Janus are generally consistent with those of human experts.

# CONTENTS

# Chapter 1:

# Introduction

## 1.0 The purpose of this thesis:

The questions addressed by this thesis have relevance to a specific aspect of human-computer interaction: that of recognition and interpretation by a computer of user facial expressions of emotion. The questions which this thesis attempts to answer are, however, more general than whether a particular application can be implemented in a specific computer-user interaction. Instead, an experimental system (Janus) has been developed to interpret an input description of facial expression in terms of an emotion label. This work attempts to answer the following questions:

(1) Can a suitable representation of facial expression be found which is sufficiently precise for manipulation by computer and yet sufficiently expressive for the intended purpose?

Two types of representation are used in Janus - geometric and verbal. The former is in terms of coordinate values of selected facial landmarks while the latter is in terms of face feature actions (e.g. mouth open). Conversion of the geometric to verbal is automatic. It should be noted that the face actions are not dynamic, but merely represent an assertion that a movement from repose has occurred.

(2) Can a computer system be designed to use the representations in (1) and provide an interpretation of the signalled emotion which is consistent with that of human beings judging the same expression? In other words, is a description of facial expression sufficient for the recognition of emotion in the absence of any other contextual information?

It is the aim of this work to show that the answers to these questions are in the affirmative. Janus is a memory-based (i.e. incorporating a dynamic memory based on the work of Schank, 1982 and explained in Chapter 7) expert system capable of interpreting facial expressions by sequential heuristic classification - sequential, because consecutive steps are traced in the mapping of pixel abstractions to face actions and from the latter to emotions, and heuristic, because the mapping is based on empirical associations rather than causal relations. Janus is also capable of learning from experience, learning new interpretations and associated face actions and incorporating them into its dynamic memory. Janus was developed as an experiment in making computers sensitive to the body language of users.

The significance of the work (in the author's opinion) can be summarised as follows:

1. It is the first implementation which transforms face geometry into concepts of emotion using its experience of similar expressions.

2. It is an extension of memory-based methodology into a new application area.

3. The approach used in this work could be extended to deal with geometric data in other areas of computer vision, e.g. to form a learning and classifying back end for scene analysis systems.

4. This work could represent a small step forward in the more general area of enabling computers to recognize and hypothesize about the psychological preoccupations of the user. These points are discussed further in Chapter 9.

## 1.1 Related domains of face research:

The possibility of using non-verbal communication as a means of human-computer interaction has attracted much attention recently. Several systems have been reported(Sheehy 1989, Mase et al. 1987, Aleksander and

Burnett,1983, Stonham 1986). The problem of recognition and recall of facial features is also of interest to psychologists and raises a number of fundamental questions relating to:

**(1)** The structure, organisation and functioning of human memory(see Bruce 1988 for a critical overview and also Bower et al. 1981, Baddeley 1979, Patterson & Baddeley 1977, Strnad & Meuller 1977, Winograd 1976, Bower & Karlin 1974, Wells & Hryciw 1984 among others).

**(2)** The way faces are perceived(Sergent 1984, Jensen 1986, Ellis et al. 1986, Courtois & Mueller 1979, Davies et al. 1977, Galper & Hochberg 1971).

**(3)** The importance of context effects (Watkins et al. 1976, Bower & Karlin 1974 among others).

While there has been a wealth of research over the past century in specifying the facial actions signalling emotions and the measuring techniques used (see Ekman 1982 for a discussion of many of these in terms of his evaluative criteria), the problem of how these are represented in memory and the strategies enabling their recognition and recall have received less attention than the related question of face recognition.

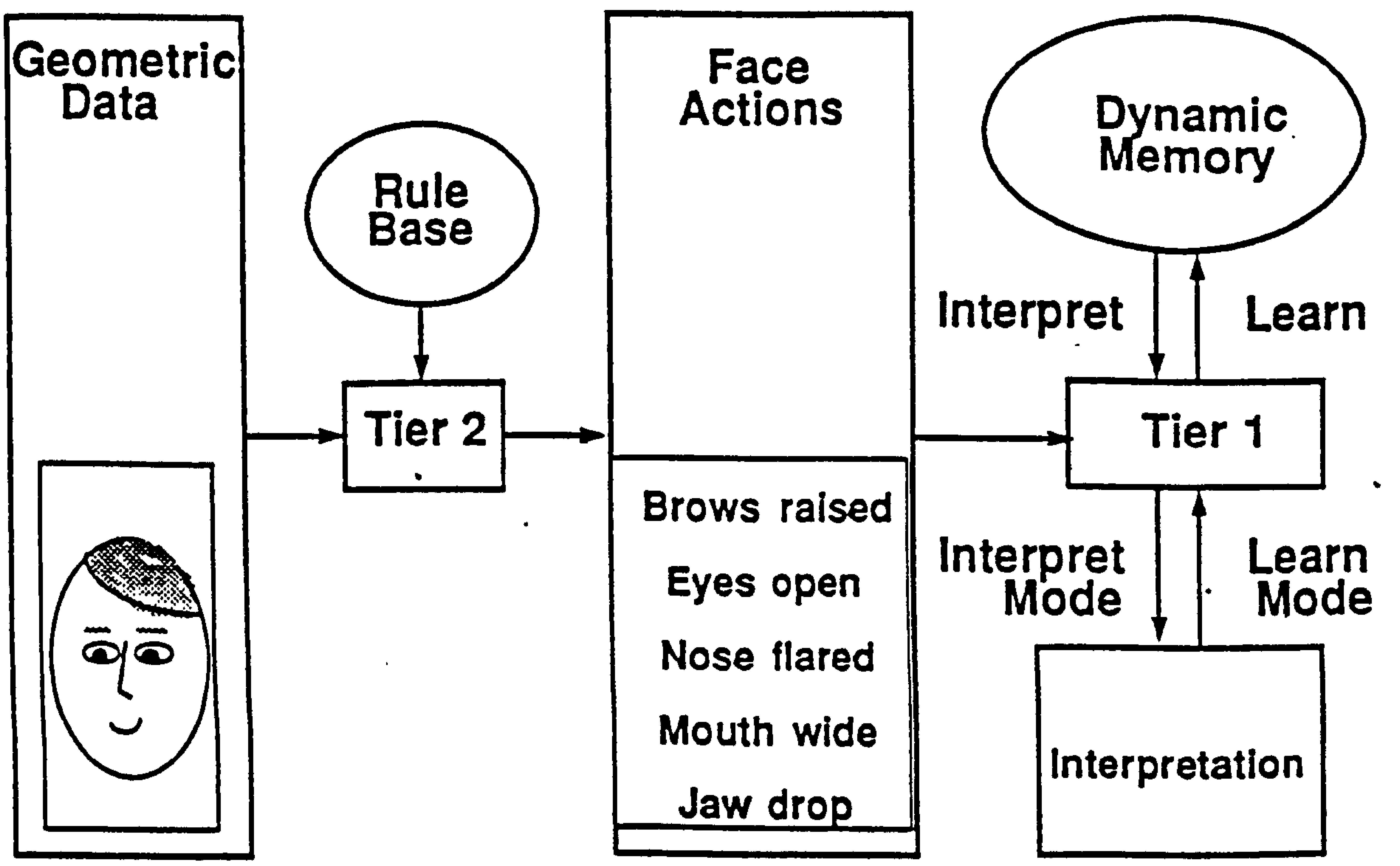
### 1.2 Janus - basic functions and components:

Janus is named after the Roman god, *Janus*, who is the god of cross-roads and beginnings. He has the wisdom of all that has gone before and knowledge of all that is to come. He is depicted as looking both ways at once: applying the lessons to be learned from the past to the events of the future. Janus, the system, attempts to do this in a very limited domain. It interprets face expressions in terms of its experience with similar expressions.

At an advanced stage in this project, the writer came across a reference to another system called "JANUS" (Day,1987) which is a hybrid system of neural networks and a production system. This work (which is believed not to have been implemented) is concerned with integrating automatic and controlled problem solving. It has no bearing on the research reported in this thesis.

Janus does not organize emotions directly. It organizes face actions. Janus is designed to accept a facial description and return an emotion label. The input

**Figure 1.0: Janus: basic components**




description may be geometric (coordinate positions of 34 selected landmarks currently obtained from measurements on a digitised full face photograph) or syntactic (a list of verbal face actions such as "mouth open", "nose flared"). The geometric description, if used, is converted into the syntactic form prior to interpretation. The conversion is done by a rule base. The interpretation is in the form of an emotion label, such as "happy" or "angry", and is accomplished by a dynamic memory based on Schank's "Memory Organisation Packets" (MOPs) and his theory of reminding and learning(Schank 1982). In addition to offering interpretations, Janus is capable of learning new emotion labels and associated face actions, thereby increasing its expertise with use. This learning results from user teaching - by typing in from the keyboard a set of face actions with a associated emotion: the result of viewing and judging a face or face-photograph. The basic components of Janus are shown in **Fig. 1.0**.

### 1.3 Janus - a distributed, self-organizing memory:

Janus differs from conventional expert systems in incorporating a dynamic memory. It is "dynamic" because it automatically re-organizes its classification of input face actions signalling discrete emotions in the light of their resemblance to those which it has experienced before. Such new organization is immediately pressed into service to process new input. This is done by juxta-positional generalization. When freshly input *face action components* (these are either a face feature such as 'brows' or the face action such as 'brows raised') are stored they may traverse the same branch of memory as some identical previous input. In this case, they form a generalized sub-category of the face action component when they are juxtaposed at a node, forming a hierarchy of *classes of face action components* (i.e. of "brows" or "brows raised" symbols). The advantage of memory-based systems is that, like human beings, they develop their expertise through experience. They also offer the possibility of successfully tackling a problem at a more generalised level if no specific rules apply. Human beings do this when faced with new situations. Janus is also a distributed memory system and has some capability for retrieving previously-stored face expressions (with their emotions) from incomplete input expressions (graceful degradation).

### 1.4 Janus - the role of validation studies:

The performance of Janus has been subjected to validation. Validation tests that a system performs with an acceptable accuracy. Such studies play an important role in the development of expert systems and indeed any software system. Validation studies on Janus have been aimed at testing both the rule-base and the dynamic memory components. Both the interpretation and learning functions have been considered.

#### 1.4.1 Gold standards used in validation:

The conclusions of Janus were compared with those of human 'lay-experts' drawn from college personnel. The use of the term 'lay-expert' denotes college folk whose everyday occupation is centred on people and who possess experience of decoding face expressions. While lacking formal tuition in the task, they have daily interaction with persons of varied ethnic cultures. An additional gold standard was provided by the descriptions given in Ekman & Friesen 1975 and their "Pictures of Facial Affect"(1976b, "PFA"). Both informal qualitative assessments and quantitative comparisons using standard statistical techniques were carried out. The accuracy of the human judges used to validate Janus' inferences may be measured, when possible, by reference to the "PFA" standard.

The chief drawback of the present system is the absence of a visual front end to scan faces and to calculate automatically the geometric description from the digitised photograph. Such a front end would need to be capable of very sophisticated image processing and is outside the scope of the present project. Nevertheless it would prove very useful.

This introduction is concluded by a synopsis of the contents of the remaining chapters:

**1.5 Synopsis of remaining chapters:**

**Chapter 2** contains a review of face research based chiefly on psychological literature. The broader domain of emotion is described in terms of current theories; its relation to human memory is discussed. The domain then narrows to face processing and two foci of research are identified: the signalling of personal identity and the signalling of emotion. Theoretical frameworks of, and experimental approaches to, face processing are described. The importance of context in this processing is discussed. The evidence for universal agreement as to expression of discrete emotions leads to the concept of typical expressions for these basic emotions.

**Chapter 3** explores the range of computer and non-computer systems involving face perception. These are seen to address such areas as tutorial systems to improve a person's perception of signalled emotion and computer systems which aid witnesses produce a description of a criminal which can be automatically matched to a computerized 'mug file' of previously convicted criminals. The Chapter ends with a statement of the questions addressed in this thesis.

**Chapter 4** describes the role of human experts both in providing initial knowledge for the development of Janus (the Knowledge Acquisition phase) and during the validation process. The use of an alternative gold standard is also discussed.

**Chapter 5** explains the representation of facial expressions in Janus in detail and compares this with the representations used by other researchers in their work.

**Chapter 6** describes the design and operation of a basic component of Janus, namely the rule base. The rule base transforms the numerical representation of positions of facial components comprising an expression on a digitized face image into the 'syntactic' representation in which they are stored and processed in the dynamic memory.

**Chapter 7** describes the structure and function of the dynamic memory. The basic lattice is represented by a list of record data-types each of which has an 'info' field which can contain a reference to an object and a 'links' field which lists the links to daughter nodes(records). Dual data-types conceptually distinguish the 'physical' knowledge hierarchy of face components from the broader conceptual knowledge which is represented as a POPLOG FLAVOURS OBJECT hierarchy.

**Chapter 8** covers the validation studies on Janus. In all of these studies the conclusions of Janus are compared with human judgments on the same expressions and tested for statistical significance. Janus' conclusions include the face actions produced by passing a numerical face description through the rule base and the emotions which are retrieved from dynamic memory by an input of an expression in search of interpretation.

**Chapter 9** critically discusses the current system vis-a-vis the aims of the project, weighing what has been achieved against the limitations and drawbacks.

**Chapter 10** outlines possible further extensions to this work involving the modifications necessary for including context as an aid in interpreting facial expression. It also discusses extensions to the system which would allow inferences to be made by the system on the cognitive associations of emotional states as revealed by facial expression.

# Chapter 2:

# Review of related research

## 2.0 A review of face research:

This chapter reviews the present state of the art in face knowledge. Although the aim of this project is to implement a computer facility for interpreting face expressions, and this could be thought of as a model of how humans go about this task, such a model is limited in its scope. It does not, for instance have anything to say about the early face processing required to define face features or to recognize a face for what it is as apart from other objects. Neither is it primarily concerned with extending the analysis of expressions to include the contextual memories, thoughts, goals, plans and circumstances of these. It is however necessary to be aware of other models in the domain so as to define the boundaries of the present project, and to realise how incomplete a model Janus is. The terminology becomes technical here. As a rough generalization, mental faculties may be described under three categories:

1. Thinking (Cognition),
2. Feeling (Affect),
3. Striving (Conation).

In the following review of psychological theories, these terms may acquire more specific meanings to the protagonists who use them. It is common to hear reference to "The Cognitive System", and this can include memory components as well as processing functions. Detailed explanations of the technical would tend to obscure rather than aid understanding. An unsophisticated glossary is included here to give the gist of the meaning.

This chapter starts by emphasizing the distinction between the signals that the face emits about emotion and those about the identity of the person. In Section 2.1, the literature concerning representation of face expressions in the human is reviewed. Three different approaches to

**Glossary of terms used in this chapter:**

| Word or Phrase | Meaning |
|---|---|
| affects | emotions, moods, attitudes |
| autobiographical memory | memory for events in one's life |
| configural | relating to pattern of parts |
| discriminanda | feature selected for recognition |
| episodic memory | memory for events in a context |
| ethology | study of animal,human behaviour |
| expression | one or a set of face actions |
| facilitation | allowing a faster response |
| functional | what function is served for Man |
| holistic | the whole is more than parts |
| innate | born with |
| logogen | word detector (Morton, 1969) |
| metric | measure |
| mood | longer lasting affective state |
| orienting tasks | tasks to direct testees' task |
| phylogenetic | evolution though phyla |
| physiological | how body organs, cell, work |
| preferenda | affectively preferred objects |
| proposition | formal abstraction of concepts |
| psychological | of the 'mind' |
| recall | past events from memory |
| recognition | present, familiar, seen before |
| reflex | in reply to stimuli, not willed |
| salient | standing out |
| semantic | having meaning |
| semantic memory | memory of facts, concepts |
| somatic | of the body, distinct from mind |
| subcortical | below the brain voluntary level |

the perception of emotion : the category, dimension and ethological are outlined. Next, the strategies that humans use in perceiving the face are reviewed. There is evidence that both holistic and feature-analytic strategies are used. The results of this review lay the ground work for the strategies which Janus adopts. The broader domain of emotion is addressed in Section 2.2 and includes summaries of theories about its purpose; functional frameworks about how face perception is organized and the relation of emotion to memory in the human. The chapter ends with the description of some representations of memory.

## 2.1 The face: A multi-signal / multi-message processor:

Ekman (1984) describes the face as a multi-signal, multi-message processor. Age, identity, health, psychological state, personality character traits, degree of excitation as well as emotion are some of the messages that can be signalled by different static and mobile elements. Static signs are, for example, skin wrinkles. Mobile signs include colour and muscle action. The skull, jaw and the teeth give the basic shape, but the soft tissues: muscles, connective tissue, skin and hair define the contours. These soft tissues are moved by muscles most of which are tethered in the bones but other muscles: e.g. those which make the hair stand up, are tethered in soft tissue.
Muscles are usually classified as voluntary or involuntary depending on whether we can command their action. The muscles of expression are voluntary in this sense but also have some involuntary, reflex action. Thus it is possible to pose expressions of emotion, and also find it difficult to hide feelings. Though clearly, the effects of voluntary muscle contraction on the elastic soft tissues of the face provide the activities used to signal emotion, there are others e.g. tears, wetness, blushes, pallor, goose-pimples, pupil-size, "soft", "surly" and "hard" looks which seem to involve the autonomic nervous system. Reflex increases in muscle tone are involuntary and some

expressions are reflex. The way in which face expressions of emotion are described in the literature has turned full circle from Duchenne in 1862 who contracted face muscles to observe their effect, through a period of use of descriptive syntactic feature actions (e.g. brows raised) to the more recent anatomically-oriented descriptions of Ekman & Friesen (1982), Ermiane & Gergerian (1978).

### 2.1.1 Identity vs Emotion message:

There are two points to note concerning message-bearing components of the face:
Firstly identity of the individual cannot depend on the same signals as does emotion display since we can classify a person to be somebody we have or have not seen before, be she happy, sad, angry, afraid, surprised or disgusted, as also we can recognize a disgusted expression whoever bears it. This is not to imply that expression plays no part in recognition; it has been found to play some role (see for instance Galper & Hochberg 1971, Sorce & Campos, 1974). As a practical result, one would not expect a system dependent on emotion signals to return the identity of a poser nor one dependent on identity signals to interpret the emotion of a person.

Secondly, the muscle contractions underlying face actions are produced over a finite time rather than instantaneously with the result that each has an onset, an apex and a decline (Ekman & Friesen, 1984). Our experience of the signals of emotion transmitted by others is therefore grounded in time sequences. Much of the research carried out to identify these signals has used still posed photographs as stimuli to be judged by experimental subjects, and their use has been criticized as irrelevant to everyday experience of emotion. However, it must be allowed that very high consensus can be obtained among judges as to the emotion signalled using posed photographs. It has been suggested that what is signalled and responded to in these displays are culturally-agreed conventions.

### 2.1.2 Dimension vs Category approaches:

It is common knowledge that categories of emotion with commonly-accepted labels occur in response to commonly experienced life events. Sadness, for instance, would be the typical response to many commonly experienced events. This "category" approach to the perception of emotions is so general that most people, if pressed to do so, could provide a shortlist of "sad" events 'out of their head'.

On the other hand, the notion of underlying common dimensions of emotions has led to the study of the number and nature of dimensions that can be perceived. Dimensions are statistical abstractions based on such measures as similarity ratings between face photographs. The following abstract of the dimensions of emotion is from Saltzen (1981). One such dimension commonly found is pleasant/unpleasant; another usually has an idea of activation/intensity and a third, attention/rejection (Dittman,1972 reported by Saltzen). However Saltzen suggests that the problem with both the category and dimension approaches to the perception of emotion in the face is that the results may simply reflect the experimental approach and the cognitions of the experimenter. Real characteristics of face expression perception are expressed in both. Saltzen refers to a number of studies in which analysis yielded clusters of judgements, suggestive of categories of emotion, which were distributed in multidimensional space. These results appear to underline this dual interpretation of perception of face expressions.

### 2.1.3 The Ethological Approach:

There is a very different view of facial expressions: that of the functional significance of the display in terms of, say, defence of the eyes in face of threat; of rejection of noxious material from mouth or nose, of feeding movements and those movements which increase alertness. The phylogenetic evolution of these functional expressions are evident in Man in the movements we recognize as a smile, a sneer, a stare, a laugh, pouting, grimacing etc. Saltzen (1981) lists the functional categories underlying such responses as:

1. Attend,
2. Incorporate,
3. Attack,
4. Reject,
5. Retreat,
6. Protect,
7. Relax,
8. Tire.

These are variously described in terms of the face, head, mouth, and respiratory, muscular and autonomic systems. He shows how a paired interaction of these functional expressions could be interpreted both in the terms commonly used for emotional categories, and also account for the three commonly measured dimensions of emotion: unpleasant/pleasant, thwarting/end-of-thwarting activation (by level of arousal of the tendencies thwarted); and the attention/rejection dimension (by whether the arousing stimulus is a stimulus for approach or avoidance behaviours).

### 2.1.4 Holistic versus feature_analytic face processing:

The problem addressed under this heading is whether humans perceive a face as a whole or gestalt (as something more than a sum of its parts) or as a sum of its parts. An analogy is whether three dots are perceived as a triangle or as three dots. The holistic perception depends on an interaction among the features; whereas the feature-analytic approach rests on the perception of individual features. It is certain that both forms of perception occur naturally, though the conditions under which each typically occurs has to be defined. Apart from understanding human perception, this has implications for the design of systems to help witnesses reconstruct a face from memory (Identikit and Photo-fit use sets of independent facial features) and if consistently one sort of processing were found to make recognition of faces easier, then the strategy would be worth teaching to those whose business it is to recognize faces. Again, if there are qualitative differences between recognition and recall in this respect, it has relevance for memory experiments aimed at elucidating the organization and search techniques that are most effective in retrieval from human memory. Interaction between inner face movements has been demonstrated (e.g. by Ekman & Friesen, 1984) to be significant for perception of facial displays of emotion.

A decision has to be made as to how descriptions of face expressions are to be represented in the proposed project: whether in terms of features or more holistic descriptions. In the knowledge acquisition phase when the chosen experts judge expressions and give reasons for their judgements, it will be interesting to note the way they do this and whether it is in accord with the research findings. It is proposed here to review the controversy arising from the findings in a selection of experiments conducted by psychologists with the object of finding out how humans perceive faces.

The way in which humans learn, recall and recognize faces is usually investigated by psychologists employing some version of a study/test sequence in which the faces are first displayed ( the study phase ) and then, mixed with an equal number of distractor faces, re-displayed (the test phase) one at a time. The basic procedure is modified in order to introduce the factor that is of interest. This may be a contrivance to orient the subjects' approach to the task. For example, requesting the subjects to decide whether the test face's most distinctive feature was the chin, lips, nose, eyes, or forehead (Courtois & Mueller, 1979) would probably ensure that the subject attends to multiple physical features. Providing a rating scale focuses the attention similarly. These instructions are called "orienting tasks". The time taken by the subjects to signal recognition is often the dependent variable.

A focus on specific face features for learning, recalling or comparing faces is called a feature-analytic approach. An orienting task aimed at forcing a holistic strategy in the subjects is to ask the subjects to decide on an abstract personality trait or attribute. The latter tasks are believed to require "deeper processing" (Craik & Lockhart 1972). Deeper processing takes longer than superficial, and tasks such as judging the honesty, generosity or likableness of a face are supposed to require deeper processing than attributes based on physical feature judgements, for instance the sex of a face, and have been found to lead to better retention in memory (Mueller et al. 1978, Strnad & Mueller 1977, Winograd 1976, Bower & Karlin 1974). A multiple physical feature orienting task however has been found to lead to better retention than the single feature case although still slightly worse than the abstract trait condition, (Courtois & Mueller 1979; also suggested by Bower & Karlin 1975). Parkin and Goodwin (1983), however found that a supposedly holistic orienting task did not influence the overall performance level in their experiment where the comparative orienting tasks were:

1. To make a decision as to the meaning of the expression and
2. To judge the most distinctive feature on the face.

However the authors urge caution in accepting this result as support for Winograd's (1981) theory that both physically- and trait- oriented strategies address the number of physical features encoded. They hold that the more features that are included in the orienting task, the more likely it is that the most distinctive feature will become part of that face's representation. Cases where the physical orientation failed may not have included the necessary features such as these distinctive features.

Baddeley (1979) reporting on the evaluation of a training course based on feature analysis as the method for person recognition concluded the course was ineffective. He concluded, on the basis of further experiments reported, that there was a small but significant advantage for recognition performance in the test phase when subjects categorized the faces in terms of personality as against judging physical features at study time.

It seems, on the basis of these results, that neither approach alone can bring a large advantage over the other to person recognition. Baddeley notes that Bruce (1977) found that the semantic characteristics of faces are important for recognition when scanning for more than one familiar face but that resemblance might be more effective as a strategy in the single case.

This leads us to the view that both holistic and feature analytic processing are used in human person recognition and when they are used is task-related.

A more pertinent investigation, from the point of view of this thesis is that of Jensen (1986). He asked whether groups of facial features are perceived holistically or independently of one another. He used a classification task on inner (the eyebrows, eyes, nose and lips) and outer (hair and chin) groups of features, making subtle changes to these to see if subjects detected changes. When the inner features were separated, either by grouping the eyebrows or the lips as an outer feature, classifications by the inner group were affected by changes in the outer group although classifications by the outer group were unaffected by changes in the inner group. These results indicate that the outer features are perceived independently of the inner features and that the eyebrows, eyes, nose, and lips are perceived as a group.

Sergent (1984) also investigated the component and configural processes underlying human face perception. Her analyses provide support for the hypothesis of a simultaneous unfolding of separate processing strategies using component and configural properties possessed by faces.

Ekman & Friesen (1984) provide numerous examples of the interaction between the inner face features in the typical facial expressions which signal common emotions. For example, unambiguous anger requires the distinctive changes in the brows, eyelids and mouth. If only the compressed mouth occurs, it can be anger but also exertion. It can also be seen that the outer face features contribute little to typical emotional expressions in spite of the reports of hair standing on end, unless one includes the movements of the jaw necessitated by opening the mouth. The outer features, for example the hairline, are more salient cues in the identification of a face than in interpreting the expression of emotion on it.

## 2.2 Theories of emotion:

What follows is not a full review of the field but a taste of the direction that research is taking. The various theoretical positions adopted by researchers in defining categories of emotions have been classified by Izard (1971). His classification is summarized here to give an idea of their diversity: thus human emotion has been seen as primarily motivational (Tomkins 1962,1963; Izard and Tomkins, 1966); as derived from adaptive biological processes (Plutchik, 1962); as related to perception or awareness (Leeper, 1965); as a function of perception and appraisal (Arnold, 1960); as physiological arousal and cognition (Schacter and Singer, 1962); as a complex response system (Averil, Opton & Lazarus, 1969) and as an aspect of personality (Izard,1959, 1960; Izard and Tomkins, 1966); these positions will not be discussed further.

The theorists who place high significance on face expressions as a component of emotion include Tomkins (1962) and Izard(1971) in whose theories it plays a central role in the differentiation and communication of emotion. Awareness of face and posture is also a prerequisite for the emotion theory of Bull (1951). Ekman and Friesen and their co-researchers over the years have developed an empirical theory of what expressions go with what emotion; they have done much as has Izard and others to elucidate the universally-recognised emotions and their typical expressions. They have perfected a comprehensive system relating expression to the causal muscle action units.

## 2.3 General levels of theory:

The levels of theory defined by Marr (1982) are used to classify the theories to be described. The "computational", level 1, involves asking what the goals are and what are the essential characteristics of the solutions to the problem, whereas the "algorithmic" level 2, describes the strategy used and the "implementational", level 3, the hardware. The computational level poses questions of the "What is being computed?" and "Why?", whereas the other levels answer the "How?".

For emotions, the computational questions are: "What purpose does emotion serve" or "What are the goals for which emotion provides the solution" and "What solution does emotion provide". It is for the theorist to specify whether emotion is a goal or a plan to obtain a goal. The algorithmic question is "What is the algorithm: how is the solution obtained?, The implementational question concerns the hardware involved, namely the neural substrate of the brain and the effector organs, muscles among them, of the body.

The treatment of theories of emotion here is extremely selective so as to present examples of the coming together of multi-disciplinary views on the role of emotion.

### 2.4 Marr Level 1 theories:

The first theory presented, that of Sloman (Sloman, 1986, Sloman & Croucher 1981a, Sloman and Croucher 1981b ) is developed within the context of required constraints on the design of human-like intelligent systems in coping with multiple goals when in an unpredictable setting. It derives from the property of all emotions to interrupt motives. The second theory is developed by Oatley and Johnson-Laird (1985) who give a functional explanation of emotions. A third Level 1 theory is the Informational Theory of Simonov (1966), which categorizes emotions in terms of the adequacy of information required for reaching a goal. Abstractions from these three theories are presented below.

#### 2.4.1 The Sloman and Croucher approach: "why robots will have emotions" (1981a):

The authors assert that the need to cope with a changing and partly unpredictable world makes it very likely that any intelligent system with multiple motives and limited powers will have emotions. There is a close relation of emotions to motives in the theory. Motives include representations of states of affairs or events to be achieved, preserved, and prevented. Processes may compare, assign, add, remove, select, suspend, reconsider and monitor motives. The system must be able to interrupt and redirect processing with new motives, if necessary, and resolve conflicts or interleave actions or planning in pursuit of different intentions. A parallel system monitors internal and external events for significance so that these processes can respond to them.

The concurrence of a strong motive with a belief about its possible fulfilment or lack of fulfilment - provided it is sufficiently pre-occupying to disrupt decision-making and distort perception and decision-criteria - normally produces an emotional state (positive or negative). The new motives or direct actions which may or may not be produced in this state depend upon qualifying factors such as whether there is doubt about the fulfilment, whether the motive is for or against something, whether satisfied or denied, past or present.

The theme of "interruption" is an important characteristic of such emotional states in the theory (cf the Oatley_Johnson-Laird theory below). Simon (1967) also emphasizes this property of emotions. Sloman and Croucher (1981b) speculate that the power of emotions to interrupt on-going cognitive processes will be subject to thresholds which can be raised by important motives and unconscious emotions. A disturbance need not spawn new motives (e.g. a desire to right what has gone wrong ) , but dwelling on it gives ground for learning or else direct action can be generated without the forming of an intention and a plan; this could be an impulse.

**2.4.2 The Oatley and Johnson-Laird (1985) sketch for a cognitive theory of the emotions:**

Emotions, according to this theory, are a biological solution to the problem of co-ordinating planned action with multiple goals in a world that is only partly predictable. In their theory, emotions have a specific communication function within the cognitive system and to others nearby. They arise from a specific separable system evolved by active animal species as part of the means whereby multiple goals can be coordinated. In social species they signal junctures in mutual plans.

Oatley and Johnson-Laird model the cognitive system in the human case as a modular hierarchy of processors in which those higher call lower modules, and emotion is one of two specific kinds of communication (the other is calling lower processors by means of propositions(propositional invocation). The Anderson & Bower (1974) definition suggests a proposition is an associative rule-configured cluster of elements which are abstract concepts the existence of which can only become known by their representation by a word or symbol. A proposition has a meaningful truth value.

In the Oatley and Johnson-Laird model, in contrast to propositional invocation, emotion signals propagate globally among the processors to set and maintain the system in an 'emotion mode'. Emotion signals are released at particular junctures of multi-goal planning sequences. They turn some processors on and some off. Emotion modes relieve the need for the highest level operating system to evaluate propositional data and reason about an appropriate action. There are a small number of these emotion modes. They are: happy, sad, anxiety or fear, anger, and disgust. Each tends to inhibit the others. The emotion mode is necessary but not sufficient for the full experience and expression of an emotion i.e. the distinctive phenomenological tone, the somatic changes, behavioural changes and courses of action. By itself it only prepares for action by focusing on some goals thereby amplifying emotions with a compulsive yet flexible attentional focus. The full emotion depends on the evaluation of the juncture in planning by the operating system and the meaning ascribed to the emotion mode so that voluntary action can be scheduled.

The five emotional modes: happy, sad, anxiety or fear, anger, and disgust are triggered by significant junctures of plans (to achieve goals) which are distinctive and recurring. The function of these modes is to organize a transition to a new phase of planned activity directed to the priorities of the mode with its associated goals, and to maintain it in that phase until another transition occurs.

The five junctures described are :

1: sub-goals achieved
2: major plan fail or loss of goal;
3: self-preservation goal violated;
4: active plan frustration;
5: gustatory(taste) goal violated.

The states to which transition is made are respectively:

1: continue with plan;
2: search for new plan:
3: stop, attend or escape;
4: try harder, and/or aggression;
5: reject substance/ withdraw

Dysphoric(negative) emotional modes usually interrupt an existing plan in response to a threat to a self-preservation goal or a blocked goal. New evaluation by the operating system is required urgently, and modifications or abandoning the plan may have to be decided.
Euphoric emotions continue while plans appear to be succeeding.

Mutual plans between individuals develop during normal social interaction, and similarly set up emotion modes. These are communicated to others by verbal and non-verbal behaviour, e.g. expressions. These are matched to the view of the self and the meta-perspectives of how the other views one. Mismatches between such views may trigger a dysphoric emotion mode. Mutual emotion modes can develop readily.

The Oatley-Johnson-Laird approach is an example of a cognitive theory (cognition, among others, includes perception, thought and memory) of the emotions pitched at Marr Level 1, as also is the Sloman & Croucher formulation of an essential role for emotion necessitated by the design of a multi-goal intelligent system. There is specific reference in the Oatley-Johnson-Laird theory to the communication of mutual goals by face expression and also to the matching of the perceptions of these and of the communicated expectations of significant others to the perspectives of the self.

Presumably there are stored representations of face expressions in memory by which communicated expressions are classified. This is the intended function of Janus. The viewer's task, if so motivated, is to infer the goals which spawned the plans which are in juncture. This is beyond the scope of the present project but would have to be attempted if a fuller understanding by a computer program was desired. At least the meaning of the expression in terms of generalities can be attributed on the basis of the expression alone. Knowledge from other sources would be needed to constrain the search.

### 2.4.3 An informational Theory of Emotion:

Simonov (1966) conceives emotions as physiological nervous mechanisms which ensures the adaptive behaviour of higher living beings in situations which disrupt their habit system. This may occur when there is a lack of the information required for reaching a goal and satisfying a need.

Positive emotions are aroused whenever the newly acquired information exceeds the previously existing need for it. They also motivate living beings to strive for their goals despite an "information vacuum" and so to overcome the vacuum. Logic alone is insufficient for the success of adaptive actions in a changing environment"

**2.4.4 Frameworks of face processing:**

These models are important as hypotheses about face recognition in the human whether in real life or from a photograph. They are "box and arrow" diagrams to explain experimental and intuitive findings relating to visual recognition of familiar faces, especially the route by which the visual code of a familiar person produces the verbal code of the name of that person. Each box represents some necessary function such as a type of encoding of visual information and the arrows represent the route followed by the encoded information. The order of the boxes is important for explaining the order in which humans can come to a decision about whether a face is familiar, access associated knowledge and lastly name.

Although they specify the order and required functional units and classify the encodings coarsely, these frameworks do not detail the algorithms involved and are therefore conceived as being Level 1 theories. The word "semantic" can roughly be taken to mean "meaningfully associative" i.e. anything meaningfully associated in memory with the person: e.g. the fact that a person is a politician makes the Houses of Parliament an association as does a photograph of the spouse etc.. This is called "person information" obtained from person identity nodes. The name can only be generated after the latter has been accessed. The frameworks are conceived as specialisations of a network of nodes and links so that the location is spoken of as the person identity node of the person.

### 2.4.4.1 The Hay and Young Model (1982):

The description of the model draws upon the description in "Recognising Faces" (Bruce 1988).

This model attempts to explain recognition of familiar faces in the human.
Hay & Young proposed that representations called "face recognition units", analogous to the "logogens" proposed by Morton (1969,1979) for word recognition, distinguished a familiar face from all others. In the original idea, a specific recognition unit has a threshold of excitation which must be reached for it to respond in an all or nothing way when a visual description of any view of a familiar face is presented. A stimulus face is encoded by representational processes into an abstract structural code and accesses recognition units directly. An indirect route to them is via other visual processes and cognitive processes including the analysis of expression. Positive recognition that the face is familiar, allows one access to an already-made store of information about that person and from there to the person's name. Access to the name of the person is only through this person information. Other information concerning salient features, clothing, gait and posture, will be stored in memory. A point to note is that this ordering is a prescribed sequential accessing of visual, semantic and verbal codes.

Important to this thesis is the framework's assumption that the analysis of expression and other visual information is independent from that of identity but parallel to it. The "other visual processes" route which processes expression also allows independent effects of visual and semantic identity e.g. information about clothing, posture etc (Bruce, 1979).

The face recognition unit was at first conceived as analogous to a logogen and inherited its properties: the threshold of a unit was lowered by presentation of the word or by exposure to a semantically-related (meaningfully-associated) word. For the face recognition unit, this means that if the same face is re-presented within a small time again, the unit signals familiarity more quickly than before. Faces and objects which are associated with the face in the mind of the subject have a similar effect. The presentation of faces and objects so as to bring about this lowering (facilitation) is called "priming": identity priming in the case of the same face, and semantic priming in the case of associations to the person.

According to Bruce (1988), the problem for the Hay and Young model was that research on humans suggested that different views of the target used as primes should produce graded facilitation according to their resemblance to the target but the face recognition unit as originally conceived responded to any view of the face unrelated to the degree of resemblance or pose. An additional problem for the model was its inability to explain the human finding that of the two kinds of priming: identity and semantic, the latter depended on the target being presented immediately after the prime. Another problem occurred in explaining resemblance since one can note resemblance without identifying a face.

There seemed to be two options for theorists: either to reformulate the model, especially the working of the hypothesized face recognition units or alternatively to explain identity priming in terms of episodic memory in which an event is held to be recalled or recognised at test to the extent that the test situation emulates the context present at study. One characteristic of such a

representation based on episodic memory is that it should show similarity effects i.e. there is more accurate response to identical than to similar episodes or faces.
A modification to the way face recognition units function was made by Young et al.(1985) which involved a change to a graded response according to the degree of match - the degree to which a face resembles the stored description. The interpretation however is aided by contextual information which could detail the circumstances under which the face was seen.

#### 2.4.4.2 The Bruce & Young Model (1986):

Bruce and Young (1986) collaborated on a further development of the above model which is summarized here. Again the source is from "Recognising Faces"(Bruce 1988).

There are two levels of visual structural code producing two abstractions of the face; namely a viewer-centred abstraction for the analysis of expression and facial movements of speech, and also expression-independent descriptions which are derived in the same code as those in the face recognition units for which they are destined.

The face recognition units maintain the modification introduced above and their activation levels are affected by recent use or by context feeding back from the person identity nodes. The person is recognised at this latter level from either face identity or all the other non-face cues e.g. stature, contour, name etc. It is when the person identity node is accessed that the identity specific codes become available. Names can only be generated via the person identity nodes.

In this model, the analysis of face expression is not described in detail. It is proposed that it requires a viewer-centred visual description and that it proceeds independently of the process sub-serving face recognition.

Bruce and Young see the expression analysis, facial speech analysis, face recognition units , person identity nodes, directed visual processing and name generation as having links with what is called 'the cognitive system' - a catch-all containing among other functions: semantic memory, selecting between different sources of information, decision-taking, initiating responses.

In this model, whether a face is familiar is signalled by its face recognition unit but more than one encounter is required (the exact requirements are unspecified) before a face recognition unit is set up. Until then a face remains unfamiliar. Faces, say, in a test/re-test experiment are recognised by a combination of pictorial, structural and visually-derived semantic information. Whereas structural visual codes are abstracted from many views of the same face, pictorial codes are formed from a particular view of the person.

**2.4.4.3 Hadyn Ellis' Model (1986):**

Operation:

Ellis (1986) sees the processing of visual information about faces as proceeding in the direct order of overall facial configuration, sex, age and race, and identity. First the face is recognised as a face; then as a face of a certain sex; then a male or female face of a certain age group; then a race is assigned to this face and lastly , an identification. Ellis includes a parallel link between "physical analysis" stage and the "person nodes" thus by-passing the "face register" (analogous to the face recognition units of the Bruce and Young model). The analysis of expression in Ellis's model draws on the early structural encoding and, becomes part of the "contextual and other knowledge" which both sends and receives information to the face register, person nodes and name register. One encounter is sufficient to include a face in the face register.

Comment:

The interest in these framework models in relation to the present project is in the concept of a face recognition unit and in the separate routes taken for the encoding sequence of identity and expression. When the method adopted is discussed it will be seen that a representation of the typical facial expression of each primary emotion is crucial to the formation of a concept hierarchy for the classification, storage and retrieval of expressions and their interpretations. It will be necessary to discriminate between proposed structures for these representations and the face recognition units above. The relatively independent route of analysis of expression in the above frameworks encourages the attempt in this project to propose a working model free from the constraints imposed by identity.

### 2.5 Marr level 1 & 2 theories:

These theories combine the "why' and "what" with the "how":

#### 2.5.1 Tomkins' facial feedback theory:

An example of theory which goes some way to cover the algorithmic and implementational as well as the computational level is that of Tomkins (1962, 1963) . He uses the term "Affect" as synomymous with "Emotion". Use is made here of Izard's summary (1971).

For Tomkins, the affect system is the primary motivational system. Affects amplify drives. For him also, affects are primarily facial responses, which act as feedback bringing their own reward or punishment. The programs for the distinctive face responses are innate and subcortical. They command specific responses from other systems such as the cardio-vascular and the endocrine (hormonal) as well as the face.

The expressions of the affects in the face, in the body and in the viscera are distinct and innate. When these become conscious one becomes aware of affect, but awareness is not a necessary condition for the development of affect. These face and system responses are stored in memory images which can be recalled. He considered eight affects to be primary: of necessity these had specific facial responses:

1. Interest/excitement (brows lowered, stare);
2. Enjoyment/joy (smile);
3. Surprise/startle (brows raised and blink)
4. Distress/anguish (crying);
5. Disgust/contempt (sneer);
6. Anger/rage (frown, clenched jaw, red face);
7. Shame/humiliation (eyes and head down);
8. Fear/terror (stare, eyes wide or avoiding, pale cold sweating skin, trembling, hair erect)

How these affects are activated depends on a single metric: the rate of change of the number of nerve cell firings per unit of time: if increased slowly: interest; if increased rapidly: startle; if remains high: anger or distress depending on the level; if dropping: degrees of joy depending on rate of fall; The part of the brain (the reticular formation) which has been found to be associated with general arousal, Tomkins postulated, was responsible for general amplification of any message including those of affects. For Tomkins an emotion is a rewarding or punishing experience produced by feedback from body ( face and other ) responses to a happening. The feedback allows one to perceive the cause of the emotion. This awareness develops a motivation to maximize positive affects(emotions) and minimize negative affects. There are innate precipitants of emotion (the human is born averse to and predisposed towards certain experiences).

Others are learned. These can produce emotion from thoughts.
Another theory pitched at Levels 1 & 2 is that of Bower (1981). However, because of the Memory focus of this theory, it is included under the Section: Emotion & Memory, below.

## 2.6 Marr level 2 theories:

This section is concerned with a very large research endeavour going back many years; that of describing the typical face expressions of discrete emotions, and ascertaining how universal is their occurrence and recognition. The rationale for including such theories at Level 2 is that, if universal, it is likely (unless they are vestigial) that specific expressions are instrumental in evolutionary terms (Darwin 1872) in achieving a goal or several goals. They seem more likely to serve, as well as be part of, a goal and thus are an answer to "How" a goal is achieved. Even in the feedback theory of Tomkins the goal served could be that of bringing awareness. Another goal served is communication of the emotional state of the signaller. This in turn communicates the goal junctures taking place. It is not intended to present all the theories of typical facial expressions of emotion. There are many labelling differences in these that may obscure the results of research that are concordant with the concept of universal emotions which are conveyed by typical face expressions. Summaries of this research are the main content of this section. It is relevant to the present project because of the possibility of using typical expressions as templates in the proposed system.

The typical expressions of common emotions which have been researched by many workers are levels of description describing how the goal of facial expression is effected, albeit on a phenomenological and syntactic basis in many cases. The anatomical descriptions (Duchenne 1862, Landis

1924, Frois-Wittmann 1930, Fulcher 1942, Hjortsjo 1970, Ermiane and Gergerian 1978, Ekman & Friesen 1982) were based on the actions of facial muscles and are on a deeper level (nearer the machine code so to speak) than the phenomenological level of description. Both are classifiable as Level 2 theories. There are now very many of such studies and their varied methodology has been reviewed by a number of workers in the field (Saltzen 1981, Ekman 1982).

### 2.6.1 Pancultural elements in expressions of emotion:

These studies are motivated by a desire to know whether the typical facial expressions of emotion which show consensus of understanding in one society are typical in their form and understanding in other cultures. The conclusion of Ekman & Friesen (1984), who have carried out many of these investigations, is that the facial appearances of at least some emotions, those of surprise, fear, disgust, anger, happiness, sadness are universal. They noted cultural differences in the circumstances in which these expressions are shown. The expressions of U.S. and Japanese students (actual facial movements, grabbed from video tape) were concordant when watching a stressful film alone in their respective countries but discordant when a researcher was present and the students talked about the experience while watching the film. Cultural rules applied and the Japanese masked their expressions of unpleasant feelings more than U.S students in this situation.

The showing of photographs of facial expressions to people of various cultures and assessing the degree of concordance achieved in a choice of six primary emotion words was another technique used. Ekman's group (summarized in Ekman & Friesen 1984) found the same facial expressions were judged the same in the U.S., Japan, Chile, Argentina and Brazil. Independently, Izard (1971) found evidence of universality in eight different cultures: U.S., English, German, Swedish, French, Swiss, Greek and Japanese.

The cultures studied above may have acquired consensus through the mass media. In order to investigate pre-literate cultures, Ekman & Friesen (1984) tested people in the Southeast Highlands of New Guinea where the technique used was to show each person three expression photographs at once, have a translator read a story and ask the subject to point to the photograph that fitted the story. The results were concordant with the literate cultures with the exception that fear and anxiety expressions were confused. As a variant other subjects were requested to show the emotion on their own faces and results were also concordant. A further culture in West Irian, New Guinea was investigated by Heider & Heider(summarized in Ekman & Friesen 1984) who also found evidence for universality, repeating the experiments in an even more visually-isolated people in the island. Ethological studies by Eibl-Eibesfeldt (1970) also attest to the universality of facial expressions of emotion. Ekman & Friesen (1984) point out that cultures do differ in what circumstances elicit an emotion; also conventions in the management of emotion in social situations may differ in response to the same situation.

### 2.6.2 Primary emotions:

These universal emotions (which are termed "primary" by Ekman & Friesen 1984) are six in number: happiness, sadness, surprise, fear, anger and disgust. Ekman & Friesen (1984) report that the appearance of the face in each is common to all people, and that the primary emotions were found by every investigator in the previous thirty years who sought to determine the vocabulary of emotion terms associated with facial expressions. Blends of these occur. Other researchers use different labels.

Tomkins (in Izard, 1971) lists Interest-Excitement, Enjoyment-Joy, Surprise-Startle, Distress-Anguish, Disgust-revulsion, Anger-Rage, Shame-Humiliation, Fear-Terror, Contempt-Scorn. Izard (1971) refers to "major" emotions sub-served by innate mechanisms. These are interest, enjoyment, surprise, distress, disgust, anger, shame, fear and contempt.

### 2.6.3 Producing typical expressions:

Since the descriptions of Ekman and Friesen (1984) influenced the conception of this project, a precis of their method of determining which face actions are associated with each primary emotion is given below. They have put together an atlas of these expressions (Ekman & Friesen 1978) and also published a book entitled 'Unmasking The Face' (hence referred to as UTF) based upon it.

First they made a table listing what other researchers had published regarding either the face muscles involved or the surface appearance of the face for these emotions. With Tomkins, those face muscles not mentioned were also classified by their action under the six emotions. Evidence of the cross-cultural studies was helpful in this respect.

Models were instructed in the movement of particular muscles listed in the table. Their expressions for each emotion's musculature were photographed in three areas of the face separately: roughly the brow/forehead, the eyes/lids and root of nose, and the cheeks, nose, mouth and chin. These areas are capable of independent movement. The photographs of these face areas were assigned to the appropriate primary emotion. Composites of these made up varied typical expressions of happiness, sadness, disgust, anger, fear and surprise.

The resulting atlas was validated by demonstrating agreement with other subjective accompaniments of the appropriate emotion using the emotions experienced by U.S. and Japanese students watching a pleasant and an unpleasant film, in the cross-cultural experiment described above (2.6.1).

Those students had reported various emotions, and the researchers grabbed frames of each expression on the three areas of the faces of the students. The atlas expressions compared with these images predicted which film was watched for each expression regardless of culture.

The heart rate had been recorded while watching the films. The pattern of heart rate change was known to be different under emotions of disgust and surprise. The pattern when it occurred could be matched to the student's expression at the same time. The atlas expression for disgust or surprise should be concordant with the student expression. This was so.

In another validation study, atlas photographs were compared with independent face photographs which other investigators had used to display emotions. The interpretations in terms of the primary emotions, which the atlas suggested for these, agreed with those made by a set of judges who interpreted the same photographs. In the case of the atlas interpretations, these predictions were made on the basis of the three face areas with their separate predictive expressions. In a further experiment the atlas predicted the emotions modelled by another set of students posing the expressions of the primary emotions.

The atlas was found to agree very well with an independently produced atlas using a different method, (Hjortsjo 1970).

### 2.6.4 Context effects on expression interpretation:

Context can be defined as "the whole set of secondary characteristics of a situation or secondary properties of a cognitive or motivational state of an individual which may modify the effect of an effective stimulation (stimulus) or an oriented activity."(Tiberghien,1986).

The crucial question is whether emotion can be inferred from the face alone?. The question was addressed by Ekman (1977) who disputed earlier studies which appeared to show that context always contributes to and improves upon judgements made from the face alone. He suggested that these studies had neglected to take into account the variations in clarity (involving considerations of ambiguity, message complexity and strength) of the information of face and context. He advocated an experimental treatment which separated the judgements from face and context alone and from their combination. These are referred to as different types of "source". The type should be varied systematically or held constant across the experimental conditions. The effects of other factors such as the distinct emotions associated with each source, the pairing of sources to yield discordant and concordant combinations, and representative sampling needed to be all taken into consideration. Such a treatment has not come to light in the writer's search.

Several researchers report context effects in the experiments on recognition of faces. Recognition is facilitated by using the same context at test as at study (Watkins, Ho and Tulving 1976, Winograd & Rivers-Bulkeley,1977). This 'context effect' is important in episodic (autobiographical) memory. Episodic Memory is that aspect of the mind, or the brain, that makes the successful completion of individual acts of remembering possible (Tulving 1983). Episodic memory is more prone to context effects than semantic memory (memory for concepts). The Encoding Specificity Principle asserts that an object is recalled to the extent that the context at recall recreates that at the time of encoding.

Conway & Bekerian (1987) investigated emotions within the context of social situations and proposed that their findings could be modelled by a hierarchical organisation in memory of situational information (situations in which emotions arise).

Their findings suggest that representations of emotional knowledge exist at different levels of abstraction. One level would not contain specific experiential records but schematic or abstract knowledge of emotional experience.

At another level, the knowledge appeared to be shared across different members of the same society suggesting a consensus in emotional judgement.

In their proposed hierarchical organisation, knowledge of conceptual aspects of emotion is context-free, containing information about physiological/behavioural attributes. Such knowledge is likely to include the face expressions which achieve a high measure of consensus in society. Pancultural studies described above (2.6.1) suggest strongly that there are typical expressions for some common categories of emotion which can be recognized by literate and pre-literate peoples alike.

It would probably be true to say that a majority of the studies dating from Darwin's time which had as their aim the elucidation of the facial expressions of emotion have used face photographs alone without further knowledge of the context.

The work of Jensen (1986) and Sergent (1984) who obtained evidence for configural processing of faces and that of Ekman & Friesen (1975) who demonstrate the interaction between face actions in differentiating the signalled emotion might lead one to the conclusion that the face provides its own context. This is not to say that situational contexts would not play a role in the interpretation of the expression, but that the effects have not been systematically investigated.

**2.7 Emotion and memory:**

There are two main approaches to modelling memory: semantic memory incorporating knowledge about concepts and autobiographical or episodic memory which consists of all the events that one has experienced. Both of these are approaches to Long Term Memory (LTM). LTM has been modelled by semantic network, connectionist, holographic and mathematical models. Bruce (1988) has reviewed a selection of these models relevant to face recognition. The modelling of emotion and of its memory representation and organization has not received very much attention. An exception is a Level 1 & 2 theory of emotion-dependent recall which was proposed by Bower (1981). This will be described now.

**2.7.1 Semantic net representation of emotion:**

Bower et al.(1981) investigate the influence of emotions on memory and thinking by inducing these emotions in subjects by hypnotic suggestion. His results indicated "mood-congruent" salience of narrative events: the mood of the learner at the time of learning causes selective learning of material congruent with the learner's mood whereas the mood at the time of recall has no effect if the learning occurred in neutral mood unless it were sufficiently intense to become a context. He suggests that an emotion serves as a memory unit that can enter into associations with coincident events.

The relevance of the theory to the proposed thesis is in its explicit representation in memory of episodic context and emotions in a propositional semantic net format.

Bower's emphasis is on the experiencing of emotion by the person. He placed emotion events in a general semantic network model to explain the results. Semantics are about meaning. Human memory is modelled as an associative net of meaningful concepts and groupings of these for describing happenings. The basic unit of thought is the proposition . A proposition organizes concepts. An event is represented in memory by a cluster of descriptive propositions. Instances of the concepts involved become "associated" or linked together uniquely to record the event.

In the model the nodes are concepts and the clusters propositions. The links are the carriers of 'activation'. Some "happening" prompts a thought and activates its propositions. Bower (1981) found that recall is improved by reinstating the same mood during test as during study of test material. This mood-state-dependent retrieval is explained by the creation in memory of nodes representing the material remembered, the context and the emotion in which the material was learned. When asked later to recall the material, activating the emotion node by inducing the emotion in the person strengthens activation of the material node indirectly by boosting that of the supplied context given as a cue.

### 2.7.2 Organizing structures representing memory:

Semantic networks however could be inefficient in retrieval if they just kept growing which they would do unless chunking (Bartlett 1932, Minsky 1975, Schank & Abelson 1977, Schank 1982, inter alia) occurred.

Autobiographical memory may be represented in this way. Such chunked knowledge structures would have some generalization abstracted from individual experiences which would guide retrieval. The knowledge structure becomes the context (Reiser, Black & Abelson 1985) within which such search is carried out and is accessed by an event relevant to that context. Recurring temporal sequences of actions met with, become invested with so much familiarity as to become the expected rule. The exceptions to this rule then become the events that we pay attention to and learn in the sense of noting them in memory. Kolodner showed how these principles could be implemented in a computer fact-retrieving system, Cyrus (Kolodner, 1984), which automatically organized and elaborated input.

Reasoning in this way Schank and Abelson (1977) proposed knowledge structures called "Scripts" to represent chunked knowledge about ordered succession of events in the service of a goal e.g. the stereotyped ritual of obtaining food in a restaurant.

From the original interest in these knowledge structures as representations which serve in understanding stories, a Theory of Reminding has been developed by Schank (1982) which places Scripts in even larger chunks of memory called Scenes; and Scenes, in turn, in Memory Organizational Packets(MOPs). He extends the family of these conceptual structures to represent generalized scenes, and to Thematic Organizational Packets which organize themes, plans, and goals. Schank uses their functions to propose a dynamic memory.

The concept of dynamic memory increases the efficiency of autobiographical memory by automatic reorganization of its functional knowledge structures with experience. As more and more events are encountered, they are organised by recording ways in which events differ from the generalizations which typify all the events contained in a chunked memory knowledge structure. New generalizations

are formed when these anomalies recur.
The question which suggests itself is: Can emotion be an organizer of experiences in human memory ?. The findings of Bower et al. provide a basis for further investigation. There is some experimental evidence that suggests humans may have difficulty in accessing events in memory if asked to recall those associated with a particular emotion (Robinson 1976; Reiser, Black & Abelson 1985; Dyer 1983).

Robinson (1976) found that humans access experiences in their memory faster when primed with activities than with emotion (cf Conway & Bekerian below). Reiser, Black & Abelson (1985) requested subjects to recall events associated with a particular emotion and report that, in response, they appear to recall a general activity in which the emotion was experienced and then enumerate instances of that activity, rather than different emotion episodes themselves. They conclude that information from other sources is likely to be necessary to discriminate experiences within an emótion. Dyer (1983) also found a similar effect in his subjects recall from narrative.

Conway and Bekerian (1987a) suggest four types of knowledge of emotions:

1. Semantic, context-free, conceptual, physiological and behavioural knowledge of categories of emotions,

2. Emotional "scenes" (analogous to those described by Schank (1982) containing general precipitating and accompanying. circumstances of an emotional reaction e.g. "accidents" causing fear etc.,

3. Emotional "scripts", detailing specific situations, and
4. Autobiographical memories of emotional experiences.

They show how these can be conceived as forming a hierarchy of emotion knowledge for each emotion with three levels: context-free, basic and specific. Basic level subgroups of emotions, they say, may be organized in terms of scene similarity e.g. illness and accidents clustering fear, terror and panic. This level, organizing what they classify as "scene" knowledge would appear to be similar in level to the class of response that Reiser, Black & Abelson (1985) obtained (i.e. "studying for examinations" as a response to a request to recall events involving the emotion of "ambition").

Schank (1982) conceives a higher conceptual chunking: the Memory Organization Packet as the memory structure which organizes scenes in autobiographical memory. Conway & Bekerian (1987b) do not appear to include these in their tentative hierarchy of situational knowledge of emotion but their work in autobiographical memory leads them to suggest a hierarchical organization of personal information ranging from abstract to more specific levels. They propose lifetime periods (A-MOPs) index general events (E-MOPs) and these index specific autobiographical memories. Their finding that personal memories "rarely came to mind in response to situational cues" leads them to suggest that scripted information may be accessed independently of the recall of associated autobiographical memories. Emotion terms used as cues however facilitated retrieval from memories of autobiographical events in contradistinction to Robinson's finding above and suggest that a combination of a situational context and an emotion name caused the facilitation.

The role of situations in the representation of emotions is emphasized by Conway & Bekerian: subgroups of emotions are determined by situation attribute overlap. Attributes are places, people and activities associated with the discrete emotions. The scenes, scripts and autobiographical memories form different types of situational knowledge. These probably differentially facilitate the recall of emotional experiences. But they are closely bound to the subgroups of emotion at the basic level whatever level of knowledge is involved. Their correspondence to the hierarchical levels of autobiographical memory suggested by Schank, suggests to Conway & Bekerian a similar hierarchical organization for emotion in the human. In this, the knowledge of emotions may be organized in terms of overlap of activities, places and people judged to be important determinants for the experience of emotions and in terms of the potency of different types of knowledge for inducing moods.

Although memory for faces has developed through an individual's experience and is therefore at least partly autobiographical and episodic, the theoretical frameworks have not been conceived in terms of its representation in these chunking structures which Schank and others have conceived. The work of Conway and Bekerian breaks new ground in this direction for emotion, but no model has appeared for the representation of perceived face expressions in memory: an omission which the present project aims to correct.

Episodic theory stresses the uniqueness of each instantaneous event in one's experience and hypothesizes that a memory trace is formed for it at the time. This is separate from the semantic memory. Both, however, have been modelled by semantic network. The event will be recalled to the extent that the retrieval information overlaps with the trace. The principles underlying recall are the same as for recognition.(Watkins, Ho & Tulving (1976). Semantic net theorists sometimes assume that a face as a whole is equivalent to a word in its representation, e.g. as a single node or a cluster of these and nodes are tagged with a context.

For the representation of expression there needs be an abstracted level of representation since the concept of (say) a smile can be abstracted from any individual's smile. Such abstractions could serve as organizing structures for emotional expression in episodic memory comparable to Conway & Bekerian's situation structures and Reiser, Black & Abelson's Activity and Actions structures. These would organize individual facial expressions by indices which allow their efficient retrieval. The organization will be built on memories of particular expressions viewed throughout the life of the individual and would include any context associated with the episode which contained the expression. As in Conway & Bekerian's formulation, where there is a level of representation at which discrete emotions are context-free in the sense that they do not contain information relating to separate autobiographical memories, one would expect a similar level of context-free expression representation at top level corresponding to semantic categories of emotion.

One characteristic of such a representation based on episodic memory is that it should show similarity effects i.e. there is more accurate response to similar than to dissimilar episodes.

**2.7.3 Network models:**

In some of these a face is represented not in discrete memory locations but as a pattern of interaction (weights) between many units of a network (hence the term "connectionist"). The emergent properties of such a network (McCllelland and Rumelhart,1985) can simulate the function ascribed to face recognition units (see 2.4.4). Wisard (Stonham,1986; Aleksander 1983) is a recognition device with a video camera front end which samples random but fixed sets of pixels from a frame and stores a "1" at the RAM addresses suggested by the arrangement of black and white pixels on the sample until the entire frame is sampled. The procedure is repeated for many instances producing an overall pattern of 1 and 0 in a discriminator. Fresh test instances will score a percentage resemblance to the maximum possible. Wisard can be trained to recognize the difference between a smile and a frown irrespective of the person. Wisard seems to fulfil the function of a face recognition unit and does this from training on many instances.

## 2.8 Relation to this work:

1. The expert knowledge with which the present system, Janus, is based on the primary emotions of Ekman et al. which initially filter the input of a set of face actions.

2. The context is limited to the co-occurrence of actions of the inner features of the face

3. The units of face measurement are at two levels: geometric and syntactic, the first being mapped to the second.

5. A dynamic memory classifies input expressions under these six primary-emotion typical expressions and creates new more specialized sub-classifications on the basis of repeated atypical input.

6.) Emotion labels may accompany face actions as input. Subsequent retrieval of these labels is effected by retracing an identical path in memory by new input. This will occur when an identical expression is input without a label.

**Summary:**

The review of research presented in this chapter has highlighted some interesting directions in which memory, emotion and face research has travelled. For a system to interpret expressions, the concept of Dynamic Memory could be adapted to organize the classification, and retrieval of the face actions associated with emotion. Also, the function which emotions are believed to serve in relation to motives and goals confirms the view that their representation in intelligent computer systems is worthwhile. In the next chapter, a review of existing face systems is given together with a statement of the questions addressed in this thesis.

# Chapter 3:

# Systems that process faces

## 3.0 Systems that measure faces:

This chapter describes some systems both computer and non-computer which are centred on the face. The main object is to see whether the proposed system has been implemented already by some other researcher either in its overall aim or in its design. Those described range from geometric/syntactic measurement systems to aids to facilitate the matching of a target face description to one of known identity in a database ('mugfile') as well as attempts to infer emotion. The term "measurement" has been applied both to recording locations on the face in coordinates relative to some geometric axes and also to the reporting of syntactically-described facial movements related usually to the major features. As noted in Chapter 2 in reference to pancultural studies, research groups have their own empirical methods for describing and validating the message carriers in terms of the components of expression. More recently the Ekman Group have turned to the description of these movements in terms of the causative muscle groups sometimes artificially stimulating these so as to record the facial result.

The geometric measurement technique has found adherents in the Pilowsky Group(1986,1985) and with Thornton & Pilowski (1982) for determining the best differential distances for a restricted range of face actions. Kaya & Kobayashi(1972) have used hand measurements of distances between points on faces to define 9 geometric parameters and to calculate the amount of information carried by them. This was used to suggest a classificatory algorithm for recognizing a face from a population of faces with a high probability of success if measurement noise is almost negligible.

The automatic measurement of facial features would find a ready application in retrieving matches from computerized 'mugfiles'. Such measurement of contours of face features on digitized images of faces has proved very difficult but the work of several researchers will be reviewed. Automatic measurement to the degree of refinement and discrimination required for detection of the expressions of all the primary emotions however has yet to be demonstrated.

The specification in the pertinent literature of particular expressions as signalling specific emotions should be regarded as a theory of the particular researcher involved. Empirical studies sometimes appear to involve judgements based more on part area or whole face expressions rather than at the level of individual skin structures e.g. a rise in the lower eyelid. Ekman and Friesen(1975 & 1984) provide a description for each of their primary emotions which is used in this project.

In the succeeding paragraphs some of these face systems will be described and judged in their relevance to the prospective project.

### 3.1. Research and tutorial Systems:

#### 3.1.1. The facial affect scoring system (FAST):

This was the first face system from the Ekman Group (Ekman, Friesen & Tomkins, 1971) a non-computerized system constructed as a tool to predict observers' judgements of emotion in terms of the six universal categories: happiness, sadness, anger, fear, surprise, and disgust; using wrinkles, and descriptions of tension or relaxation in specific features and also descriptions of positions of features. The descriptions were made in the three areas:brows/forehead; eyes/lids/bridge of nose; and cheek/nose/mouth/chin/jaw. This was to prevent judgements of components being influenced by the whole face and also to emphasize their functional characteristics. Intensities of the emotions were represented. The choice of components within these areas was based on the empirical findings and theories of previous writers in the field and the group's observations and intuitions.

Models were photographed while producing the particular appearances in each area of the face. Verbal descriptions were attached to indicate head orientation and gaze direction. The three areas of the face were each represented in a set of photographs. A scoring procedure was constructed which could be applied by judges matching a test photograph area by area to the standard sets. The system was validated in this way, using a priori procedures to combine the results of three independent trained scorers into a single set of scores and also to determine whether there would be either no prediction, a single emotion prediction, or a blend prediction of an emotion category. The reliability between scorers was considered satisfactory and 45 out of 51 faces were correctly predicted by the system in the first validity study with only fear among the categories attaining a low percentage hit rate. Problems were encountered with permanent age-wrinkles mimicking actions, and individual differences in facial repertoire. Since this is a mostly visual match technique, it will have relevance to the present project only if that technique is adopted.

**3.1.2. The facial action coding system(FACS):**

Criticism of the methodology of FAST came later from Ekman himself(Ekman,1982). The method had required deciding which face actions go with which emotion but then used them to discover the function of those very actions. Another limitation was that the system could not find out any signs it did not already know about.

The basis for deriving units of face expression were considered to have special importance for Ekman(1982) because there was much variation in this factor among published research and often the basis for selection of such was not stated. He considers that the lack of an accepted, standard ready-for-use technique for measuring facial movement was perhaps, the most important obstacle to research in this domain. Ekman and Friesen's Facial Action Coding System (FACS) sought to redress these limitations and to widen the scope from just those movements signalling emotion to all expressions.

In planning FACS, Ekman & Friesen set out to represent all the actions of which the facial muscles are capable provided that these actions could be reliably distinguished from their facial appearance, otherwise they were combined. They systematically recorded the change in muscle, identifying it when necessary by inserting a needle into it, voluntarily contracting the muscle and verifying that electrical activity can be recorded from the needle. By this approach they claim to have produced a comprehensive measuring technique in the sense that it claims to be able to record all visible facial actions. In support of this they produced 7000 different combinations of facial muscular actions on their own faces. The following account is based on Ekman(1982), Ekman & Oster 1979, Ekman & Friesen (1976a,)

The intensity of some actions was represented. Some emotions in their theory are signalled by overlap of different muscles, and overlap can involve the coincidence of any phase of muscular contraction: onset, offset and apex. Instruction on measuring these aspects is included in the system. The method can be used to describe any facial movement whether displayed on still photographs or video.

Facial Expressions are described in terms of anatomically based minimal Action Units (of which there are 43), and in combinations of these, detectable on the face, each of which is described in non-inferential terms that describe the facial behaviour. This excludes descriptions which, like 'frown', 'grimace' or 'leer' appear to infer a prejudged psychological state. Each Face Action Unit is illustrated by still photographs and filmed examples, in a self instructional package published by the authors. The task for the user is to learn the muscular basis of face movements and the changes in appearance they produce, described in words and diagrams. Instructions are included on how to reproduce a unit on the face (to aid learning) and for the minimal changes necessary to score it present. The scoring is taught with practice faces. Relationships between Units are scored according to rules addressing overshadowing, overcoming, similarity and impossibility of co-occurrence. When single Action Units and the rules governing their combinations have been learned, it is quite possible to decompose complex expressions into these elements. The scoring system is in terms of these. The training can take many hours, and in use may require repetitive examination of motion film or video. It is thus unsuitable for real-time coding(Ekman & Oster 1979. p 540). A version of FACS which measures single emotions has been produced by Ekman & Friesen(Ekman 1982), and measures the occurrences of actions typical in happiness, distress and/or sadness, fear, disgust and/or contempt, surprise, and anger. Strategies have made the scoring more efficient than the parent system: the confirmed presence of one of the typical actions focuses the search for the rest. Intensity and timing are not included.

Comment:

Tremendous achievement though FACS certainly is, it is designed with the human trainee in mind and relies upon the latter to identify the presence of a action unit. This requires a fair amount of training and one could not see the system in its present form becoming part of an expert system. One could perhaps anticipate Wisard (Stonham 1986, Aleksander 1984, 1983) being trained to recognize a reduced number of Face Action Units, with a back up module to deliver an emotion interpretation.

**3.2 Computer-aided investigations of face perception:**

**3.2.1 Quantifying expressions of emotion:**

Quantifying the face by measuring distance parameters was employed by Kaya & Kobayashi(1972) as described above. They used facial distances to classify types of faces rather than to classify emotions and one would expect that distances chosen to classify emotional expressions would be different, reflecting more, perhaps, the pattern of the inner features rather than the "fatness" of the face.

Thornton & Pilowsky(1982) measured distances which would indicate the position of a target face along a happiness/sadness dimension. Using microcomputer graphics facilities, they constructed a line drawing of a face from linking sixty strategic locations measured on a face photograph. They represented each facial action muscle group contributing to the same action as a straight line(muscloid) which could be shortened along its direction of action towards a fixed point at one end representing the fixed attachment of the muscle to the face bone. The shortening end would be the muscle's attachment to the skin of the face at one of ten selected points. These points are connected to one another by skin in the real face so that one muscle group's shortening can affect another point by dragging the skin, and a similar effect was contrived by links between the points in the model e.g. the narrowing of the eye in smiling. Using posed photographs, they created a line-drawn image of the face to match by manipulating these muscloids. They depicted extremes of happiness and sadness (influenced by the description of these emotions in Ekman & Friesen (1975). They chose an arbitrary neutral face to reference the interpolated expressions. The scale value is obtained as a function of the activation of the muscle groups. From a neutral expression the model can traverse the range of expression to each extreme.

The authors saw the model as being useful in recognizing and estimating mood. Although the 1982 publication deals with the one dimension, that of happiness/sadness only, the authors believe extension to other axes should prove straightforward.

Pilowski, Thornton & Stokes(1985) were interested in relating facial measures to specific expressions(a smile) so as to identify those with the closest relationship. For purposes of comparison purposes between faces the distances were normalized by dividing vertical ones by nose length measured on each photograph, and horizontal ones by the distance between the external angle of the eyes. Faces were digitized during a smile sequence and three distances(mouth-width, end-lip-raise, and eye-opening) measured. The first two measures were highly correlated (Spearman correlation coefficients) with the degree of smile. Twelve distances were compared for posed expressions of happiness, sadness and fear. Mouth-width was greatest, and end-lip-raise was shortest for the smile. The twelve distances measured include distances between apices of eyebrows; between inner ends of eyebrows; upper and lower lid iris intersection of right eye; width of mouth between angles; right angle of mouth to external angle of right eye; between facing margins of upper and lower lips in centre; width of upper and lower lips in centre; between internal angle of right eye and centre upper margin of upper lip; between internal angle of left eye and centre upper margin of lower lip; length of nose to mid brow; between upper and lower eyelids at centre.

In their 1986 paper they compare face measures from a smiling face with their counterparts in a neutral face and find that those which differ most are "lower eyelid iris intersect", "mouth width", "end-lip-raise" and "mouth opening". They look at the possibility of using the Ekman & Friesen pancultural expressions of the primary emotions as templates against which to match any facial configuration and quantify a degree of match. The authors envisaged the model as being useful in establishing criteria, norms and a taxonomy for expressions and in human-developmental and medical monitoring.

The approach is relevant to the present project in adopting a geometric measuring approach to representing emotional expression and also in seeing in the pancultural phenomena a potential for their use as templates for the primary emotions. The present project was conceived independently.

In relation to face points chosen for measurement, The Pilowski Group papers do not explicitly explain why these particular key face locations were chosen and it is of course possible that other measures could give as good or better correlations with the degree of smiling. Also, as the authors point out, the measures that distinguish a smile best may not be specific for a smile.

The normalization procedure which allows a face from one person to be compared with that from another by dividing vertical distances by the nose length and horizontal distances by the bi-ocular width may be compared to Kaya & Kobayashi's normalization procedure of dividing all the parameters horizontal or vertical by the nose length. Such normalization is aimed at neutralizing variation in size of the different faces. It is assumed that the length chosen does vary directly with size of the face and that it is readily measurable. In dealing with expressions, it assumes that it is not altered in length by the expression. It may be that the external angle of the eye is difficult to locate in expressions where the brow is lowered since the upper lid covers it. The distance between the inner angles of the eyes is perhaps more readily located.

It is of interest to compare those distances chosen above for representing emotion with the 9 distance parameters Kaya & Kobayashi(1972) used for classifying faces for identification: Three are the same: the external bi-ocular breadth, the length of nose, and the mouth width; two were collapsed: each lip's height into one mouth height; The eyebrows, iris / eyebrow intersections, and mouth/eye oblique distances and open mouth gap are not represented in Kaya & Kobayashi whereas the measures of the breadth of the face at two points and the vertical distances between upper lip and nose and lower lip and chin find no counterpart in the Pilowski representation. It would appear from Kaya & Kobayashi's census(1972), upon which their choice of parameters rest, that the characteristics of faces which impressed their informants were those which represent the height and breadth of a face rather than the relationships between inner features which characterize emotion.

Townes(1976) also used ratios of vertical and horizontal distances as parameters for representing faces in a sequencing algorithm and used a least squares distance between database ratios and those of the target as a similarity measure by which to order retrieval of similar faces which might contain the target identity. His parameters were similar to those of Kaya & Kobayashi except that he did not use mouth height preferring the distance from the lower level of mouth to pupil level. Townes also measures the distance to the chin from the pupil level.

Another application of quantifying faces is the interactive computer system for retrieving faces from a police database system, FRAME (Shepherd 1986). FRAME codes faces in 50 parameters: many of these are coded in terms of direct physical measurements from an image. The face points used were based on the work of Jones et al.(1976). Seventeen of these were spaced around the outside of the 2-D image of the head, the others were located at: the ends of both eyebrows, the angles of both eyes, the eyelids above the pupils, the sides of the nose, the junction of lips in the midline; and three for the hair outline. Each face was coded in the coordinates corresponding to these points by projecting its image onto a bitpad and touching each point with a stylus. Attributes (about the general shape and individual features of each face) were rated on a 5-point scale mapped to corresponding 5-point scaled line and shape measurements derived from these points. The 50 parameters representing a target are matched to the population in a databank to produce an order of retrieval according to similarity.

### 3.2.2 Generating Caricatures:

The previous quantifications representing faces are modest in comparison with the 186 points with which Susan Brennan represents a face in her Caricature Generator - a program which generates caricatures by comparing a to-be-caricatured face(TBCF) with an averaged face (made up of the averaged values of the points on a number of faces) and exaggerating the differences. The points are connected by interpolating curves and their choice was made with the principle of the minimum required to produce a recognizable face. The lines are organized into 39 face features.

### 3.2.3 Computer aids to face research:

The principles underlying perception and recognition of faces are still poorly understood. Screening and manipulating images quickly and easy storage of the images can be of help. Several techniques have been published(Jensen 1987). Jensen photocopied and digitized composite faces and individual features using Identi-Kit superimposable foils. Once stored he manipulated these features with graphics editors on a Mackintosh computer. Specialized software packages can be purchased which do geometric transformations such as screening chimeric faces made up of two left or right sides.

Some previously unreported findings relating to the salience of the inner eyebrows and upper lip in face recognition were discovered by Haig(1986) using an image-processing computer with secondary storage. He sees the advantage conferred by a computer system as enabling the researcher to tackle different tasks demanding large data collection. Also a high reproducibility is obtained. Examples of the tasks undertaken with this equipment include:

1. Measurement of the sensitivity of subjects to slight positional changes of prominent facial features ("mouth up" was greatest);

2. Constructing a face feature saliency by interchanging features (for example different faces or features could be substituted on different or the same head. The results suggested that head outline was most salient followed by the eyes, whereas the mouth and nose played little part in frontal face recognition. They do in emotion recognition);

3. Whether identity recognition depended on other parts of the face not usually thought to be salient to recognition. By masking various individual parts of the face, saliencies were found. The surprise was the hitherto unsuspected salience of the inner eyebrows and the upper lip area.

Haig sees his results as reinforcing the objection to simple feature-listing as a means of understanding the recognition of faces.

### 3.3 Systems aiding criminal investigation:

Witnesses to crime aid police in describing the faces of the law-breakers. The witness may search the "mugfile" (collection of photographs of known criminals), but if a positive identification cannot be made there are aids to helping the witness recall the face. A police artist may draw a likeness to the specifications of the witness, or the witness may use one of the marketed systems to reconstruct a face appearance by choosing different interchangeable feature-examples from a set ( Identi-Kit and Photofit). Computer aids to recall a suspect allow witnesses to modify stored features or to build up a likeness from an average face.

#### 3.3.1 Computer systems which retrieve faces from databases:

Computer systems have been devised to speed the retrieval of similar faces to that obtained by one of these aids from a database of faces on the basis of a pattern-matching program. Another method is to code a verbal description from the witness's memory in the same format as the faces in the database are coded(e.g. the Home Office "Faces" system where the description of each feature is coded on a five point scale which is entered in numerical form and amendments can be made on the basis of retrievals made by matching). Shepherd's system Frame, described above, is of this type. That of Rhodes & Bargainer (1976) is another example. They devised a system based on measurement of distances on an image derived from a witness who recalls the face either with the help of an artist or using Identikit. Using an algorithm devised by Townes(1976) ratios of nine vertical and horizontal distances on this image are compared with those of each image in a mugfile to produce a similarity rating. The output is made up of the mugfile images in the descending order of similarity between the 'mugshot' and the target image. This is determined by a 'least squares' function of the Euclidean distance between the ratios of the target face and database faces.

Such aids to identification are of especial interest to cognitive psychologists because they provide a means of studying such questions as whether humans learn faces holistically or feature-analytically and what configurations of features are important(Jensen 1987) and what features are salient for recognition(Haig 1986).

Since they are widely used, the question of whether training in paying attention to features increases correct recognition has prompted research(see Baddeley 1979). Whether the human uses a holistic or feature- analytic strategy would influence the decision of what face primitive is represented in a face recognition model and is also pertinent to the representation used in this project. As noted earlier, there is evidence that both the configurations and features can be used by the human - perhaps for different tasks.

### 3.3.2 Non-computer aids to face recall:

There are two such aids in current use: Identi-kit and

Photofit. With one method, the witness is asked to recall their impression of the face to be reconstructed (FTBR). An operator constructs a face from the components which is then refined by attention to detail. With the other, components are selected one by one, choosing the nearest fit from the kit. Original Identi-kit has line drawings of hair, eyebrows, eyes, nose, lips or ears/jaw and so on. Each is on an acetate transparent sheet which can be superimposed on other sheets to build up a face. Photofit(Penry 1971 ) has photographs on pieces of card of hair-styles, face shapes and other features such as eyes, noses, mouth, jaws and cheeks

The following points were raised (among others) in a discussion in a BBC radio programme: "About Face" featuring Hadyn Ellis and John Shepherd then of Aberdeen University: The reconstructions can only reproduce types. They may not produce a good likeness. Whether the sheer variety of component features obscures the recalled impression is not known. Choice of the right face feature is only one task with which the witness has to cope. The features have also to be placed in the correct relationship to each other. Even minute differences in distance between features can destroy the similarity with the FTBR.

The limitations of the effectiveness of these aids are due to people being poor at recalling faces. They are much better at recognizing suspects from "mugshots".

The use of one of these aids has been investigated by Ellis et al.,(1975). Davies(1986) describes some newer aids. These include a photographic version of Identikit with more realistic features. There is also Magnaface: a magnetic-backed natural coloured composite of features which is overlaid with a blending film and can be "finished" with cosmetic pencils.

### 3.3.3. Computer aids to face recall:

The BBC's Videofit is an application of the standard Quantel 7001 'paint-box' computer graphics package to an electronically processed Photofit picture. A computerized system influenced by Photofit, E-fit developed at the Home Office allows one to manipulate an unlimited variety of facial features into any desired configuration on a video screen. Another computer system SKETCH (Rhodes & Klinger,1977) modelled on Photofit and Whatsisface (Gillenson, 1974, Gillenson & Chandrasekaran,1975) builds a face from stored components on a minicomputer and uses an interactive language with imprecise or fuzzy input via the keyboard and defines these in terms of drawing routines which alter these components. Sketch can converse with the user through screened messages. Continued discourse about a feature is possible even in the absence of its name (or substitution for it of a pronoun) because the context of a feature template is retained until it is changed. The Whatsisface system achieved modification of its 17 basic face features through the user-operated key presses and analog devices with dials which controlled transformations of the graphic line features of the initial average face and the dimensions of the latter itself.

### 3.3.4 Automatic measurement of faces for retrieval:

Laughery et al. (1981) describe the early work done by Bisson(1965a,b) in this direction but they accredit the first successful automatic measuring algorithm to Sakai et al (1972). They describe a related feature-measuring algorithm developed by Bromley(1977).

The relevance of such a facility to the present project is limited, since a front end visual system is considered beyond its scope. The intention is first and foremost to provide a transduction of measurements on the face to the emotion signalled and to that end, would anticipate that a front end could deliver the face points required by the rule base.

In each of the above methods a full face photograph was digitized and a line drawing obtained by thresholding the rate of change of the intensity of illumination at every point relevant to its neighbours so that only the very abrupt changes are left as dark pixels against a back cloth of white.

Sakai et al. use a 9 x 9 Laplacian two dimensional secondary differentiation operator on a face digitization array of 140 x 208 picture elements with 32 gray levels. To detect the necessary face and feature contours they introduce a variable sized window slit which is applied to the binary image in steps: horizontally for the top of head and cheeks and nose tip and vertically for the nose, mouth and chin and radially for the chin contour. The number of dark picture points in the slit is the metric for locating the contours of the inner and outer features. As the slit moves from one side of a line to the other the number of dark picture elements goes from 0 to a positive number and then to 0 again.

The histogram profile produced from the number of dark picture elements in each slit across the depth( or width, if the slit is vertical ) is called the integral projection of that slit at that position. Because of characteristics of different peoples' appearance, these are distinguishable in form from those of other people and can be used to classify faces. Their use in automatic measurement is that they identify points on the face which can be expressed in terms of xy coordinates from which distances can be measured.

The narrow slit is then moved down or across the face. The order of search for Sakai et al. is: head, sides of face, nose mouth and chin contour, width of eyes and nose, and axis of face. The order for Bromley is the left and right sides of the face, the centre line, the top of the head, the hairline, mouth, chin-line; eyebrow and eyes, the tip of nose and facial outline. The Sakai program compares shapes of integral projections by coding them in symbols for the types of peak and its length and comparing these with various standard stored shapes which can be added to. An exact match pinpoints the feature: no match means re-tracking to the previous detected point and there is room for relaxation of match criteria in between these results. The plan of the method is to locate the area filled by each feature first then to refine the processing within that area.

The results in classifying faces are very impressive: giving correctly all the sought feature points (sides, nose-mouth-chin, eyes, nose width and face axis) in 552 of 607 photographs. However, judging from the binary pictures illustrated, without further refinement, the prospect looks uncertain for the usefulness of this approach to the measurement of those facial features which Ekman & Friesen(1984) describe for the primary emotions ( e.g. It would appear very difficult to measure the amount by which the lower eyelid is raised in comparison with a neutral expression: such difference is of the order of a few picture elements). Again the message appears to be, as it has at several points in this short review, that the feature information required for classification of people differs from that required for classifying their expressions and is more in terms of static(Ekman & Friesen 1984) signs. Buhr's algorithm(Buhr,1986), like that of Kaya & Kobayashi(1972) and of Sakai et al.(1972) give very impressive success rates in classifying faces and perhaps could be refined to detect the fine measures required for classifying emotional expressions. The field is still active(Craw, Ellis & Lishman(1987); Wong(1989), Petajan(1985)).

A newer approach to classifying faces is to extract the variance in a set of faces independently of any judgment of features, A set of Eigenfaces(analogous to eigenvectors) characterize the variation among faces. A single face is represented in terms of a linear combination of these eigenfaces(Turk & Pentland,1990, Abdi, 1986).

**3.4 network models:**

Although these are not generally purpose-built for face input some of these sytems in which the information stored is represented in the strengths of the interconnections between the units of the network (connectionism), can learn and retrieve faces. Bruce has reviewed these in respect of their competence and theoretical implications and it is upon her review that the following brief account is based. The McClelland and Rumelhart(1985) model is able to form prototypes from many inputs which themselves deviate little from the prototype. Then it responds more strongly to the prototype than to distortions of it but also responds strongly to its more recent training instances. It is natural to surmise that aspects of face recognition in the human could follow the same form e.g. different views of the same face as training instances labelled "George" producing an effect like that of a face recognition unit, or different persons' faces labelled "person" producing a "person prototype". However, what would happen when the same sad expression were used as training instances with varied labels: gloomy, miserable, unhappy, depressed etc. each accompanied by the additional label "sad", and then independently, different expressions were presented labelled 'gloomy', 'miserable' etc. without the additional label 'sad'?. Perhaps the strongest response would be to the expression nearest to the prototype formed for 'sad' but also a response to close matches to each instance: 'gloomy', 'miserable', 'unhappy' and 'depressed'. Such response seems intuitively plausible in the human as long as one acknowledges that one could have prototypes for 'gloomy', 'miserable', 'unhappy' and 'depressed' also.

Kohonen(1981) using digitized faces as stimuli, taught his distributed network different orientation poses of the faces of 10 persons and obtained identity-specific responses to a novel view in the same range of orientation. The face image is represented as light intensity values for each picture element in a vector in which the relative positions of these are maintained, from which an autocorrelation matrix is produced. This, according to O'Toole et al., 1988, is a correlation of all possible combinations of the vector elements encoding the intensity relationships of all possible spatial positions in the image. Autoassociative learning produced the correct person even with incomplete or degraded stimulus of the person.

'WISARD'(Aleksander & Burnett, 1983, Aleksander et al., 1984. Stonham, 1986) is a general purpose object recognition device, based on neural net principles, which can be taught to recognise faces. The training instance is digitised by a video camera and thresholded into a binary image. A 'tuple' of from 2 to 8 picture elements are selected from the binary image using a random but fixed mapping and the whole image is sampled. A 512 x 512 binary image contains 2**18 bits and for n = 8, 2**15 n-tuples are sampled(Aleksander & Wilson). The sequence of 0 = black and 1 = white in each tuple makes an address to a ram element each of 256 bits.( The 2**15 RAM elements, each of 256 bits are implemented using much larger RAMs. RAM arrays organized as k-bit words provide k 'discriminators' simultaneously ). In training with examples and their label, a 1 is stored at all the locations addressed by that input image within the 'discriminator' selected by that particular label. In testing, if the image is presented again. a '1' is produced at all its data output terminals of that 'discriminator'. Graded responses where not all the output terminals produce a '1' result from similar images. The label of the 'discriminator' with the strongest response is output as well as it's actual response and also the difference between this response and the 'runner-up' to the test image. The test face is given a statistical histogram corresponding to its similarity to the learned patterns contained in discriminators representing the response of learning to various images, and the label is output if a significant match is present. The match is a percentage of the number of 1's achieved by the test object compared to the number of 1's which would result if every tuple from the test object addressed a memory location where a 1 was stored. The label associated with the input could be a primary emotion-term or the name of the person. Wisard trained on a framed full face to a 95% response would give a much lower percentage match if only 25% of the face were tested in the frame, and the response to a profile in the frame would be a poor match also. Wisard can "window" an area of the face thus restricting input to that area. After training on a window around the mouth area an upturn or downturn can be discriminated. Although Wisard, taught to discriminate a frown from a smile, may generalise that discrimination to a person on whom it has not been taught, it has yet to be shown that teaching the expressions signalling the six primary emotions could be so generalized so as to discriminate these successfully on all faces.

**3.5 Systems designed to improve Man-Machine Interaction:**

The idea behind these systems is related to the aim of this proposed project: to enable computers to understand the non-verbal behaviour of the user so as to facilitate the mutual interaction.

The first system described: "Head Reader" is designed as part of a "future total human motion understanding system" by Mase, Suenaga & Akimoto(1987) which coordinates the positioning data from face parts with that of head motion so that a lip reader will find the mouth of the user. The second system(Sheehy, 1989) is part of an ESPRIT project to implement dialogue with the computer user. These systems are described briefly below.

Head Reader extracts 3-D positional information from digitized images of the user and interprets it as movements: "tilt up", "tilt down", "turn left", "turn right", "forward", "backward" etc. A "yes" or "no" can be interpreted from these. A 3-D face model constructed of triangular patches can rotate around three orthogonal axes. Head movement is roughly detected by the ratio of face area to head area. The movement is replicated on the model, and the latter overlaid on the input image of the user so the possible positions of the face features are delimited for refinement. The movement is interpreted by noting whether the movement along and around the 3-D axis exceeds pre-defined thresholds. Flags are set in the six cases that the corresponding motion is true. A "yes" is when both "up" and "down", and a "no" is when both "right" and "left" are set within a prescribed period. The recognition rate was 63% with 31 samples and 93% with 28 learning samples.

As part of its programme to engage users in a dialogue, the ESPRIT Project on "Communication Failure in Dialogue and Techniques for Detection and Repair" detects non-verbal behaviour in the form of user signals of surprise, puzzlement and doubt and interprets them as requests for more information(Sheehy,1989). Head nodding is also detected and interpreted as 'handshaking', requests for turn-taking, and pacing the flow of information. A video camera supplies a digitised image of the computer user which is 'windowed' on the top of the head, eyes, sides of face (to detect head movement). To correctly position these the user must not slump. The camera is positioned over the VDU. Frames are grabbed while the user looks to the screen, to the right and left, down and also when absent. A template is stored for each of these user-positions and the image processor can respond to the dialogue manager's enquiry: whether the user is present, is looking at the screen, has nodded or shaken the head.

### 3.6 Conclusion:

The prime object in making this review was to see if the goal of the proposed system has already been achieved in an existing system. Only the tutorial systems deliver an interpretation in terms of a category of emotion and these differ from the proposed system in the following ways:

1. They are not computer expert systems.
2. They do not convert a geometric input into an emotional term. 3. Their object is to train people to make face discriminations. 4. Such knowledge has no representation in software either in terms of face geometry (although the scoring may be automated) or in automated methods for transducing this into emotional categories.

The Pilowski group research is closest to the proposed system in its aim: that of quantifying the Ekman & Friesen(1975) expressions to make templates for their recognition; but in its published form it is not a computer system applying templates for all the primary emotions.

Wisard possesses the generalized learning and discriminative capability to discriminate emotions. It would need to be trained on typical expressions and the corresponding labels of the primary emotions. It is uncertain whether it would be able to recognise and label the primary emotion expressions on any presented face.

The connectionist systems likewise, it would appear, have the potential to discriminate emotions (although to the writer's knowledge there are, to date, no published reports of this), provided the input features are carefully selected.

Automatic measuring of face geometry is essential for a running system of the type envisaged in this thesis, even though not included in the proposed prototype. There is however a doubt whether it could identify all the needed features. Some of the image areas required for discrimination of the primary emotions may be very noisy. This is especially likely for the eye regions where shadows occur and fine measurement is needed.

**3.7 In the light of this review, the aim of the present project is to test the following hypothesis:**

It is possible to quantify face expressions and produce machine-made interpretations of emotional states from such quantities alone, such that these interpretations are rated no worse than those provided from humans beings judging the same facial expressions.

**3.8 Possible Applications:**

Janus is an experimental system. It is an investigation into a measuring approach to face expression categorization. Applications for a computer approach to face expression categorization might include any situation in which a computer is purposefully scanning human beings and coordinating information about them from various sources such as speech, posture, behaviour. Non-verbal dialogue is less specific than verbal dialogue (i.e. the depressive posture cannot give the reason for itself, but its sudden lifting when, say, the loved one hovers into sight, can give enlightenment). Head movements can pace and signal turn-taking in dialogue - functions which the Esprit researchers (above) are attempting to emulate in human-computer interaction. The task for humans and computers alike is to infer the possible junctures in the goal planning that the emotion signals. It is quite likely when the dialogue is between a computer and the user that the emotion felt and expressed by the latter are triggered by significant junctures of the "mutual" multi-goal planning sequences that the user has in mind involving the computing task. The user may signal happiness if a procedure works( a sub-goal achieved) or sad if it fails, anxiety as time runs out, etc. Because of the one-to-many relation which exists between a facial expression and possible causes, it is extremely difficult to infer the cause unless the search is constrained severely. Other knowledge sources, e.g. situational context, will need to be drawn upon to supply such constraints.

The review of face systems undertaken in this chapter indicates that the field is very active Face recognition has an appeal to researchers interested in its practical application to criminal investigation and non-verbal behaviour is entering the domain of human-computer interaction. The review also reveals that no extant system apparently has the same aim as that proposed for this thesis.

In the next chapter, the Knowledge Acquisition phase of development of the prototype, Janus, is described.

# Chapter 4:

# Knowledge acquisition

## 4.0 Knowledge acquisition for Janus:

Janus is a memory-based expert system which builds up and structures its specialist knowledge of face actions and interpretations incrementally through use. The memory is dynamically self-organizing. Its manifest knowledge is small at "birth" but it has the capacity to develop its knowledge to the capacity of its working memory. What distinguishes it from many other expert systems is that its inference procedure was not derived from the relations between the concepts in the domain but from a general theory of reminding and analogic thinking. In particular, the way it interprets an expression is by analogy to past experience of the same expression. Since this is applied generally across the domain and past experiences are always kept, the conclusion is that if an expression is not in the store, the user gets the next best in terms of the interpretation.

In the context of traditional expert systems, knowledge acquisition usually refers to the acquisition from the expert of knowledge upon which the knowledge engineer bases the system design and implementation. In this chapter the term also covers the acquisition of knowledge by which the prototype is validated, since the expertise is essentially of the same type.

### 4.1 What signals emotions that we should know them?

Many researchers from Darwin to the present day have described face components of discrete emotions. Many studies used face photographs which they presented to experimental subjects with a request to name the emotion depicted. When a consensus judgement results for particular judgements often the photographs and a syntactic description of the face actions attributed to the particular expression are published. Yet the reader can remain unclear as to whether such descriptions originate from the subjects or the researcher. If the latter then such descriptions form the latter's personal explicit theory as to the face components which were indicative. This, based on a very deep empirical knowledge of the domain is expert judgement without doubt - but of the researcher not the consensus of subjects. The reader is entitled to believe that the researcher holds the same judgment as the consensus as to the emotion depicted. The case is subtly different from the expert system paradigm in which the Knowledge Engineer focuses on the expert's stated reasons for the judgments or upon the experts thinking aloud as they tackle problems - rationalization though these may be - without the intermediary of another expert. There is no a priori reason why face actions alone or at all should be the most telling vehicles of emotion. Other phenomena such as wetness, dryness, colour, heat, electrical conductivity of the skin, blood volume of the skin, "hard", "merciless" or "tortured" eyes which defy anatomical description, as well as pupil size, flaccidity, relaxed appearance, wrinkles etc. all appear to be recognised as signalling some component of emotion.

### 4.2 Design constraints on the primitives chosen:

The problem for a system like Janus which aims to derive an emotion from a list of x/y coordinates is that its choice of the face vehicles of emotion is constrained by what can be represented by x/y coordinates. This could constrain the experts unduly in their descriptions in a traditional knowledge acquisition. Against this was the weight of much of the reviewed literature which represent a face expression in terms of face actions. Syntactic face actions such as 'brows raised', when attributed to face photographs, mostly expressed actions in two dimensions and could therefore be represented in terms of x/y co-ordinates.

## 4.3 A degree of nature constraining the nurture:

The decision was taken to acquire and incorporate into Janus a "core" of knowledge about *what* face actions signal *which* emotion and to acquire this knowledge from a published well validated source.

This core Knowledge would have a classifying role in filtering each input face expression into different branches of the memory tree. It comprises six sets of face actions from which can be drawn the typical expressions of the six primary emotions: happiness, sadness, fear, anger, disgust and surprise. The descriptions of these can be taken from the research literature without the need for acquiring them from people. Indeed, if definitive descriptions are required, this is much more satisfactory, since the field has been well researched and published tutorials exist containing face photographs well validated for the emotion displayed. Researchers who have spent their professional lives studying such expressions have formed theories about what the signs in the faces are for each emotion. Without this validity one would have to photograph models and validate the expressions by demonstrating that they evoke a sufficiently high concordance of judgement for each emotion, and then posit one's own theory as to what it is in the expressions that signal the adjudged emotion.

The need in Janus, however, is not for definitive descriptions with cast iron validity but for sets of likely face actions. Specific emotions can be signalled by more expressions than one and in more intensities than one, with the result that it is over-restrictive to have just the one definitive expression for an emotion. .

The idea is that the chosen sets would have a likelihood of containing those features which could be abstracted from many examples of expressions of that emotion so that

Table 4.0: Face actions associated with basic emotions

| | |
|---|---|
| Happy | eyes l-lid-raised; mouth bared; mouth up; mouth open; mouth wide; cheeks raised. |
| Sad | brows contracted; brows in-raised; brows centre-raised; eyes l-lid-raised; eyes inlid-raised; eyes down; mouth down. |
| Angry | brows low; brows contracted; eyes narrowed; eyes inlid-raised; eyes l-lid-raised; eyes l-lid-tensed; eyes u-lid-lowered; eyes u-lid-tensed; nose flared; mouth compressed; mouth l-lip-tensed; mouth u-lip-tensed; mouth bared mouth open; mouth wide; mouth square; cheeks n-l-vert. |
| Afraid | brows raised; brows in-raised; brows contracted; eyes l-lid-tensed; eyes-wide; eyes u-lid-raised; eyes l-lid-raised; eyes inlid-raised; mouth wide; mouth open; mouth pulled; mouth u-lip-tensed; mouth l-lip-tensed; mouth bared. |
| Disgusted<br>- | brows low; eyes l-lid-raised; eyes narrowed; nose screwed; mouth l-lip-everted; mouth bared; mouth u-lip-everted; mouth u-lip-raised; mouth l-lip-lowered; nose flared; mouth l-lip raised; cheeks n-l-vert; mouth open; cheeks raised; mouth compressed; mouth l-lip-tensed; mouth u-lip-tensed. |
| Surprised | brows raised; eyes wide; eyes u-lid-raised; eyes l-lid-lowered; mouth open; mouth slightly-open; mouth wide; jaw drop. |
| l = lower, u = upper, n-l-vert = naso-labial grooves vertical. | |
| | |

expressions which have most face actions in common with a set would be 'classified' under it. This would have the effect of classifying each input expression under a basic emotion label regardless of the interpretation attributed to it by the user. This 'primary emotion-label classification' does not over-rule the latter; it is only to be used as a default. The input label will be preserved and accessible along with other labels entered by other users for identical expressions.

The face action contents of these sets may change with the system's experience. In a human analogy, this might be the experience, for example, of one working solely with the victims of lower facial paralysis. Such a person might discover in time that the upturn of the corners of the mouth was an unreliable sign of the emotion: happiness in those people. In like manner, if Janus found that a 'core' face action was persistently unsupported there would be a possibility that this face action would be dropped from the 'core' set of that emotion only to be re-instated when its input re-included that face action with a sufficient frequency.

## 4.4 Knowledge sources:

Janus is built around this small endowed knowledge and its memory classifies all input in relation to it. The source reference for this face expression knowledge has been the 1984 edition of "Unmasking the Face"(Ekman & Friesen). This source has influenced the descriptions used in Janus for the six emotions of the same name as those authors' primary emotions. However the actual descriptions in Janus are somewhat different from the source and by being so, lose all claim to the validity of those of the source. Their use in Janus is in no sense definitional nor definitive but it is hoped that they merit a measure of credibility. Their aptness must be established by validation of the system.

Because Janus is a dynamic system, these descriptions can gain or lose a component through the vagaries of input, and their description in Janus is also modified by the constraints of the design as explained above. They are expressed in a restricted natural English language subset. What is represented is a set of face actions for each emotion from which a typical description may be drawn. These adaptations are given in **Table 4.0.**

## 4.5 The knowledge:

The knowledge built into Janus is of two kinds: firstly, how to transform measurements taken from a digitised face photograph into syntactic face actions and secondly, how to process the face actions to produce an interpretation.

Suggestions for the answers came from the published results from two diverse domains of psychological research: from the empirical findings on face expressions of emotion on the one hand (Ekman & Friesen, 1984) and from a theory of reminding (Schank, 1982) on the other. Both sources of knowledge came from the published literature of the researchers concerned.

It was necessary to bring the pertinent research findings of what expressions go with what emotion to bear on what Euclidean distances best describe an expression to solve the first problem. Creating the hypotheses of how one face expression can remind one of another and suggest its interpretation helped to solve the second. This suggested itself to the writer on reading Schank(1982) who suggested a theory of how human memory might be organized to explain the phenomenon of how one social event we encounter brings into mind some previously experienced event which has some analogy with it. The two events must be organized by a common explanation of how they both differ from the expected.

The first requirement was an inter-lingua in which these different sources of knowledge could talk with one another. The intermediate level representation of this knowledge was suggested by the research on face expressions of emotion. It would take the form of restricted English descriptions e.g. "brows raised". These can be expressed in terms of distances between standardized points located on a 2D image of a face.

The knowledge of how many points and where these needed to be located on a face in order to represent all the face actions involved in the selected emotions was derived by a hypothesis and test( trial and error ) technique.

## 4.6 Rule-based production of face actions from geometry:

An IF...Then rule was hypothesized for each face action. The antecedent conditions involved the horizontal and vertical co-ordinates of points located on a digitized image of a face and normalized by dividing each by the between-eye distance or nose-length respectively. This disposes of the first knowledge problem. Once the face actions representing the six sets of likely expressions were known, they would need to be defined in terms of horizontal or vertical co-ordinates or of distances on the face by a rule for each. The definitions were suggested by common knowledge of the face and the tacit understanding with which the action terms are invested. Thus, "brows raised" is tacitly understood to mean that the eyebrow has assumed a higher y-position than that prior to the action, and could be defined as an increase of a point on it in the y-direction over a certain minimum or in relation to a neutral expression. The definition then becomes the condition for a rule. The presence of a particular face action in an input face representation would be concluded by backward chaining: the antecedent conditions being tested in turn. The input geometric representation of the expression was converted into a list of the face actions it represented by applying the whole rule set to the geometric input in this manner. The process will be discussed in more detail in the chapter describing the rule base.

### 4.7 Acquiring domain knowledge from lay-experts:

So far the system has been conceptually designed with the only knowledge coming from books, and the tacit knowledge that one has about the face. The system will have no "deep" knowledge about how or why brows "rise", and such knowledge it has about these is composed of heuristics. This seems in order. Few humans, if any, could tell us why the brows rather than the nose rises with surprise nor how it is done. Knowledge about how faces were interpreted in terms of the emotion signalled and the reasons given for these was collected from various volunteers for the following purposes:

1. Acquiring knowledge of the domain,

2. Training,

3. Validation.

These are discussed in turn below:

1. How the lay-experts tackle the task of interpretation of face photographs may suggest a consensus among people of what they take into account.

2. Training is required in order to give the memory its experience. Training consists of entering a sequence of face expressions with emotion labels. The memory organizes these incrementally forming sub-generalisations as the sequence proceeds. The resulting organization reflects the particular knowledge (face actions) which the sequence held. The order in which the individual face expressions are input is only important in determining the order in which the organization is set up.

A trained Memory will consist of a tree with six main branches at the root of each of which is a node holding the set of likely face actions for a primary emotion. Indexed below each of these "likely" nodes are all the input face events (expressions plus their emotion label), some directly off the "likely node" (these have their constituent face actions completely subsumed by the "likely" face actions), or indexed by the face actions which differ from the "likely" face actions, creating a distributed memory of this event. Examples of the tree together with diagrams are shown in Chapter 7.

Sub-classifications are formed among these by the juxtaposition of two identical anomalous face actions in the same node. The roots of all branches contain the full event of all inputs which followed that branch, from which the emotions can be suggested for further input of expressions in search of interpretation by analogy to those already there. Training is incomplete, because it can always be added to.

3. Validation tests that the system produces the "right" answers. Janus is compared with human judges with respect to:

a) the face actions present in an expression. This requires the collection of judgments from human subjects of what face actions are to be seen in a series of face photographs, and,

b) the emotion signalled by an expression. This involves collecting the interpretations made by human subjects while studying face photographs.

The human judges involved in (b) above would be compared in their judgemental capacity with some well validated 'gold standard' judgments of expressions different from those employed above.

Acquiring the knowledge for training the memory and validation involved several acquisition techniques including interviews, questionnaire completion and scoring procedures. The procedures will now be discussed in more detail.

### 4.8 Eliciting emotion constructs from triads:

(see Kelly, G. A., 1955 for an exposition of Construct Theory)

**The Subjects:**

**Number:** 30

**Status:** College lecturers, researchers, students, administration staff

**Age:** range 20 - 65 years; average 33.5 years; mode 30 years

**Gender:** 12 male, 18 female.

**Culture:** academics and paramedics, pastoral professionals and students of Commonwealth origins.

**Stimuli:** 14 black & white face photographs, size 10x8 inches of male actor filmed during evoked response to various events. The face occupied about 70% of the area.

**Method:** Each subject was interviewed separately by the writer. Ten triads (3 photographs at one time in set combination and order) were examined by each subject, one triad at a time. In each case the task for the subject was to say in what way two of these three were alike in showing the same emotion and differed from the third who was showing a different emotion. The emotions were recorded. Later, subjects were asked to give cues for their attributions. The cues were recorded.

**Use:** In the selection of face points, obtaining rules converting measures to face actions, & training the prototype.

**Table 4.1: descriptive cues given for interpretation:**

(Subjects were asked to state anything in the photo that gave them the interpretation.)

**a) Cues with a possibility of representation in 2D co-ordinate values(Identical cues are not repeated):**

| | |
|---|---|
| Brows: | contracted; raised; pointing down; flat; horizontal; together & up; puckered between; narrowing of gap between. |
| Eyes: | narrowed; half-closed; slitty; eye pulled up; wide-eyed; down-cast; closed; hooded. |
| Nose: | screwed-up; horizontally-wrinkled at bridge; pulled-up; flared nostrils. |
| Mouth: | gritting of teeth; upturned; open; closed; longer horizontally; showing teeth; lower lip protruding; asymmetric height of mouth lower lip thicker; teeth together; top lip up; pulled back at sides; corners drooping; corners down; line between mouth and nose is at 45 degrees; slightly open. |

**b) Cues unlikely of representation in 2D co-ordinate values:** downward arc of naso-labial grooves, bunching of cheeks, chin merging into neck, softened jaw line, direct eye contact, taut muscles, lips stretched, stare, sneering, grimace, fixed eyes, deeply marked naso-labial grooves, smile, downward tilt of head, frown, deep naso-labial grooves, wincing, screwing of eyes, wrinkles around eyes, yawning, slightly wrinkled brow, deep marks under eyes, animation, double chin, tense chin, pointed chin, bunched muscles of nose, focused eyes, furrow arched down in centre, muscles not relaxed in eye, nose and mouth, needs a shave, fixed expression, looking straight at you, tension lines, rigidity of neck, less pouch, chin forward, chin back, long neck, more straight jowls, cheek furrows chubbier, more strained, contraction of middle face, tension of right side, tears, surly eyes, dull eyes, nose - mouth grooves drawn in by mouth, highlight in eyelid, humour lines side of eyes, dimples, bunching of cheeks, mouth relaxed, turned-up nose, yawn, twinkle in eyes, concentrating.

**Evaluation:**

**a) Interpretations:** The unrestricted interpretations evoked were eminently suitable for training the memory with a variety of emotion labels. These with their accompanying face actions would be taught to Janus and form part of the experience brought to bear on new similar expressions input in search of an interpretation. Such system-produced interpretations in later evaluation studies would be compared with the interpretations made by human judges.

**b) Cues:** Cues were not requested for all the interpretations. Their large number would make the task onerous for each subject. Cues were requested for a selection of them which seemed to the writer to be associated in meaning with the six primary emotions. The request was quite open: "Cues are anything in the photo that gave you the interpretation". **Table 4.1** gives a selection from the amassed responses, from which can be gained an appreciation of the type of cue which the writer felt could be described in terms of 2D co-ordinates. There is clearly a loss of information in representing an expression solely by such measures but knowledge of the underlying muscle actions which produce the appearance helps to re-phrase some of these cues in simpler terms e.g. the "screwed" nose referring to the horizontal wrinkles at the root of the nose can be approximated by the brows being lowered and the cheeks being raised. There seemed to be sufficient generality in the face actions expressed by this group of subjects to confirm their use as an intermediate representation in Janus.

The cues were interpreted by the writer with respect to the appropriate photograph(s) and the recurring cues which could be expressed in 2D coordinates were given optimal image points by which to represent the movement. Comparison of these points with points identically placed on the face features of a neutral expression of the model enabled the writer to express a criterion for the movement in terms of their difference when measured on a digitized screened image of the model. These points and distances were measured in picture elements or pixels. In this way a rule base was constructed which appeared to be capable of expressing a geometric representation of an expression for the six primary emotions in terms of syntactic face actions. This rule base is constructed only for converting a VDU-grabbed image into face actions. It is used in this thesis for validation purposes. With user descriptions, expressions are input in syntactic format, by-passing the rule base. Any face action of the brows, eyes, nose, mouth, cheeks and jaw can be input.

**4.9 Knowledge acquisition by questionnaire:**

The 'evoked' nature of the emotional expressions of these 14 poses(i.e. the model is all unsuspecting not knowing with what stimulus he is to be confronted) produced varied tilt and slant on the model's head with the result that distances were distorted on the face, and indeed some could not be made because points were obscured. This did not affect the logic of the rules defining each face action in terms of measures between points. These were still intuitively valid, but the actual distances which occur in their preconditions could not be measured with any accuracy on the 'evoked' poses. After trying mathematical corrections without being convinced of the accuracy of the resulting rules, a new set of posed expressions of a new model was photographed with the head constrained to a full-face presentation by a device. The new set of posed photographs were used to refine the rules. The rules were believed to have sufficient generality to measure all faces provided a normalisation was carried out on the image point values to correct for difference in size. The validation of the rule base utilized these new 'garryphotos'. The validation of this rule base will be discussed in Chapter 6. Here it is viewed as a form of knowledge acquisition since the validation proceeded by collecting human subjects' accounts of what face actions were present in the new series of photographs(called garryphotos) considered one at a time, and comparing the results with the face actions Janus derived from their digitised images by measuring the set of definitively-located points in terms of their x/y value using a tip-of-nose origin and passing this representation through the rule base.

The Subjects:

**Number:** 4

**Status:** 3rd Year medical students

**Age:** 22 - 24years

**Gender:** 2 male, 2 female

**Culture:** suburban English

**Stimuli:** full-face photographs

17 black & white face photographs, size A6 of male model filmed during posing happy, sad, angry, afraid, disgusted and surprised and one neutral expressions.

**Method:** The correspondence of the depicted expression to the poser's emotion at the time was not an issue. Each garryphoto was the subject of the same questionnaire which was completed while the photograph was examined by each subject in the absence of the writer. Thus 17 questionnaires numbered with reference to the appropriate numbered garryphoto were obtained from each subject, one at a time.

In each case the subject was asked to state which of the actions specified are present compared to photo

2 (the neutral expression) by putting a tick against those judged present.

**The Questionnaire:** There were 5 eyebrow actions (pulled together, lowered, raised, raised and pulled together, strait across); 7 eye actions (closed, upper lids down, upper lids raised, inner upper lid raised, lower lid raised, widely-open, narrowly-open); 2 nose actions (widened nostrils, screwed up; 13 mouth actions (teeth show, corners down, corners up, lips and mouth wide-open, lips slightly-apart, lips shut, lips pulled back, lips compressed, lower lip raised, lower lip everted, upper lip raised, upper lip tensed, mouth wide-open with thinly-stretched upper lip as in shouting); 3 cheek actions (raised, dropped , nose-mouth grooves curving down & accentuated due to nostrils pulled up or mouth corners down or both); 1 jaw action (dropped). Many actions were illustrated by explicit diagrams and positive or negative photographic examples of the action by another model. In each feature-section there was an "other" option to add new observed actions in addition to those described. The final questionnaire is included in the appendix.

**Use:** Refinement and subsequent validation of rules converting measures to face actions.

**Verification:** Preliminary questionnaires of a similar type had been designed and filled in by College personnel and the feedback from these had indicated the need both to alter some rules and the wording and format of the questionnaire. It was clear that many of the judgments required were very difficult to make and that diagrams would be helpful to make more explicit the measurement upon which the face action was based (Janus of course would go only by the rule). While the rules had been based on the writer's judgment of the tolerance allowed in the comparison between a neutral face and an expressive face, it was necessary to establish that this tolerance was a sensible one. This could only be done by testing the rules against successive batches of human judges, deriving what pointers one could as to what might improve performance, implementing the changes whether in rule or questionnaire, until there came a time when a fresh set of judges produced the desired result. The desired result is demonstrable consistency between the judgements of human experts and those of Janus. This will be discussed in Chapter 8.

**4.10 Knowledge acquisition of emotion; judging photographs by forced choice:**

The Subjects:

**Number:** 4.

**Status:** College personnel: computer science graduates.

**Age & (gender):** 21(m), 24(f), 38(m), 38(f) years.

**Culture:** fully exposed to mass media.

**The Stimuli:**

17 black & white full-face photographs, size A6 of male model filmed during posing happy, sad, angry, afraid, disgusted and surprised expressions.

**Method:** The subjects were asked to examine each photograph and complete the statement: "I think he is feeling: -", by selecting from the list: Sad, Happy, Angry, Afraid, Surprised, Disgusted. First and second choice were requested.

**Use:** For validating the interpretations given by Janus for the same photographs. These were derived from measurements on the digitised photographs and compared (using various statistical tests described in Chapter 8) with the emotion labels assigned by the human judges.

### 4.11 Evaluation of the nature of the knowledge acquired:

There were a few explicit indications in the acquiring of interpretations to indicate a judge's strategy of fitting an expression to a context recalled from memory e.g. "He doesn't want to know!", and "Not wanting anyone to be here - the sort of person who if you met him in a pub you wouldn't ask him to have a drink". More directive interviewing would possibly bring out more of these 'analogic' responses, but the desire was to be as non-directive as possible in order to observe the natural way of tackling the task. Most of the subjects tended to approach the task in a feature-analytic way. The following protocols may give a flavour of this approach:

**subject s.g.**, interpretation: sad: "It's more the lower half of the face .... The set of the mouth .... The drooping corners .... The eyes are downcast .... The relaxed forehead .... The closed mouth ....Relaxed ... Corner not upturned ...Mouth not widened! ... Nose lines obvious without being deep - lengthened(/ \) rather than pulled-back(< >) ... In a happy face the distance from nostril to corner of the mouth is shorter than in sad ... No wrinkling(humour lines) at corners of the eyes".

**subject j.w.**, interpretation: distaste: "Frown ... Tension lines... Contraction between the eyebrows ... The grooves from the nose to the mouth are very marked ... The mouth is down-turned... The features are down-turned and the jowls are straighter (draws lines to depict the downward slope of all the contours in the face converging to and from the bridge of the nose contrasting with the mixture of up and down in the contrasted faces). ... The chin is held back (generalized this - its forward in positive and backwards in negative encounters making the neck short whereas in thrusting the jaw forward the neck is long and this is non-aggressive and gives lightness)".

**subject n.d.**, interpretation: happy: "The face looks relaxed ... Mmm... why? ... The eyebrows are raised - at least not drawn down! ... Relaxed mouth - no pout, et cetera - seems like a lack of negative cues ". (comment: this seems more a holistic impression initially with rationalization).

**subject j.b.**, interpretation: surprised and questioning: "The eyebrows are lifted causing horizontal furrows on the brow ... The eyes are wide ... Widening of the nostrils ... Lifting of the cheeks ... Slight opening of the mouth ... Upward tilting of the head in the opposite direction ... Turning ... the jaw line is softened".

**subject m.g.**, interpretation: angry : "The eyebrows are drawn together - frowning ... The eyes are half-closed - slitty ... The mouth is pulled back and slightly down ... the grooves(between nose and mouth) are (/ \) rather than (< >) and deep and pronounced ... The tilt of the head - slightly forward(jutting) and to one side".

**subject a.t.**, interpretation: fear:

"Narrowed eyes ... open mouth ... dilated nostrils ... lowered eyebrows - asymmetrical".

The cue knowledge acquired contained many more descriptions than could be handled in terms of 2D geometry, which suggests that there is a loss of information in restricting cues according to this constraint. Examples of this lost information are concepts of softness, rigidity, dullness, chubbiness straightness and tension etc., applied to face areas; another group involves the focus of the eyes, and another the tilt and slant of the head. The terms that people use could be very varied. The main purpose of acquiring peoples' cues is to have sufficient to train the memory sufficiently and to validate this small experimental system with everyday knowledge. Provided that there are sufficient usable cues for this purpose, their adequacy must be assessed in relation to those results.

**4.12 Other knowledge acquisition of the emotion depicted in photographs:**

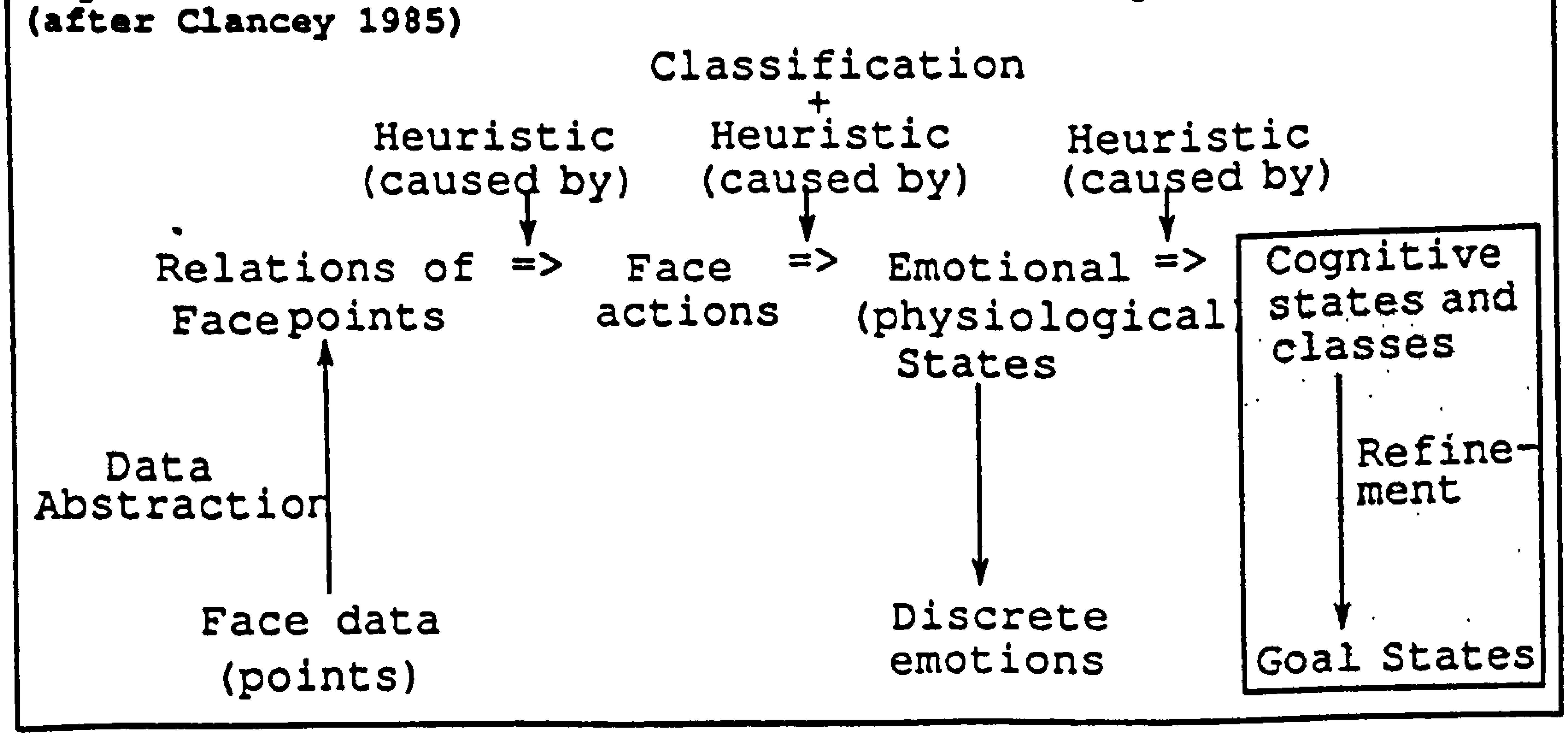
Meta-judges were used in validating Janus's interpretations. The task of each of these judges was to rate the aptness of the interpretations of Janus and the other judges in a "blind" study (without knowing that one set of interpretations came from a computer). To be in a position to do this they had to examine the appropriate photographs together with the interpretations and assign ratings according to a prescribed scheme. The credibility of these meta-judges was established by asking them to interpret a series of face photographs for which validated interpretations were available. This will be discussed in **Chapter 8**.

**4.13 Inference structure of knowledge in Janus:**

The knowledge involved in Janus can be analysed along the lines of the analysis of heuristic classification by Clancey(1985). Janus reasons about emotional states in terms of face expressions. The reasoning is heuristic in that it is based on knowledge of physiological structure and behaviour which is largely empirical. The inference structure (after Clancey) is shown below in Figure 4.0. The toned area represents an extension which has utilized Roseman's(1982) theory of the cognitive dimensions of emotion to show how the analysis of facial expressions could be extended to form theories about the goals motivating their appearance. In Chapter 10, a separate program, dyad, is described which uses Janus-supplied basic emotions in this manner.

**Summary:** In this chapter the role and acquisition of domain knowledge for Janus has been described. The source of this knowledge has been both textbook and human. Both specialized knowledge and common, heuristic associations were required. In a context-free setting, with only a face photograph to go on, interpretations in terms of emotion are not always easily made but are certainly obtainable. In these context-free conditions, most reasons for individual interpretations are sought in movements or changes in state of face features (brought about as we know by action of the voluntary and automatic nervous systems) and these form the main substance of the knowledge acquired. A loss of information is apparent when the representation of cues is constrained to what may be represented by 2D co-ordinates. This is necessary however in constructing rules to convert a digitised image of a face into a set of face actions.

**Figure 4.0: Inference structure of knowledge in Janus**
**(after Clancey 1985)**



In the next chapter the representation of face expressions will be described.

# Chapter 5:

# Representing face expression

## 5.0 The choice of face actions to represent expression:

In the last chapter the causes which lay-experts gave for their judgments of the emotion signalled by a photograph of a face were noted to be in terms of the positions and appearance of the everyday features: brows, eyes, nose, mouth, cheeks, head, neck and jaw as well as the grooves and wrinkles and relaxation and tension of these. In Chapter 2 where the findings of research were reviewed, it was seen that expressions have been mostly described in terms of facial action, although the more comprehensive anatomical representation was also favoured. In a situation where subjects were free to state their reasons, the finding that face actions figure prominently in so many protocols and so consistently for the same expressions made them the representation of choice for face expressions in Janus. This step is taken knowledgeably: they are rationalizations but such descriptions seem natural to people describing faces. The task described in this chapter is that of deriving a geometric representation of face expressions in terms of measurable distances on a digitised face photograph, which can be readily converted to the verbal form that people find natural.

## 5.1 Selecting points on the face:

The choice of face point location is intuitive and influenced by their potential to map to the 'given' face actions required by the Dynamic Memory. With the exception of the jaw and hairline and temples, they pinpoint inner face features. These are the features described (e.g. **Ekman & Friesen 1975/1984**) as being involved in expression. Jensen, (1986) presents evidence that the brows, eyes, nose and lips are perceived as a group. The hair and the jaw form a separate group that is classified independently. The jaw could, however, be grouped with the inner features.

The locations of the face points used in Janus were dictated by the face actions which needed to be represented in the classifying sets discussed in the last chapter. Distances measured on the face would form the criteria for these. Common knowledge of the moving and relatively unmoving features in relation to each other provided hypotheses for which distances to measure and these in their turn suggested where the face points should be placed.

The refinement of the number and positions of the face points was done by a hypothesize and test procedure in a verification of the rule base described in the preceding chapter. The early knowledge acquisition had provided a set of face actions upon which the locations of the points were based and the relations from which the early rules were drawn. Subsequent analysis of the text-based descriptions of the face actions of the primary emotions led to the establishment of "likely" sets of face actions as main organizers of knowledge in Memory. This required that these "likely" sets be represented in the rule base so that Janus would be able to infer their presence in an expression. Some additions were thereby needed to the points already incorporated in order to represent these. The face actions which needed to be represented are listed in **Table 5.0.**, and the positions and enumeration of the face points adopted are depicted in **Figure 5.0** and described in **Table 5.1**.

### 5.2 Comparison of the points used with those of other studies:

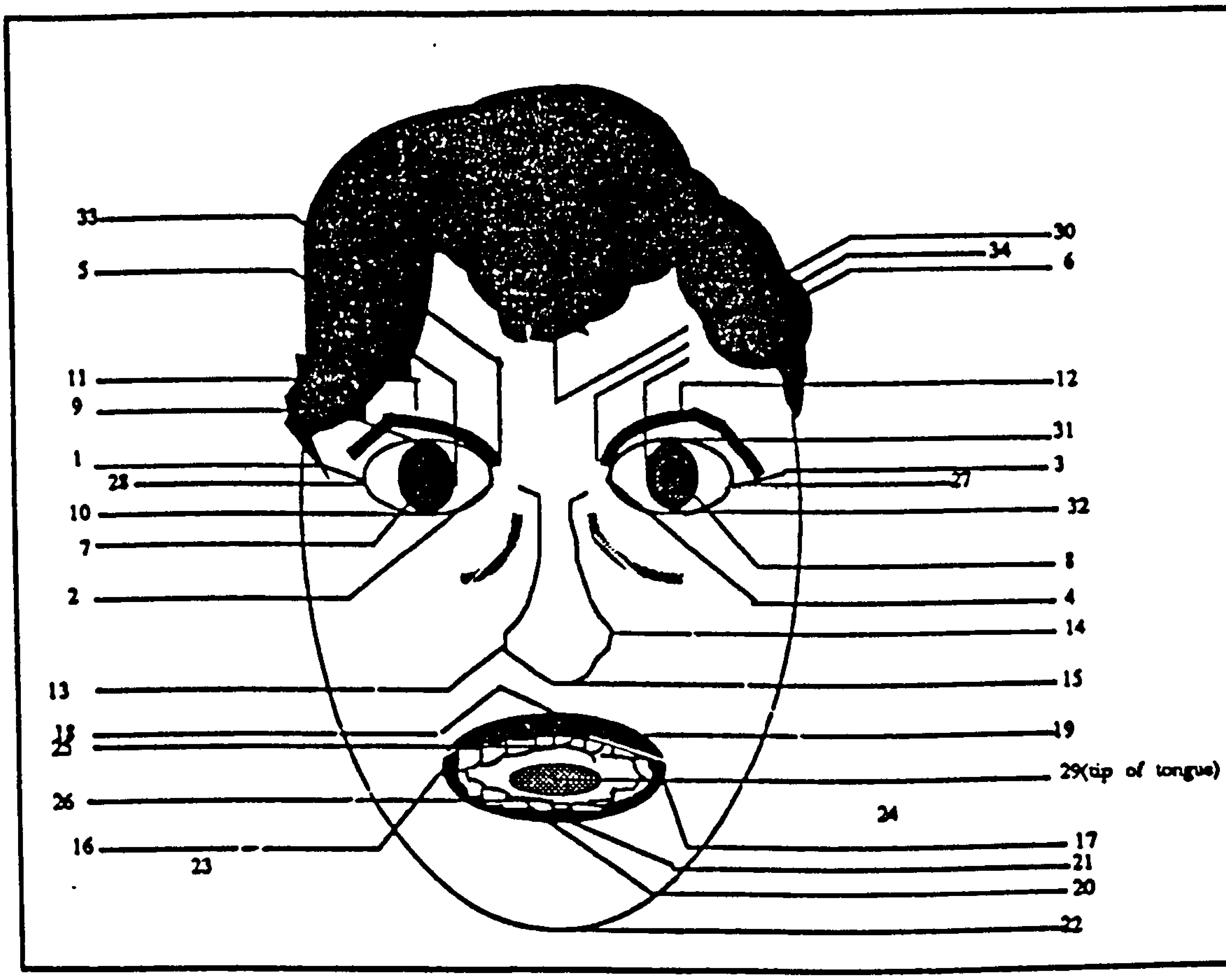
Janus is not alone in representing expressions as a series of normalised measures, nor in relating changes in distance between such to emotions. In Chapter 3 the work of the Pilowski, Thornton & Stokes (1986) was described. They present statistic correlations of such measures to the development of a smile over time and anticipate using pancultural expressions of emotions as

Table 5.0: Face actions used in Janus

| Brows | Eyes | Nose | Mouth | Cheeks | Jaw |
|---|---|---|---|---|---|
| low | l-lid-raised | screwed | open | raised | drop |
| contracted | inlid-raised | flared | down | n-l-vert | |
| in-raised | narrowed | | up | | |
| centre-raised | u-lid-tensed | | bared | | |
| raised | l-lid-tensed | | wide | | |
| | u-lid-raised | | pulled | | |
| | l-lid-lowered | | square | | |
| | u-lid-lowered | | sl. open | | |
| | down | | compressed | | |
| | wide | | l-lip-everted | | |
| | | | u-lip-everted | | |
| | | | l-lip-lowered | | |
| | | | l-lip-raised | | |
| | | | u_lip_raised | | |
| | | | l-lip-tensed | | |
| | | | u_lip_tensed | | |

n-l-vert = nose-mouth grooves more vertical and accentuated, sl. = slightly
l = lower, u = upper, in-raised = inner part raised, inlid = inner part of lid.

**Fig. 5.0: Face points: their positions and order in representing an expression.**



templates against which to match any facial configuration. Up to 100 face points are represented in their computer graphics model of a face. They do not explicitly specify the locations nor the rationale for selection of these 100 points but some comparison with those of Janus is possible from the account and diagrams of the distances which they compared for degree of smiling. These include points at the upper and lower eyelid/iris intersections which Janus does not have. This horizontal measure between the two points in the lower lid/iris interactions case correlates positively and significantly with the degree of smiling, and the vertical distance between this point on the lower lid and its counterpart on the upper lid is negatively and significantly correlated with the degree of smiling.

**Table 5.1: Locations of face points. Tip of nose is used as the origin.**

| No. | Location |
|---|---|
| 1 | external angle LH eye |
| 2 | internal angle LH eye |
| 3 | external angle RH eye |
| 4 | internal angle RH eye |
| 5 | iris-scleral junction midway between lids, LH eye |
| 6 | iris-scleral junction midway between lids, RH eye |
| 7 | LH pupil |
| 8 | RH pupil |
| 9 | upper LH eyelid border above pupil |
| 10 | lower LH eyelid border below pupil |
| 11 | top LH eyebrow above pupil |
| 12 | top RH eyebrow above pupil |
| 13 | LH side of nose at widest |
| 14 | RH side of nose at widest |
| 15 | tip of nose |
| 16 | LH angle of mouth |
| 17 | RH angle of mouth |
| 18 | centre upper lip upper margin |
| 19 | centre upper lip lower margin |
| 20 | centre lower lip lower margin |
| 21 | centre lower lip upper margin |
| 22 | point of chin |
| 23 | projection of LH eye under-crease meeting projected RH mouth line |
| 24 | projection of RH eye under-crease meeting projected RH mouth line |
| 25 | upper teeth lower margin central |
| 26 | lower teeth upper margin central |
| 27 | RH contour of face at level of external angle of RH eye |
| 28 | LH contour of face at level of external angle of LH eye |
| 29 | tip of the tongue or mid central lips or teeth line |
| 30 | centre hairline |
| 31 | upper RH eyelid border above pupil |
| 32 | lower RH eyelid border below pupil |
| 33 | LH eyebrow inner end |
| 34 | RH eyebrow inner end |

LH = Left Hand of viewer     RH = Right Hand of viewer

Presumably, both distances measure the rise of the lower eyelid in smiling, an action which Janus also measures in the happy face but using a different metric: the vertical distance between the lower lid below centre pupil and the internal angle of the eye. By comparing scattergrams of measures likely to be affected in a smile to those of a neutral expression Pilowski were able to decide which of twelve measures maximised the probability of a smile. (These are lower eyelid iris intersect, mouth width, end-lip-raise and mouth opening). The success of such a procedure, if carried out on the wide spectrum of face actions, would provide a good basis for choice of measures of facial actions (but only of those chosen) to incorporate into a rule for the presence of a smile. Such a detailed approach was not adopted in Janus. Janus would however be able to measure all but one(the eyelid/iris intersect) of these distances using its set of face points, were rules written to do this and added to the rule base.

It is clear that any set of distances measured on the face can not be complete. Even without a detailed statistical analysis, successful choice of points can be achieved using the results of validation studies with human judges and refinement can be carried out by trial and error.

**Shepherd (1986)** used physical measurements(37 xy co-ordinates, mainly based on the work of Jones et al.,1976), in "Frame": a computerized face retrieval and matching system prototype for police identification work.

"Frame", it must be understood, is for identification, not for interpreting emotional expression. It possesses a database of 1000 records of different faces coded on 50 attributes or parameters. A program compares a set of input face parameters with each of the records in the database, ranks records in their order of similarity to the input, and displays the corresponding facial images via a video disc player on a tv monitor.

Table 5.2: Evoking Stimuli for the T-series Photographs of male model numbered by photograph(see Appendix )

---

1. Continuing discomfort from an unexpected taste of salt.

2. A sudden noise: someone entering room and saying "Hello". Pleasant surprise, wanted to know who it was, curious.

3. A smell: button polish - reminded model of furniture polish - liked it; reminded him of some pleasant past house he had lived in.

4. Model was given offal to touch which was concealed in a bag. Model was a little alarmed and felt disgusted.

5,6. Acted frown: Model summoned frown by thinking of unpleasant memories, felt cross and perturbed.

7. The male model was shown a photograph of a female model in a flattering pose. The male model felt slightly embarrassed but mostly amused.

8. A smell: of bleach. Model took a good hard sniff and felt most uncomfortable - verbalized many disgusted feelings e.g. Revolting!, Horrible!.

9. A stretch following a yawn - model felt relief.

10. Looking at an unpleasant picture (war photograph): model felt distressed and bitter.

11. A pleasant picture of a female model, not such a powerful image as in (7) and not such a flattering pose. The model was interested and found it a pleasant image and again amused slightly.

12. Given salt to taste, model took a little too much and had to spit it out. He felt a slight nausea and discomfort.

13. A yawn - genuine early morning session- model was a little tired and felt a bit rushed, slightly grumpy.

14. Smell: aftershave - he thought it wasn't as pleasant as expected, wasn't sure whether he liked the smell or not, undecided.

In the representation used by Shepherd, thirteen more face points are allocated to outer features and hairline than is the case with Janus. This reflects the greater salience of the outer and upper features in classifying identity as opposed to expression. The linear and area measurements derived from the 37 xy coordinates comprised 21 of the 50 parameters coding descriptions of faces on file. Expression does not appear to figure; nor do the attributes express actions.

In JANUS, since it is by passage through the rule base that actions are produced from xy co-ordinates, errors can be due to point selection, point measurement or an inappropriate rule.

**5.3 Obtaining face photographs:**

a) The T-series:

An actor was photographed in evoked emotion situations. The photographer, a leader in the College photographic society, was asked to provide the model( she had an actor friend, M.T. ) and the evoking stimuli were left to her with the direction that the stimuli should be chosen as likely to evoke emotions of happiness, sadness, fear, anger, surprise and disgust, and these for the most part were 'sprung' on the actor who was then photographed at the moment of reaction.

The evoking stimuli are described in **Table 5.2**. The series was called the "T" series (see Appendix ) and were 14 in number, each 19.8 x 25cms, black & white. An estimated 1/2 of the photograph area was occupied by the head. The T-series photographs were used for the triad differentiation knowledge acquisition described in the last chapter.

b) The garryphotos:

These 18 black & white A6 size photographs of a male model, the writer, are full-face views of posed expressions of happy, sad, angry, afraid, surprised and disgusted emotions. They are full-face with the face area occupying about 1/4 of the print. Head movements were restrained by a device and the distance of the camera was kept constant. The shots were taken by a professional photographer. Knowledge of the facial expressions for these emotions as detailed in the pertinent literature was in mind while these were posed.

c) The Ekman(1975)series published by The Consulting Psychologists Press:

These are full-faced black & white studies of well validated facial expressions of the six primary emotions and of neutral expressions of the several models photographed.

**5.4 Digitising photographs:**

The photographs in the T-series and the garryphotos were digitised on a flat bed scanner using the Mackintosh C-scan bitmap format producing 16 dithered gray scales. Subsequently the image was converted to Sun rasterfile format and displayed on the black and white Sun 2 workstation on which Janus was implemented. The images gave passable definition of the face features.

**5.5 Measuring the location of face points:**

The T_series images were screened in a 606 x 779 window using the Poplog Window Manager(PWM) in Suntools. The PWM provides image processing facilities, one of which is mouse tracking. This was used to obtain the coordinates of the points required by moving the mouse cursor to the position and pressing the appropriate mouse button. A procedure was written to subtract the x/y values of the tip of the nose from these values so that all measurements were made relative to this origin.

## 5.6 Normalizing face points:

The measurements for each face were made in the order described in Fig 5.0 and were then normalized to take account of the differences in size between different faces. Horizontal distances were divided by the distance between the inner angles of the eyes of the same face; vertical distances were divided by the length of the nose.

## 5.7 The Face representation and Face Base:

The two coordinates for each point were enclosed in a list format, the complete face expression representation being a list of 34 sublists. The order of sublists was maintained for each face so that any specific point could thereafter be retrieved directly. The entire set of faces was stored as a list which formed a face expression database. The neutral expression was also included in the database.

Summary:

In this chapter the geometric representation of face expressions in Janus has been described. Each expression is represented by a list of lists. Each of the 34 enclosed lists holds the coordinate values of a facial landmark as measured from a digitized image. There are many problems in adopting a face point approach. Even with a high-resolution, grey-scale image, the position of some points owe something to guesswork due to shadows and blurring. One cannot be sure that one has chosen the optimal points to capture the action. The locations of the face points are determined by their usefulness in defining the face actions which, in various combinations, can describe the likely facial expressions of the primary emotions: happy, sad, angry, afraid, surprised and disgusted. In the next chapter will be described the construction of the rule base by which this geometric representation is converted into a syntactic form in terms of face actions.

# Chapter 6:

# The rule base

## 6.0 Introduction:

In the last chapter the geometric representation of face expressions was described. The rule base transforms and decomposes this geometric representation of a face expression into a set of 'syntactic' face actions. It is in this format that the expressions are stored in Memory. The transformation is effected by passing the list of the 34 normalized x/y positions as a parameter to each rule in turn and collecting those face actions that are returned as 'true'.

The reason for having a geometric input and rule base is to allow for automatic visual scanner input capability. Sakai,T., Nagao,M. and Kanade,T.(1972), Bromley(1977), Buhr(1986) and others have developed algorithms which attempt such automation. Sakai et al. determined the positions and dimensions of face features with a view to recognizing and classifying faces. The algorithm was designed to locate in sequence the top of the head, the sides of the face, the nose, mouth and chin by counting the non-white pixels in a thin horizontal or vertical strip as it is moved respectively down and across the image.

However it is not proven that the procedure could measure the 34 points needed by Janus with sufficient accuracy. The development and validation of an error-free algorithm capable of automatically measuring 34 face points with sufficient accuracy was considered to be beyond the scope of this project. The emphasis here is on demonstrating that points chosen can represent the face actions typical of the primary emotions.

Face features to be used in a pattern recognition system, with say, identity as the goal, may be selected on the basis of a number of criteria viz: their usefulness to discriminate among faces; the ease with which they can be measured, their stability over time or their contribution to the efficiency of the system (Laughery,K. et al.,1981). Face features to be used in expression analysis, on the other hand have to move from a neutral position, and the same movement may take part in more than one expression; they are not necessarily the same points that discriminate between people e.g. the hair line and head shape. In Chapter 5 the saliency of inner features for emotional expression was noted. These include the eyebrows, eyes, nose, lips and, since the level of the chin is related to the degree of opening of the mouth, the chin.

### 6.1 The rationale for the choice of the rules:

The Face Points were displayed in **Figure 5.0** in the last chapter. The rules use the co-ordinates of specific points to reason whether their under-lying feature positions have changed sufficiently from the rest position to conclude that the specific feature action has occurred.

The rules are required to cover those feature actions in untrained Memory which make up six special groupings or clusters. Each represents a set of face actions from which "likely" expressions can be drawn of one of the six universal emotions(Ekman, P. et al.,1972; Ekman & Friesen, 1971; Ekman et al., 1969, Izard, 1971) i.e. those emotions that all cultures will associate with some specific expression. These are the primary emotions of happiness, sadness, surprise, anger, fear, and disgust.

The choice of these clusters of feature actions was influenced by the work of Ekman & Friesen(1975) who have published a comprehensive account of such facial cues and associated emotions together with illustrative photographs. The Janus clusters are however different in many details from these owing to design constraints. They therefore make no claim to the pancultural validity of the Ekman & Friesen descriptions.

Ekman & Friesen (1978, 1976a) adopted an anatomically based approach in terms of 'Action Units' which have superseded face actions. However this latter approach is considered too technical for an expert system meant to be used by untrained lay users. Each input expression("an event") is matched against these six clusters. The degree of overlap determines in which of six divisions in Memory the new "event" is stored. If Janus is to maintain an input option in geometric format then clearly the rule base needs to define these "likely" feature actions so that the input may be correctly channelled. Table 4.0 tabulates the face action clusters associated with each basic emotion.

### 6.2 Design constraints on the rules:

The design constraints are imposed by the use of a 2-D geometric representation. Some qualitative cues such as 'bunching', 'deep' and 'rounded' for example could not be defined. Others, such as 'tensed lips', can be defined heuristically from their thickness or associated measurable face actions. Tension in the lower eyelid, however, cannot be defined directly and is inferred from a combination of lower-eyelid-raised, cheeks not raised and the mouth turned down.

### 6.3 Construction of the Rule Base:

The construction and testing of the rules requires accurate face-point location data. In chapter 4 (Knowledge Acquisition) the different series of photographs used were described. The first series commissioned were evoked responses of the male actor and were referred to as the "T" series.

The tilt and slant of some of the T-series poses raised problems with measurement. For some time the rules were defined incorporating procedures to measure tilt and slant and mathematical corrections were incorporated into the rules. Still some of the required points were just not visible and their positions would either have to be written out of the rules or estimated on a presumption of symmetry with the visible side of the faces. Clearly one can not write such rules without constraints of view and expect them to have generality over all views.

The decision was taken to limit the phrasing of the rules to apply to the full-face condition. This has the disadvantage of excluding head movements from the poses and of limiting a vision front end to full-face views only, but would not detract from the thesis aim which is to determine if it were possible to map from the geometric to the semantic representation of the expression.

A new series of full-face, posed emotion expressions were commissioned. These are the "garryphotos". The rules were phrased for the full-face condition. Tilt and slant corrections were omitted from the rules. These 'garryphotos' are described in the last chapter. The writer posed the six basic emotions trying to reproduce the six emotion-specific expressions described by Ekman & Friesen(1984) for happy, sad, afraid, disgusted, angry and surprised emotions and a neutral expression. The head was restrained from moving. The series comprised 18 full-face black and white positives.

The concern was to identify displacement of face points from the rest position, measure the displacement, and make a rule to relate the displacement to the feature action. There was the problem of basing such definitions on the evidence of just one subject but this was partially rectified by a normalization procedure performed on the measurements to correct for differences in size and scale and by including, in appropriate rules, a comparison with the neutral expression.

## 6.4 Hypothesize and test refinement:

Having digitized and measured the 34 face points on each face and normalized them by the procedure described in the last chapter, the required face action rules were defined provisionally, using common face knowledge of the movement and with reference to the neutral expression. Promising distances between points were tried.

Each rule took the name of the face action it defined. Each was tested against a set of coordinate positions representing each face and "wrong" results (the writer thought them wrong) indicated need to re-question either the logic, the face points involved or their measurement or all or any of these. Refinements in the tolerance allowed in measurement, or selecting another distance to represent the action, might make a more sensitive rule.

To make this procedure clearer, consider that a feature action such as "mouth up" can be conceived of as an upward turn to the corners of the mouth from rest. One knows the relative positions of face features well enough to reason that the distance between the corner and the outer eye angle will be shortened; but so probably will the distance between the corner and the inner eye angle, the mid-point of the hairline, the tip of the nose, the mid brow point et cetera.

If the rule incorporating one of these distances compared to its counterpart in the neutral face does not give the "right" results in every photograph, then one can either try another definition over the whole set of photographs or alternatively, re-check one's measurements on the digitized image.
In the refinement of the rule base, four human experts

were presented with face photographs and asked to identify face-actions. This information was elicited from them by using questionnaires(Chapter 4). The same photographs were digitised and measurements of facial landmarks were input to the rule base of Janus. The resulting face actions were compared with those obtained from the human experts. Discrepancies suggested modifications to the format of the questionnaire in the form of diagrams and examples of the face action being judged and by altering the phrasing and order of questions. Another group of four lay-experts were used to test these modifications and make final refinements. The final form of the questionnaire maintained the six sub-categories of 'brows', 'eyes', 'nose', 'mouth', 'cheeks' and 'jaw'. Many alternative choices under these were illustrated with diagrams indicating the precise measurement at which the choice was aimed, and different model depicted examples.

### 6.5 Description of the rules:

The rule base contains 39 "IF...Then" rules which deliver decisions of "true" or "false" as to whether a specific face action has occurred. Many require tolerances which can not be generalized over the whole rule set. When each face action rule in turn is applied as a query to the list of lists containing the 34 normalized x/y values representing a specific face, the particular mathematical relations required for the antecedent conditions are computed with reference to the numerical values for that expression. The result is a true/false decision on the presence of that feature action in the expression. Several such relations may be required involving, maybe, other face action rules by backward chaining in an effort to collect evidence. Many of these(22) are independent of other face actions(apart from comparison with a neutral face of the same person), while the rest(17) use the context of other actions on the same face. For instance the rule for 'brows-contracted' is context-free and uses only the distance between the inner ends of the eyebrows. Tension in the lower eyelid, however, cannot be defined directly and is inferred from a combination of lower-eyelid-raised, cheeks not raised and the mouth turned down. An example rule is given below.

### 6.6 An example rule in POP11 and the natural language equivalent:

```
define eyes_l_lid_raised(mug) -> i;
;;;The vertical distance between the inner angle of the
;;;LH eye and the point on the lower eyelid below centre
;;;of pupil is less than that of the neutral face.
;;;'mug' and 'norm2' represent the 34 x/y values of the
;;;target face and its neutral counterpart respectively.
vars l_eye_i_ny = norm2(2)(2), l_l_lid_ny = norm2(10)(2),
l_l_lid_y = mug(10)(2), l_eye_i_y = mug(2)(2);

(
(l_l_lid_y - l_eye_i_y) < (l_l_lid_ny - l_eye_i_ny)
) -> i;
enddefine;
```

comment:
"l_eye_i_y" is a variable holding the second sublist in the face description: that of the inner angle of the LH eye; "l_l_lid_y" is that of the 10th sublist: that of the LH lower lid at the pupil. "(2)" = 2nd element or y-value; the value of 'i' is 'true' or 'false'

**Natural Language equivalent:**

**IF**

the vertical distance between the centre of the lower lid margin and the level of the inner angle of the LH eye is less than the corresponding distance in the neutral face,

**Then**

the lower eyelid is 'raised'.

**note:** The "greater than" relation in the rule is due to the fact that the origin is at the tip of the nose and the y-values of the eyes are negative in sign.

Encoding such geometric functions into a rule requires a comparison with the same measures on a neutral face of the same person, This would drastically reduce the scope of practical applications for a system of this design even assuming that automatic measuring of face points were made accurately. The problem might be ameliorated by a probability approach based on a population distribution of the measures used. In a system which gains expertise from experience, the user population distribution of measures for the rules concerned could be automatically updated and averaged to serve as norms in rules for comparison purposes at different probabilities to determine the face actions present in the next encounter.

## 6.7 The rationale of specific rules:

The rationale of the rules is illustrated in the following examples and the rules are listed in full in appendix(). The variable norm2 denotes the measurement on the neutral face

**Contracted brows:**
The distance between innermost points on the eyebrows is less than that of norm2 The brow is contracted towards the mid-line above the nose causing vertical creases above bridge of nose.

**Low brows:**
The brow is lowered onto the eye. Y-distances between top of the LH eyebrow and inner angle of the eye are less that of norm2.

**In-raised brows:**

The vertical distance between the medial end of the LH eyebrow and the uppermost point of it above centre-pupil is less than the norm2.

**Raised brows:**
The LH eye-brow is raised compared to norm2.

**Centre-raised Brows:**
The medial end of the LH eyebrow is higher than that of norm2 and the brow is contracted.

**Raised upper eyelid:**
The vertical distance between the centre of the LH pupil and the upper eyelid is greater than norm2.

**Tensed upper eyelid:**
The outer upper lid is lowered and the upper lid is medially raised putting strain on the eyelid. The lower lid is tensed

**Raised inner upper eyelid:**
Centre-raised brows or in-raised brows and contracted brows have to be true.

**Raised lower eyelid:**
The vertical distance between the inner angle of the LH eye and the point on the lower eyelid below centre of pupil is less than that of norm2.

**Lowered upper eyelid:**
The LH upper lid margin mid point is lower than that of norm2

**Lowered lower eyelid:**
The vertical distance between the inner angle of the LH eye and the point on the lower eyelid below centre of pupil is greater than that of norm2.

**Tensed lower eyelid:**

The lower lid is raised and the cheeks are not raised.

**Eyes down:**

The upper lid margin mid-point is lower than the inner angle of the LH eye

**Eyes open:**

The distance between points on the lids vertical to mid pupil in comparison to norm2 define grades of opening from 0 (= shut) though a lesser distance(narrowed) to a greater distance(wide).

**Flared nose:**

The width of the nose measured along the x axis at its widest point is greater than norm2.

**Screw nose:**

The brows are lowered and the cheeks are raised.

**Mouth corners turned up:**

The y-level of the corners must be above that of mid upper - lip lower boundary and their vertical distance from the inner angles of the eyes must be less than that of norm2.

**Mouth open:**

Degrees are defined by measuring the vertical gap between upper and lower lip margins(y mid-line values). The cut-off values are arbitrary and distinguish between 'slight' and 'wide' and, with the addition of the inter-teeth gap, 'square' and 'open' and 'shut'.

**Compressed mouth:**

The mouth must be shut and the vertical measure of the upper lip red margin in the centre must be less than that of norm2. Alternatively, the mouth must be bared, the mouth open, and the upper and lower teeth together and the mouth corners must not be turned upwards.

**Mouth corners turned down:**
They must be lower than the centre of the lower margin of the upper lip and the y-distance between the inner angles of the eyes and the respective corner must be greater than norm2.

**Pulled mouth:**
The x-value of each corner must be displaced laterally compared to norm2 and the corners must not be "up".

**Raised upper lip:**
The centre point of the lower margin of the upper lip is displaced upwards compared to norm2.

**Tensed upper lip:**
The mouth is compressed and the corners not turned up.

**Tensed lower lip:**
The y-distance in the mid-line between the margins of the red part of the lip is greater than norm2. The corners of the mouth must not be turned upwards. Alternatively, the mouth is compressed.

**Raised lower lip:**
The centre y-value of the upper margin of the lower lip is displaced upwards compared to norm2.

**Everted lower lip:**
The centre y-value of the upper margin of the lower lip is displaced upwards and the lower lip red part is greater vertically compared to norm2.

**More vertical naso-labial grooves:**
The nose is flared and the point at which two projected lines from top cheek and mouth line meet is below a point close to the origin on the y-axis.

**Raised cheeks:**
The points at which two projected lines from top cheek and mouth line meet are above a point close to the origin on the y-axis and the lower eyelid is raised. The jaw is not dropped.

**Dropped cheeks:**
The points at which two projected lines from top cheek and mouth line meet are lower in y-value than norm2. The lower eyelid is not raised.

**Jaw dropped:**
The mouth is wide open.

**Summary:**

In this chapter the process of obtaining face actions from geometric distances has been described and illustrated. Each syntactic face action is made the conclusion of an "IF ... Then" rule which specifies the geometric relations which have to be satisfied for the conclusion to be true of the face under consideration. Each of 39 face actions in turn is tested against the face description and those verified are enclosed in a list format for input to the Dynamic memory. The rationale behind a sample of these rules is explained and a rule is presented both in POP 11 code and in natural language. In the next chapter, the Dynamic Memory is described.

# Chapter 7:

# The dynamic memory

## 7.0. Overall functions of memory:

The dynamic memory performs two functions. In interpret mode, it accepts a list of face actions and returns the appropriate emotion label. In learn mode, it accepts both a list of face actions and the associated emotion label and adds them to its repertoire for future use. How it performs these functions will be discussed in this section.

### 7.1. Psychological foundations:

The organisation of the dynamic memory is based on Schank's theory of reminding and learning(Schank 1982). Schank's model of the human memory explains how autobiographical or social events are stored, organised and remembered. He introduced the concept of Memory Organizational Packets(MOPs). These abstract conceptual structures organise other generalized structures called "scenes" related by a common goal.

The "scene" describes a generalized situation which is instantiated in perhaps many diverse standard situations which are about people, things and objects. Personal episodes often involve the enactment of these standard situations with some idiosyncratic variations within an autobiographical event. These functional storage structures act dynamically with one's experiences keeping track of similarities and differences from what has become standard. An anomalous event commands one's attention and is indexed from the expected by the anomaly. The typical ceases to remind us of other situations but were the same anomaly to recur, one can be reminded of the first event. At the level of the MOP this phenomenon could account for analogy reminding. The point is that for one happening to remind one of another, the two must be similarly indexed in the same larger memory structure so that an explanation is possible for the reminding. Within each Mop, experiences would organise memory to the extent that their temporal, physical, social and personal aspects become, with repeated encounters, so generalized and abstracted as to be expected over a wide range of events whereas deviations disturb expectations and are indexed at the point in the MOP at which they occur. Should the atypical become supported by further encounters in other experienced events, one tends to be reminded of the prior occurrence indicating that they are contained in the same memory structure and have taken on an organisational role as a sub-classification. This is sub_mop formation. Henceforth all experienced events of this type will be represented in this sub_mop. The former exception might so much become the rule that it then creates the expectations for the person. This ability to re-organize automatically with experience is what is meant by Dynamic memory. Schank's theorizing goes far deeper than this, involving themes and scripts goals and plans. He has more recently moved the emphasis of reminding to explanation. Janus applies an interpretation of the basic idea of the theory to a domain and at a level of event(that of face actions) for which it was not primarily conceived.

**7.2 Memory-based expert systems:**

The concept of a memory which generalizes from experience provides a different model for expert systems, called "Memory-based" . In "Memory based expert systems", Schank (1984) refers to the Alfred Economics Learning Project,(Riesbeck, 1984), and a case-based reasoning project, Judge, as systems using this concept. Lebowitz(1986) also embodies a dynamic memory in his "UNIMEM" system. The principle has also been used in a news item understanding system on terrorism, IPP(Lebowitz, 1980). Here the generalizations (standard situations) are made on the basis of similarities in news events. Following their instantiation by a news item, the causal relationships in these general structures act as top-down expectancies of what the news item will contain and interpret it in relation to them. As in Janus, some endowed knowledge about the domain is required initially before automatic generalizations take off.

Schank's ideas were applied to the domain of political events by Kolodner(1984), who developed a memory based retrieval system(Cyrus) based on reports of the diplomatic activities of two U.S. Secretaries of State. Some of her ideas and representations have been used in the design of Janus. Knowledge-based self organizing generalization lies at the heart of Cyrus where the features have to be extracted from the input before they are compared for similarity. If similar to a feature already in a MOP a new generalisation is made of that feature and added to the MOP. This is emulated in Janus as are also system designs to promote and demote generalizations.

**7.3 Face knowledge in memory:**

The dynamic memory is initially endowed with six MOPs(called face_mops) which represent the basic expressions typifying happiness, sadness, anger, disgust, fear and surprise. This is the only explicit face knowledge built into Janus. Any other knowledge is acquired from the users. These basic expressions were chosen as being typical in the sense that they are based on emotions which have pancultural validity (Ekman & al. 1969, Ekman & Friesen 1971, Izard 1971). The six clusters of face actions endowed in Janus were influenced by those specified by Ekman & Friesen(1984) with such modifications as were necessitated by design constraints. They each represent a pool of face actions from which typical expressions of the emotion can be drawn. Following Kolodner's nomenclature the part of the MOP which holds these is called the "Frame". The six Face_mops Frames are the level 2 nodes in an embracing tree(**Figure 7.0 & Figure 7.1**), where the root node is level 0. Related but atypical input events(in the sense of having some face actions which are typical and some atypical) have their atypical face actions stored as sub-trees below the Frame, each auto-indexed using its label as index in a two tier design as shown in **Figure 7.2.** This design gives an overall uniformity of depth of six nodes to the system. The links indicate how the events lower down differ from the typical. Each learned face expression is viewed as an autobiographical event and is represented as a list with an unique identifier(ID). The learned interpretation is also stored in the event node, e.g.

[brows raised eyes wide mouth wide jaw drop interp surprised].

**7.4 Representation of emotion:**

It is important to realise that there is no explicit representation for the interpretation which is normally input along with face actions when an expression is to be learned. Emotions are not indexed. Memory is face-action indexed since it stores face expressions, but in the leaves, the input events(but only their IDs) are stored which include the emotion interpretation. It is from

**Fig. 7.0: Basic tree of pre-defined action components:**

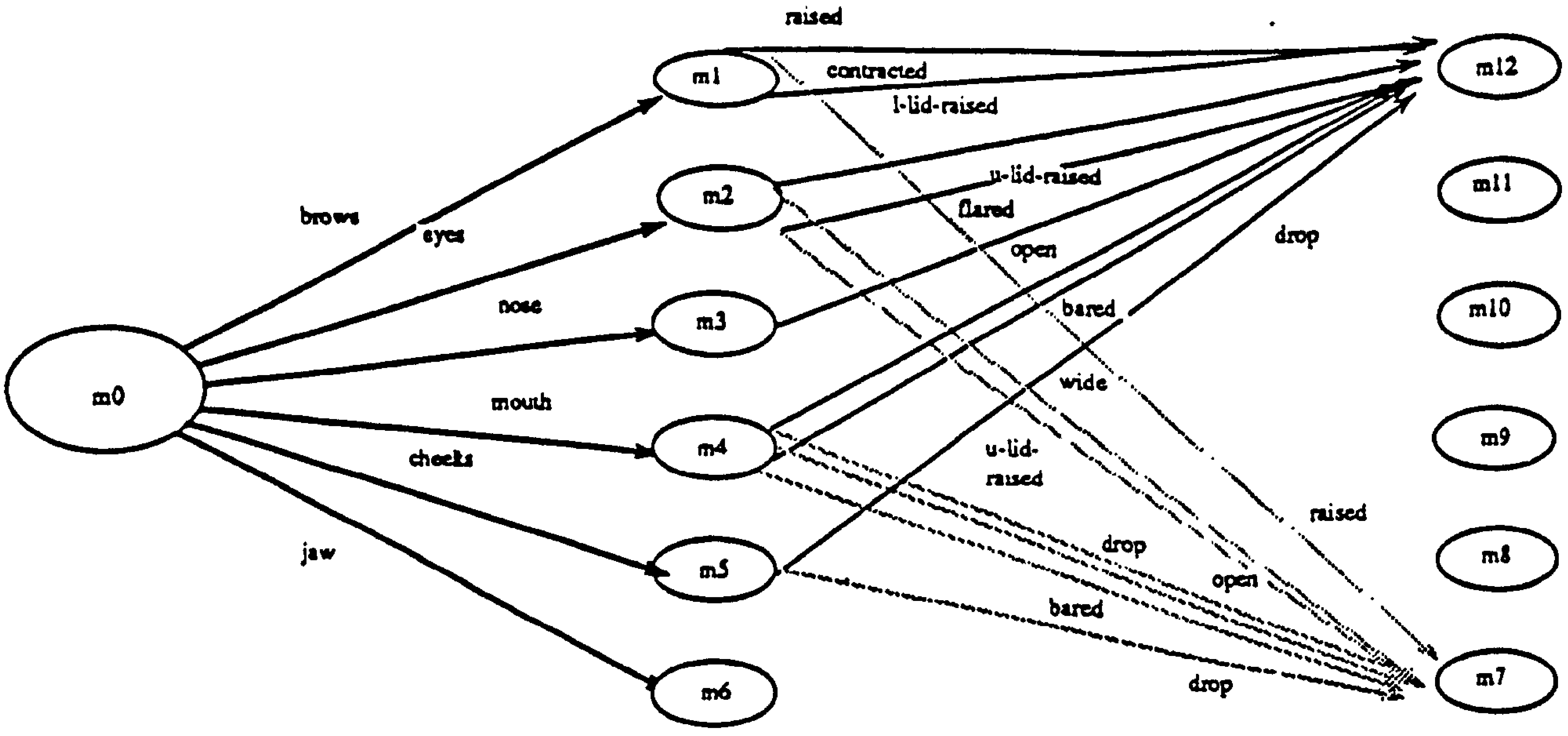


Figure 7.1: Schematic representation of Face Knowledge in Janus.
A branch of the conceptual tree is shown (centre) with the the records which represent it on the right. On the left is the overlay Flavour hierarchy classes, instances of which are made when matches take place between input and node content.

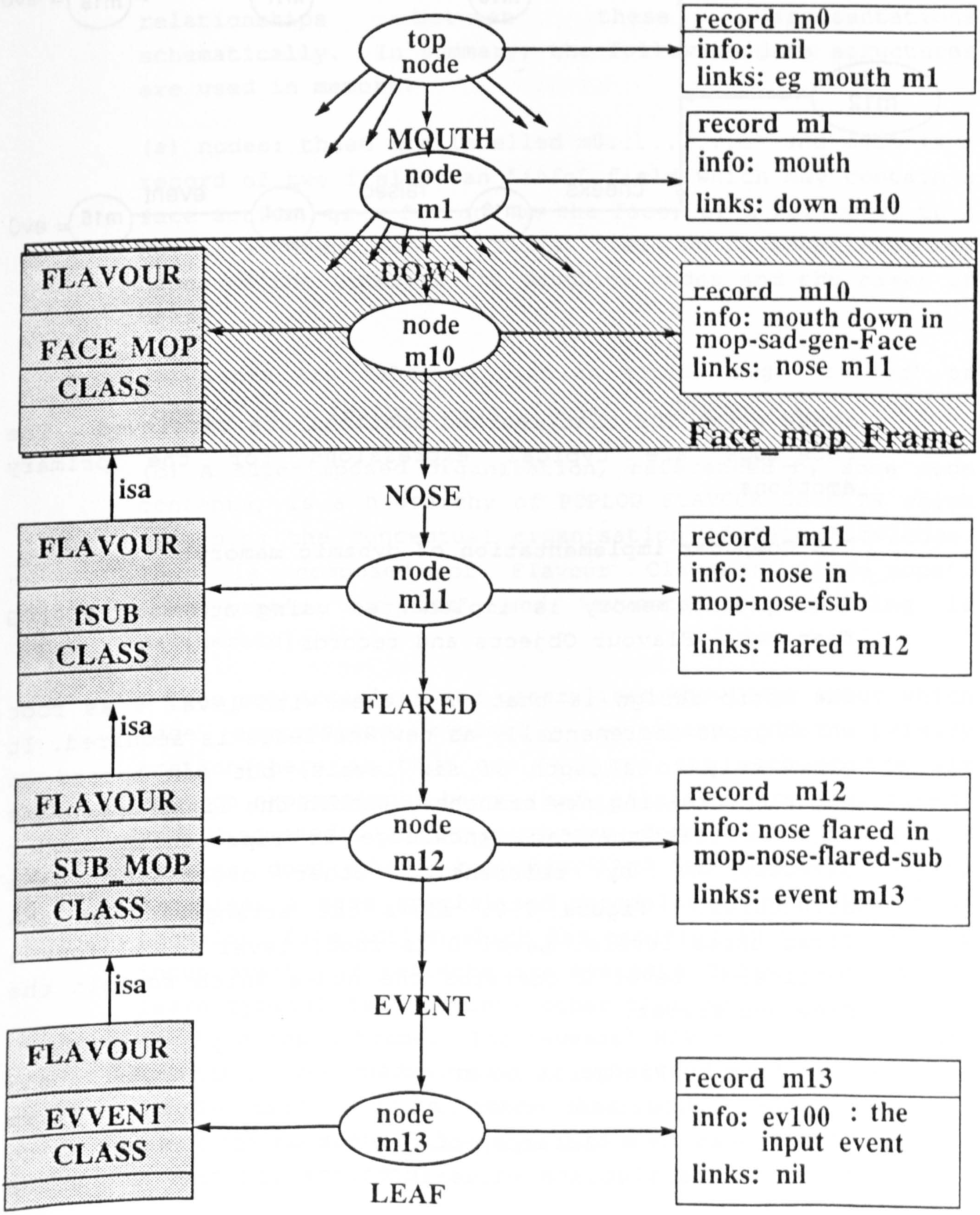
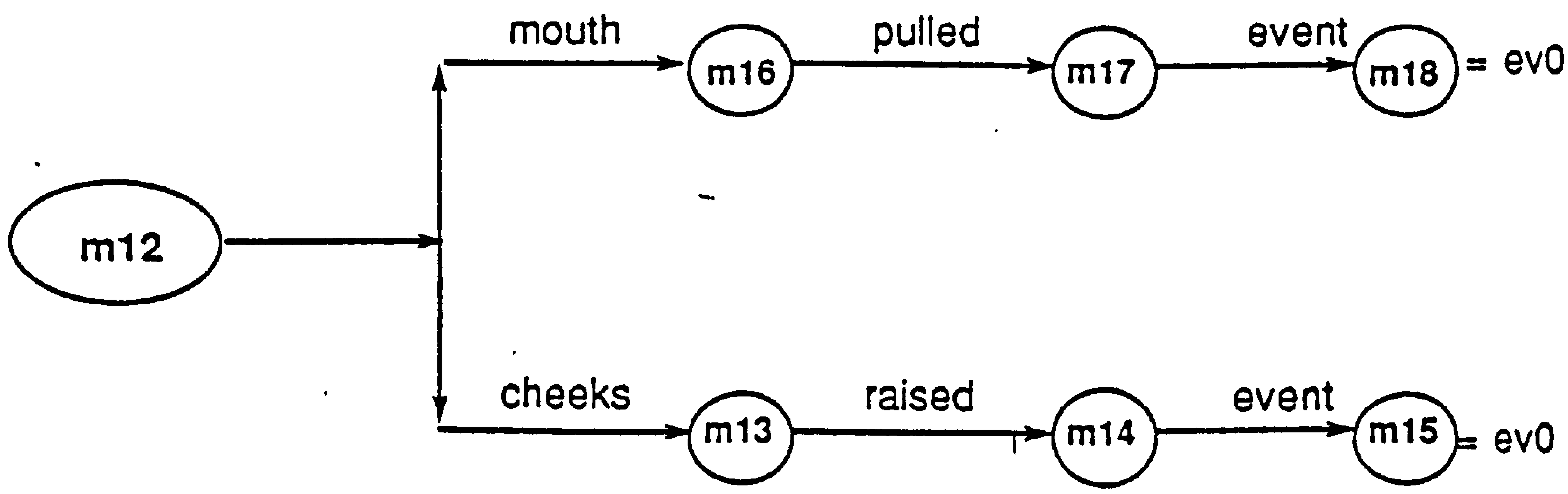


Fig. 7.2: A new event, ev0, differs in two actions from those in m12. Each indexes ev0 below m12.




these event IDs that learned emotions are retrieved. The Face_mops are typical expressions for the primary emotions.

### 7.5 Computer implementation of dynamic memory:

The dynamic memory is implemented using standard POPLOG features - Flavour Objects and records.

The basic design is that of a tree with level 0 as root which grows incrementally as new knowledge is acquired. It grows only to a depth of six levels, but can spread in width by creating new branches. Within the links and nodes of this tree the face knowledge is represented either directly or by referencing other organising data structures. **Figure 7.0.** shows the arrangement of the first three levels. Level 0 is root, level 1 & 2 form a lattice. Level 2 contains the nodes which contain the Face_mop Frames.

The face knowledge is to an extent explicit in the labels of the links and nodes of the tree, but is also represented in a hierarchy of Flavour Object Classes which provides the slot and value functions and the demons for tally keeping, updates and tree reorganizations and possibilities for inheritance of values. Each node and its links is represented by a record. **Figure 7.1** shows the relationships between these representations schematically. In summary, the following data structures are used in memory:

(a) nodes: these are labelled m0.........m* and each is a record of two fields: an 'info' field which may contain a face action, or a feature on the face, or a reference to a MOP or an event, and a 'links' field which enumerates the names of the links to the daughter nodes and the names of these.

(b) links are labelled with an action e.g. 'raised' or 'open', or a feature, e.g. 'nose' in **Figure 7.1.**

(c) A superimposed organization, referenced by some node contents, is a hierarchy of POPLOG FLAVOUR OBJECTS which underpins the conceptual organisation of Face Knowledge. This is composed of Flavour Classes: 'Face_mops', 'fsub_mops', sub_mops' and 'evvents'(the spelling is deliberate).

A Face_mop organises the generalised knowledge about which face expressions typically accompany which primary emotion(there is thus one Face_mop for each of the six emotion expressions: happy, sad, afraid, angry, surprised, and disgusted). Face_mops may develop sub_categories further down their branches. An fsub_mop or sub_mop organises a more specialized generalised knowledge of an anomalous face action which has occurred in more than one input event and reflects the system's dynamic ability to learn typical face actions other than those contained in the Face_mop's Frame. The 'evvent' MOP plays a vital role in the dynamic re-organization of memory by keeping tallies of each input face description stored within a specific Face_mop. **Listings 7.1-7.8** show POP11 code for these Flavour Classes and for their instantiation.

## Listing 7.1: POP11 code for creating a Face_mop Class Flavour

```
flavour Face_mop;
ivars name  Face_event_count = 0, Face_event_list = [];
ivars node = false, support = [1000 1000 1000 1000 1000 1000 1000 1000
1000
1000 1000 1000 1000 1000
1000 1000 1000 1000 1000 1000];
ivars Face_sub_list = [], Face_fsub_list = [];
ivars gf_a = [], interps ;
ivars common = [], ivars interp ;
defmethod printself;  ;;;prints name of instance
   pr('<Face_mop_' >< name >< '>\n' )
enddefmethod;

defmethod support_attribute(attribute, update);
/*keeps check on frame f_a support: if support for a
frame f_a falls from 1000 to 0, the f_a is removed
and events indexed directly off the frame which
contain this f_a are re-indexed below the frame
using that f_a's labels as indexes.*/
enddefmethod;
endflavour;
```

## Listing 7.2: POP11 code for creating an instance of a Face_mop Class Flavour

```
consword('mop_happy_gen_Face') -> mop_happy_gen_Face;
make_instance([Face_mop name mop_happy_gen_Face gf_a
[cheeks raised  mouth bared mouth up  mouth open mouth u_lip_raised
 mouth slightly_open mouth wide mouth pulled eyes l_lid_raised  ]
interp happy common[cheeks raised  mouth bared mouth up  mouth open
 mouth slightly_open mouth wide mouth pulled eyes l_lid_raised interp
happy]
Face_sub_list [] Face_fsub_list [] Face_event_list []
Face_event_count 0
support [1000 1000 1000 1000 1000 1000 1000 1000 1000]
interps [pleased cheerful calm non_aggressive loving mesmerized
captivated relaxed   good_humoured pleasant_to_talk_to interested
not_angry friendly good_humoured  amused open pleasant not_serious
weighing_up thoughtful] node m8])-> mop_happy_gen_Face;
```

**Explanation:** All the Flavour classes have already been defined in a different file and the Flavours library package has been loaded in. The above code creates an instance of the Face_mop class. Its name is 'mop_happy_gen_Face'. A Flavour Class has a list of variables(ivars) which can take default values or remain undefined. There are only six instances of The Class: mop_happy_gen_Face, mop_sad_gen_Face, mop_angry_gen_Face, mop_afraid_gen_Face, mop_disgusted_gen_Face and mop_surprised_gen_Face, each of which have the same variables but the values of course differ. When an instance is formed, some or all of the Class variables are given values which overwrite the Class default variables. Others are updated in the course of processing user input. The values are accessed by sending a message to the ivars in question of a specifically-named instance.

In the Face_mop Class, the variables are: name, interp, common, Face_sub_list, Face_fsub_list, Face_event_list, Face_event_count, support, interps and node. All the Face_mop Class instances have these variables. These are explained in natural language below for the 'mop_happy_gen_Face' Face_mop: the other five in the Class correspond:

| | |
|---|---|
| 'name': | 'mop_happy_gen_Face'; |
| 'interp': | the basic emotion label which occurs in the name; |
| 'gf_a: | : the list of typical face actions of a happy face; |
| 'common': | has these typical face actions as well as the emotion label |
| 'Face_sub_list' | the name of a sub_mop formed is added to this list. |
| 'Face_fsub_list' | the name of a fsub_mop formed is added to this list; |
| 'Face_event_list' | the ID of an Evvent Class instance formed is added to this list; |
| 'Face_event_count' | number of Evvent Class instances under this Face_mop |
| 'support' | list of numbers, originally a list of 1000s: one for each typical face action in the same order as they appear in 'gf_a'. Each is the support given at the outset to its corresponding face action. The level of support varies with the user input. |
| 'interps' | a list of interpretations associated with the T-series photo graphs which were also attributed as 'happy' in the knowledge acquisition. |
| 'node' | the node in the Janus conceptual tree(actually the ID of a record) in which the Face_mop instance is referenced |

The Face-mop Class also has demons: the 'support_attribute' demon monitors Frame face action support. If this falls to 0, the face action is removed from the Frame and events indexed directly off the Frame which contain this action are re-indexed below the Frame using that action's labels as indexes.

The reader might wish to keep the LHS of **Figure 7.1.** in mind so as to appreciate the Flavour hierarchy. Below the Face_mop Class comes the fsub_mop Class which is described below: The Flavour Classes are pre-coded. The Face_mop instances are also pre-coded to be formed automatically when the Janus code is loaded. Instances of the Flavour Classes lower in the hierarchy, the fsub_mop, the sub_mop and the evvent instances are formed during runtime as required by the processing of face-knowledge input. Thus their names must be uniquely compounded by this knowledge and the name of the Face_mop Flavour instance under which they are created. The various values also are assigned to variables and have meaning derived from the processing of input. All instances of this Class use some of these.

### Listing 7.3: POP11 code for the fsub_mop Class Flavour:

```
flavour fsub_mop isa Face_mop;
ivars name Spec  fsub_sub_list  = [], path  fsub_event_count = 0;
ivars fsub_event_list = [], node = false,  common gf_a ;
defmethod printself;
   pr('<fsub_mop_' >< name >< '>\n')
enddefmethod;

defmethod after initialise;
;;; adds the name to the list of fsub-mops that the Face_mop has
unless Spec = false or Spec = undef then ;;;  '<-' means send a message to ...
^name :: (valof(Spec) <- Face_fsub_list) -> valof(Spec) <- Face_fsub_list;
;;;  'Spec' is the Face_mop.
endunless
enddefmethod;
endflavour;
```

**Explanation:** Variables such as fsub_event_list & _count(listing & count of events in the leaves under this fsub_mop have to await developments before they are upgraded. The defmethod after initialise is a demon which upon instantiation, adds the name of the new fsub_mop to its Face_mop's list of fsub_mops in its scope.

### Listing 7.4: POP11 code for creating an instance of a fsub_mop Class Flavour

```
substring(5,(length(lastmop) - 13),lastmop) -> mopem;
consword('mop_' >< mopem >< b >< '_fsub') -> tag;
```

```
make_instance([fsub_mop name ^tag common [^b] gf_a [^b]  Spec ^Face   path ^road
node ^nodename]) -> valof(tag);
```

**Explanation:** 'mopem' is assigned the 'happy' part of the Face_mop name and this is conjoined to 'mop', the node match and 'fsub' to automatically name the fsub-class instance, e.g. 'mop_happybrows_fsub'. This is assigned to 'tag'. Some of the Class variables are instantiated: Only the 'name', 'common', 'gf_a', 'Spec', 'path' and 'node' Class variables were instantiated, so the default values of these are overwritten for this instance. The fsub_mop is instantiated when an input anomalous face action e.g. 'brows raised' traverses an existing path indexed below the Face_mop and the 'brows' component enters a node already containing 'brows'. This match triggers the instantiation and the matching component, 'brows', is assigned to 'common' and 'gf_a', the Face_mop label is assigned to 'Spec', the road traversed, to 'path, and the node at which it happened, to 'node'. The make_instance command has a parameter list which first names the Flavour Class and then a set of ivars value pairs.
The instance is assigned to its name.

## Listing 7.5: POP11 code for the sub_mop class flavour

```
flavour sub_mop isa fsub_mop;
ivars name Spec fSub path sub_event_count = 0, sub_event_list [];
ivars node = false, gf_a = [], common = [] ;

defmethod printself;
pr('<sub_mop_' >< name >< '>\n')
enddefmethod;

defmethod subb;
lvars v;
unless Spec = undef or Spec = false   then
(^name :: (valof(Spec) <- Face_sub_list)) -> (valof(Spec) <-   Face_sub_list);
endunless;
unless fSub = undef or fSub = false then
 (^name :: (valof(fSub) <- fsub_sub_list)) -> (valof(fSub) <- fsub_sub_list);
endunless;
enddefmethod;
endflavour;
```

**Explanation:**

The Class variables: name, Spec(= Face_mop), common(=matching f_a), gf_a(= matching f_a), support(=20), fSub(the fsub_mop), and node(= node in the tree) become assigned during runtime.

**Listing 7.6: POP11 code for creating an instance of a sub_mop class flavour**

```
........< other code>.............
elseif ((gg matches [?x : nevtec ?y : nevtec]) and
(Info matches [?w : nevtec ?z : nevtec]) and
(entry matches [ == ?v3])) and [^x ^y] = [^w ^z]
then
substring(5,(length(lastmop) - 13),lastmop) -> mopem;
(consword('mop_' >< mopem >< w ><'_'>< z >< '_sub')) -> tag;
make_instance([sub_mop name ^tag  Spec ^Face common [^x ^y]
gf_a [^x ^y] support [20]fSub ^fmopp node ^nodename]) -> valof(tag);
```

## Explanation:

A sub_mop is instantiated when there is a juxtaposition in a node of two identical face actions e.g. 'brows raised' and 'brows raised'. This happens when an input f_a traverses the same path in a mop tree as did a prior identical f_a. The match is coded in (a), where the match excludes a matching of event IDs(e.g. ev0 & ev0). The label-stem of the Face_mop name is bonded to the f_a and the terminator '_sub' and the new sub_mop(e.g. 'mop_happybrows_raised_sub') assigned to it and given a support of 20. (This will not be updated unless it gains promotion to its Face_mop). The variables instantiated in this instance have been explained before except fSub which is given the name of the fsub_mop parent mop_happybrows_fsub which is the value of 'fmopp'.

**Listing 7.7: Skeletal POP11 code for the evvent class flavour(the 'vv' is intentional).**

(full code in Appendix)

```
flavour evvent is a sub_mop;
ivars name Spec= false, fSub = false, path node input sub_match ;

defmethod printself;
pr('evvent_'><name><'>\n) ;;;prints the name of an instance
enddefmethod;

defmethod before initialise(i) ->i;
('i' is a list of ivars instantiations for the particular input event. This procedure prints out details of the storage of the input face event. It is listed in the Flavour code in the Appendix).
enddefmethod;
```

**defmethod after initialise;** (this complex code is omitted; its purpose is described)
/*the number of events indexed under a sub_mop is monitored: if it exceeds 2/3 of those under the parent Face_mop the sub_mop is moved into the content frame of the Face_mop. This does not apply to the first 6 events.

If the event is indexed directly off the Face_mop Frame, the name of the event is added to the Face_mop's list, of events under it and this count is updated; else if indexed below a sub_mop, both the Face_mop and sub_mop's list and count are updated. The calculation for triggering promotion is monitored and if triggered, the following steps are carried out: the promoted sub_mop face action is given a support of '20' and added to the front of the Face_mop Frame f_a list; the support to the front of the support list; the sub_mop is collapsed: its link in the fsub_mop's node's record is deleted; the sub_mop's node's record's 'info' field has the sub_mop's reference removed; the sub_mop's database record is removed and its daughters added to the Face_mop's db list of daughters and the sub_node's record is rendered undefined*/.
**enddefmethod**

**endflavour;**

## Explanation:

The evvent Flavour instances are made mainly to perform the monitoring of the events indexed under sub_mops, so that sub_mop promotions can be detected rapidly in a branch and carried out immediately. Their ID is constructed automatically, as each input of face actions is processed, out of the string 'EV_' concatenated with the unique ID (e.g. ev0) given to the face description on input. In the case that all of the input face actions are matched in a Face_mop, there is only one representation of that input(ev0) in the system. The case is different if several atypical face actions each identify the same input in a distributed way in the leaves of different branches of the tree under the one Face_mop. Different evvent Flavour instances are made for each under the same name(e.g. EV_ev0), the first being overwritten by the second etc.. The promotion tallies are updated as each is made, so that they are dispensable after this has been done.
The Class ivars are Spec (= Face_mop instance at the root of the branch of the memory tree to which this leaf belongs), fSub and Sub(the fsub_mop & sub_mop), path(the path traversed by the input to arrive at this leaf), the node of the leaf, paths(other paths leading to leaf nodes where the input is represented) and sub_match( the face action in the sub_mop).

## Listing 7.8: Skeletal POP11 code for creating an instance of the evvent class flavour(the 'vv' is intentional).

```
(consword("EV_" >< hd(E) )) -> tag; ;;;;'E'= an enlisted event ID e.g. [ev0]
make_instance([evvent name ^tag Spec ^r fSub ^q Sub ^pppp sub_match ^ub_match
path [^^way ] node ^nodal ]) -> valof(tag); ;;; 'tag' holds the unique ID of the instance.
```

Although conceptually, a Face_mop organises all face expressions which have much in common with a specific primary emotion and which have traversed that branch of memory headed by its Frame, it is represented in Janus as a hierarchy of Flavour objects in one perspective and as a branch of a tree in another. The dual representations maintain a comparable ordering of level. Each input is automatically given an unique ID which comes to rest in a leaf node in memory indexed by "event" a link reserved for whole input event IDs. It is represented in its entirety as an instance of the 'evvent' Flavour, but is distributed in the branch of memory where it is stored according to its anomalous face actions(those differing from the typical) if any. So, as well as being distributed, it is also held intact, and preserves its identity as well as forming part of abstractions.

**7.6 Traversing a MOP:**

Any new event, having selected a Face_mop, traverses the branches below the Frame with any atypical face actions it might possess (interpretations only feature in the full event in the leaves from which they are retrieved). Each face action component(i.e. 'nose' first then 'flared' at the next level) in turn traverses the tree if the links labelled identically exist (else they are created) until it reaches an identical match in a node content previously encountered (so descent is by links and matching by node). This results in 'reminding' and either fsub_mop or sub_mop formation (or both) located at that point in the tree. Events which have not been encountered before are automatically incorporated into new branches of the tree.

The leaf of each branch represents the input events which have pursued that route (in total) by their IDs. The one event is thereby distributed in memory according to the number of anomalous face actions which the event contains(if the event face actions are all 'swallowed' by the Frame, its ID is indexed directly off the Frame only). Frame face action components whether in Face-MOP, fsub_mop or sub_mop(they each have a Frame) are recognised as being typical for a section of experience comprising the events organised beneath them and the memory is automatically restructured to reflect this. Repeats of the generalised face action following the same pathway do not need to be re-indexed. This limits complexity of the tree.

Face_mops demonstrate another advantage of generalisations: being able to give a more general interpretation to an event input in search of an interpretation, when a leaf node interpretation cannot be given because the query does not contain the requisite face actions with which to unlock paths to a leaf. The predictive power of generalisations is not made use of in the current system ( though it is very important in IPP(Lebowitz, 1980) ) nor is their power to assist in the elaboration of input lacking the requisite indexes (prominent in Cyrus ).

**7.7 Selecting a Face_mop to traverse:**

Face_mop frames contain a varying number of feature actions. The details are given in **Table 4.0.**

The input event is channelled to traverse only one Face_mop in Memory: the number of matches between input event and Face_mop feature actions decide this. The number of matches, divided by the number of face actions in the Face_mop, is compared for each Face_mop. The winner is traversed. If this results in a tie, a heuristic decides the issue(**Listing 7.9**). This compares the input list of face actions to a few salient ones picked out for each emotion (happy, sad, afraid, angry, disgusted, surprised) based on Ekman & Friesen(1984)'s descriptions. If this also fails the user is informed that there is insufficient input on which to make an interpretation. At least one MOP-listed feature action must be input for the System to run.

The decision to traverse a particular Face_mop interprets the input as a general category of emotion (remember that the frame feature actions are based on pancultural emotions).

**Listing 7.9: Heuristic rules for Face_mop selection:**

Selection depends on specific face actions in the input face expression event viz:

```
IF                eyes lower-lid-raised
          and     brows raised
THEN              MOP_surprised_gen_Face

ELSE IF           eyes lower -lid-tensed
          and     eyes lower-lid-raised
          and     brows raised
THEN              Mop_afraid_gen_Face

ELSE IF           mouth up
          and     not(nose screwed)
THEN              Mop_happy_gen_Face

ELSE IF           brows lowered
          and     brows contracted
          and     eyes inlid_raised
          and     (mouth compressed or mouth wide)
          and     not(mouth upper-lip-raised)
THEN              Mop_angry_gen_Face

ELSE IF           (mouth upper-lip-raised
          and     mouth upper-lip-tensed
          and     (mouth lower-lip-raised or mouth lower-lip-lowered))

          or      (nose screwed and cheeks raised and eyes lower-lid-raised
                  and brows lowered)
          or      (nose screwed and cheeks raised and eyes lower-lid-raised
                  and mouth upper-lip-raised)
          or      (mouth upper-lip-raised and mouth lower-lip-everted and
                  cheeks raised and nose screwed)
          or      (mouth upper-lip-raised and (mouth lower-lip-raised or mouth
                  lower-lip-lowered) and (nose screwed or cheeks n-l-vert))
THEN              Mop_disgusted_gen_Face

ELSE IF           (brows centre-raised and eyes inlid-raised and eyes lower-
lid-                  raised)
          or      mouth down or brows centre-raised
THEN              Mop_sad_gen_Face
```

**7.8 Learn mode:**

In learn mode, if no atypical feature actions are present in the event the new event-ID is indexed directly off the frame of the Face_mop. If atypical feature actions are present for which a path already exists, the sub-tree is traversed with these as before and the event-ID is added to the existing leaf node. Any atypical actions not already present are created as separate branches and a new leaf node is added indexed by "event" in which the event-label is represented. In all cases a new instance of the 'evvent' object class is also created (the 'vv' is intentional). This details all the useful information about the event including its sub_mop and Face_mop, and also monitors the number of events gathered under these and re-organizes the knowledge structure automatically. The new event ID is added to the information field of the relevant node records. Remember that the input event in learn mode contains the interpretation as well as face actions. This is its only representation in the system and is therefore stored distributively with the intact input event.

The initial state of the tree is shown in **Figure** 7.0 where some of the links have been omitted for clarity. The root node (m0) is linked by six 'feature' links to the first level nodes(m1 - m6). These are in turn connected by action links to level 2 nodes(m7 - m12) each of which reference the typical face actions of one of the six basic emotions via the Face_mop frames named after their labels. This is the Face_mop Frame level.

New events are incorporated as shown in **Figure 7.2** An event, ev0, related to m12 but differing from it in two respects (mouth pulled, cheeks raised) is entered. The differences are indexed below m12 and two branches are created, each consisting of the sequence:

feature~<feature>~action~<feature-action>~event~<ev0>

where < ... > represent nodes in the tree and ~ are links.

All sub-trees below second - rank nodes(Face_mop frames) have this sequence, The same event is indexed twice(at nodes m15 and m18), and could be accessed(remembered) if either of the two actions occur in a subsequent event traversing the same path.

The dynamic reorganisation of memory is further illustrated in **Figures 7.3, 7.4 & 7.5.** Identical events ev1,ev2 which differ from the typical (Face_mop) in having 'cheeks raised' are indexed below that node(**Figure 7.3**). The two are then collapsed into a single branch(sub_mop) in **Fig. 7.4.** After five further occurrences of the same event, Janus decides that this is a 'typical' situation and promotes the action 'cheeks raised' to Face_mop level and indexes the events directly off that node(**Figure 7.5**). The changes involved in reorganizing Memory and the procedures which bring changes about are discussed in detail under "Reorganization of Memory" below.

### 7.9 Retrieval mode:

In retrieval mode, face actions only are entered in search of an interpretation. If the input actions are subsumed by those in the chosen Face_mop frame then the basic emotion in the MOP name is returned. If some of the input actions are atypical, the sub-tree is traversed in search of similar learned events and if any are found, the corresponding interpretations are returned.

### 7.10 Control structure:

The control structure in LEARN MODE labels the input with an unique ID. The input event is then matched to each Face_mop Frame to determine the Face_mop to be traversed. Constituent feature actions (f-a) which differ from the Frame are used to create two-tier indexed branches which are labelled with the names of the particular face action components. Depending upon the circumstances the anomalous face actions forge branches of their same label(see **Figure 7.2)** , creating links, nodes, records, and flavour class instances where these are required, and if these do not exist already as a result of previous input(in which case the input feature actions(f-a) are used to unlock links in the tree. The full input event is preserved and is indexed by a unique "event" link to a leaf node. When a node is encountered with a content, a juxtaposition occurs: the arriving new input f-a component is compared with the existing node content. **Figures 7.3 & 7.4** illustrate the process. The legal comparisons in a juxtaposition are detailed in **Table 7.1.**

**Fig. 7.3:** Identical events differing from the typical are indexed below m12

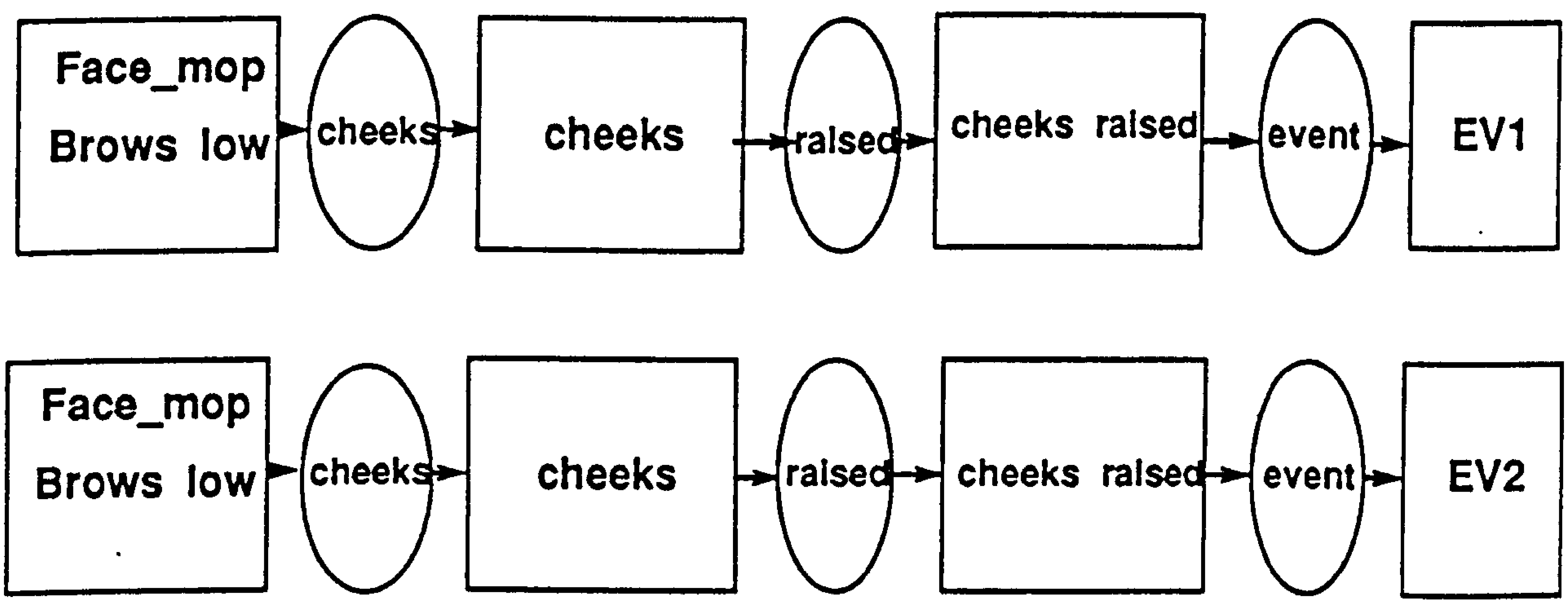



**Fig 7.4:** fsub_mop and sub_mop formation

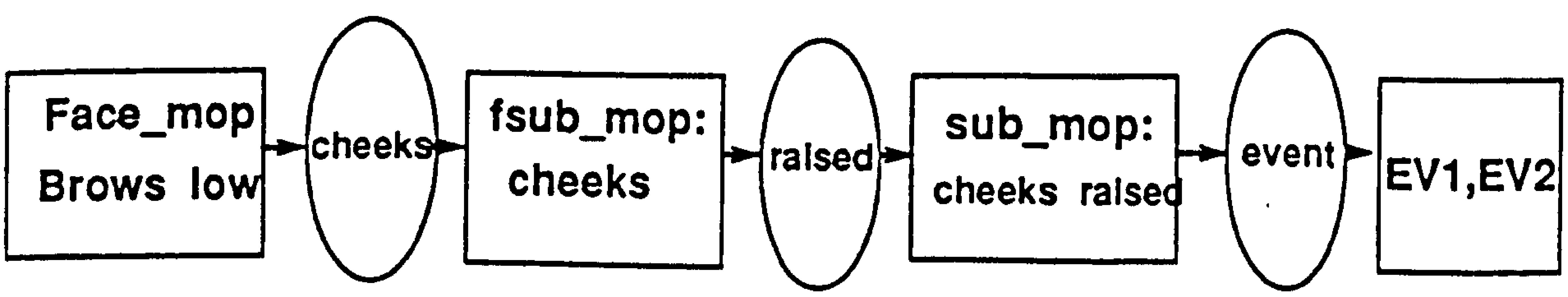
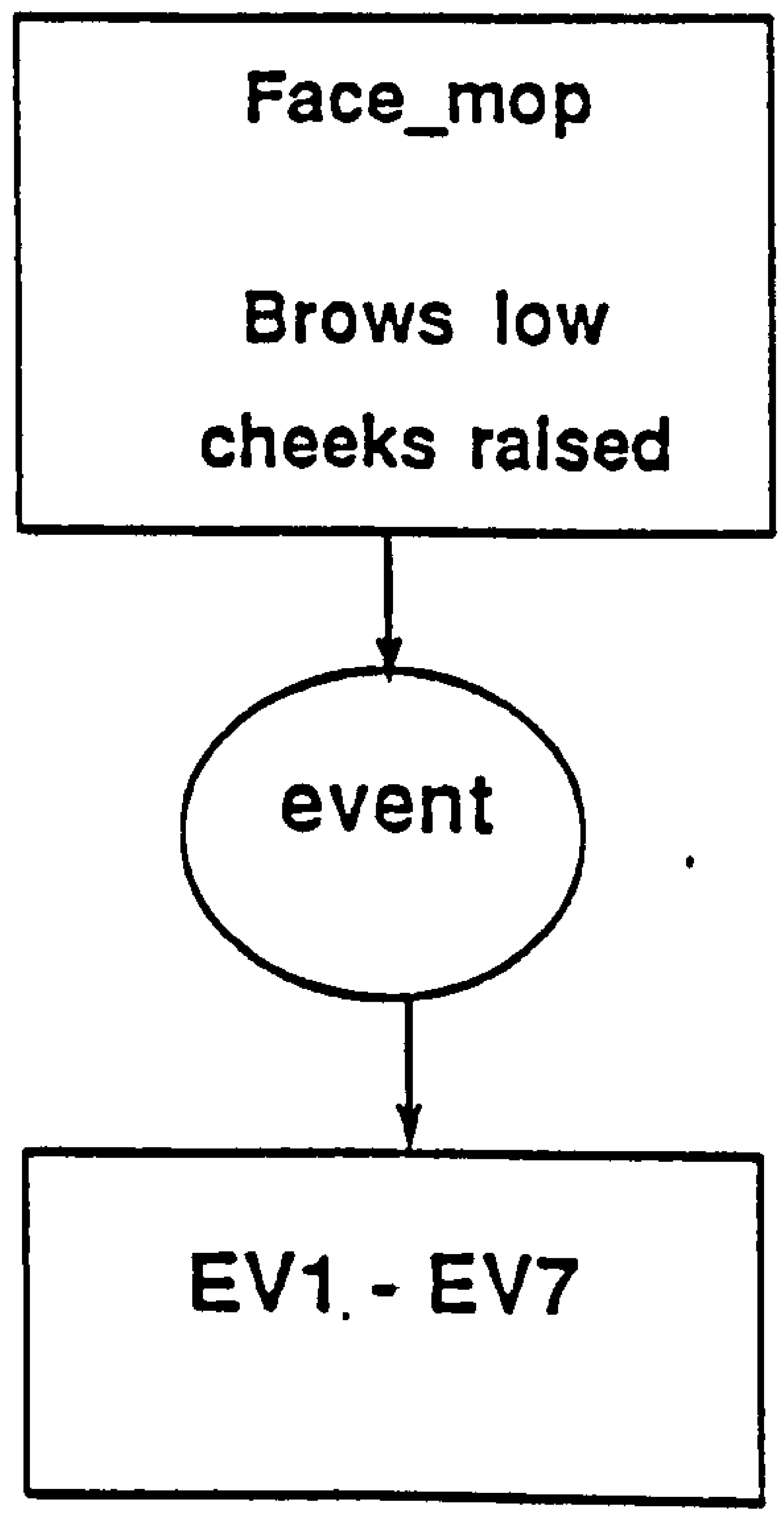


Fig. 7.5: promotion of sub_mop




The control structure in RETRIEVE MODE uses the input feature actions(f-a) to unlock existing links of the same labels in the tree. When a leaf node is encountered, the emotion labels of the contained events are returned. If progress to a leaf is denied because the input lacks the necessary keys, only the general emotion of the specific Face_mop traversed is returned.

### 7.11 Re-organization of memory:

Two types of re-organizations are involved - demotion of a face action from a (level 2) Face_mop Frame and promotion of a face action from a sub_mop to a Face_mop.

Following Kolodner(1984), the following tallies are kept: **a**: of how many times a Mop Frame f-a is supported by the input and **b**: of how many times an atypical feature-action

Table 7.1: Juxtapositional interaction in nodes: Input events are decomposed by the control system and traverse the tree coming into nodes alongside what content is there already. The consequences include MOP formation, etc.

| Input | Node content | Consequence of match |
|---|---|---|
| feature | nil | feature placed in node. |
| feature | feature | instance of fMOP Flavour Class formed. It's ID replaces feature in node and is added to database. |
| feature | fMOP | no action - subsumed |
| feature action | nil | feature action placed in node and added to database. |
| feature action | sub_mop | subsumed into MOP |
| feature action | feature action | instance of sub_mop Fla vour Class formed. It's ID replaces feature action in the node and database |
| event ID | nil | event ID placed in node and instance of Flavour Evvent Class formed. Event ID added to database |
| event ID | event ID(s) | both event IDs placed in the node. An instance of Flavour Evvent Class formed for the new event and the ID is added to database |

in a sub_mop Frame is supported by the input. Thresholds are associated with each and if the tally count exceeds them, re-organization automatically follows. **In the case of a:** each Face_mop f-a starts with a tally of 1000 points and 1 is added for support and -1 is added for non-support. A f-a whose tally drops to 0 is relegated from the frame and indexed below it. **In the case of b:** each sub_mop f-a starts with a tally of 20 points and will not be updated unless it is promoted to its Face_mop. A tally is kept however of how many events are indexed under it. A sub_mop f-a whose tally reaches 6 is kept under continuous scrutiny for possible promotion to Face_mop status. When the events indexed under such a sub_mop account for 2/3 of those indexed under the Face_mop( which include these and those indexed under any other sub_mop of that Mop and those indexed directly under the Frame of the MOP ), then the f_a is promoted from the sub_mop to the Face_mop frame and its events indexed below the latter. In b, the sub_mop is collapsed, since only one f_a is represented in the Frame of the sub_mop; in a, the demoted f_a is indexed by its label directly below the Face_mop Frame. The re-organizations can be seen by inspection of the database before and after the triggering events.

In the listings (**Listing 7.10** to **Listing 7.16**) which follow, the selected changes in the database made by subsequent input are displayed and discussed. Each Face_mop f_a is given a support of 1000 in the untrained system. Since removal occurs when support falls to 0, for this demonstration the support for one of mop_sad_gen_Face's Frame f_a (brows in_raised) has been started at '1'. If a face action of an input event's f_a matches one in the Frame, the support of the latter is increased by 1. Frame face actions not supported by that input lose 1. Since the database below mop_sad_gen_Face is unchanged in these manoeuvres, it will only be displayed once, thereafter its presence will be indicated by <.....>.

**Listing 7.10: the untrained database in Janus:**

**Comment:** The database consists of lists.
Each list represents a record which represents a node and its links in the tree.
The format is: [[content][links] node name].
The database is used only to recreate the records

```
[
 [[mop_angry_gen_Face][]m13]                        |
 [[mop_afraid_gen_Face][]m12]                       |
 [[mop_disgusted_gen_Face][]m11]                    | level 2
 [[mop_sad_gen_Face][]m10]                          |     (the Face-mops)
 [[mop_happy_gen_Face][]m8]                         |
 [[mop_surprised_gen_Face][]m7]                     v
 [[][raised m8 n_l_vert m11 raised m11] m6]
 [[] [drop m7 tense m11]m5]
 [[] [drop m7 smile m8 open m7 open m12 open m13 compressed m13
 u_lip_raised m11 l_lip_lowered m11 l_lip_lowered m11 pulled m12
 down m10 down m11 shut m8 slightly_open m7 wide m7 wide m13
 l_lip_everted m11 u_lip_everted m11 l_lip_raised m11
u_lip_tensed m12 l_lip_tensed m12 square m13
  open m8 open m12 open m13 bared m8 up m8 bared m7 bared m13]m4]
 [[] [screwed m11 flared m11 flared m13 ]m3]
 [[] [  narrowed m11 wide m7 wide m12 l_lid_tensed m13 l_lid_tensed m12
      u_lid_tensed m13 l_lid_raised m13 u_lid_lowered m13
    l_lid_raised m10  u_lid_lowered m8  narrowed m13
    l_lid_lowered m7 u_lid_raised m12 l_lid_raised m12 wide m7
    u_lid_raised m7 l_lid_raised m11 in_lid_raised m10 down m10 ]m2]
 [[] [raised m7  in_raised m10 low m11 low m13
   raised m12 contracted m12 centre_raised m10 contracted m10
  contracted m13 straight m12]m1]
 [[] [brows m1 eyes m2 nose m3 mouth m4 jaw m5 cheeks m6 ]top1]]
```

**Comment:** This corresponds to Figure 7.0. The Face_mops in m7 - m13 are in the level 2, where the "top1" node is level 0. The order of f_a in mop_sad_gen_Face(in bold type, above) is important since the support, each f_a has, is in the same order in a list. One of these f_a ("brows in_raised") is to be demoted. Its support will be set at 1 to start with. Input events which contain this f_a will increment its support by 1, whereas input events which do not will cause its support to be reduced by 1. When its support reaches 0 its demotion will be reflected in the database.

**Listing 7.11: the changes in the database of Table 7.2 following input of a face event, ev0:**

[brows contracted brows centre_raised brows in_raised eyes l_lid_raised mouth down eyes down nose drooped cheeks dropped interp sad]

[[ [ ev0] [] m19]
[[cheeks dropped] [event m19] m18]
[[cheeks] [dropped m18] m17]
[[ev0] [] m16]
[[nose drooped] [event m16] m15]
[[nose] [drooped m15] m14]
**[[mop_sad_gen_Face] [cheeks m17 nose m14] m10]..............**

**Comment:** This represents new growth in mop_sad_gen_Face in response to an input face description. The input event contains two f_a which are not in the 'sadMop' Frame: 'cheeks dropped' and 'nose drooped'. These are indexed below the Frame in two levels: 'cheeks' in m17 and 'nose' in m14 both at the 3rd level and 'cheeks dropped' in m18 and 'nose drooped' in m15 both at the 4th level. The event as a whole is represented in m16 and m19 in the 6th level, indexed by 'event' in the 5th level. The input event adds 1 to the support of all Frame f_a it contains and 'brows in_raised's support is now 2.

**Listing 7.12: the changes in the database of Table 7.3 following input of an event with the same two anomalous f-a as that of ev0: ev1:**

**Comment:** The input face descriptions are identical and are as follows:

[brows contracted brows in_raised brows centre_raised eyes l_lid_raised eyes inlid_raised eyes down mouth down nose drooped cheeks dropped interp glum]

[[[ev1 ev0] [] m19]
**[[mop_sadcheeks_dropped_sub] [event m19] m18]**
**[[mop_sadcheeks_fsub] [dropped m18] m17]**
[[ev1 ev0] [] m16]
**[[mop_sadnose_drooped_sub] [event m16] m15]**
**[[mop_sadnose_fsub] [drooped m15] m14]**
[[mop_sad_gen_Face] [cheeks m17 nose m14]m10]............

**Comment:** mop_sad_gen_Face is again traversed. This time the atypical face actions find identical components to themselves in the nodes of the 3rd and 4th level. This results in 'reminding' and MOP formation: mop_sadnose_fsub in m14 and mop_sadcheeks_fsub in m17 at level 3 and mop_sadnose_drooped_sub in m15 and mop_sadcheeks_dropped_sub in m18 at the 4th level. In each case, the two events contributing to the new MOPs are indexed below. The support of "brows in_raised" is now 3.

### 7.11.1 Demoting a Face_mop frame face action:

When the support for a Frame f_a falls to 0 it is automatically demoted from the Frame and indexed below it. The process is described below:

**Listing 7.13: Changes in the database of Table 7.12 following the input of three identical face descriptions:**
**[mouth down interp sad]:**
**These are identified by ev2, ev3, & ev4.**

**Comment:** Since these do not contain 'brows in_raised' the support for this f_a will be reduced to 0 triggering its expulsion from the Frame of mop_sad_gen_Face. The following is the database after this has taken place. Each line represents a record of a node and its links in the Face_mop tree in the format:

[[ info field ] [ links field: index & node pointer ] node]

[[[ev4 ev3 ev2] [] m20]
[[ev0 ev1] [] m23]
**[[mop_sadbrows_in_raised_sub] [event m23] m22]**
**[[mop_sadbrows_fsub] [in_raised m22] m21]**
[[mop_sad_gen_Face] [brows m21 event m20 cheeks m17 nose m14]m10]
[[ev1 ev0] [] m19]
[[mop_sadcheeks_dropped_sub] [event m19] m18]
[[mop_sadcheeks_fsub] [dropped m18] m17]
[[ev1 ev0] [] m16]
[[mop_sadnose_drooped_sub] [event m16] m15]
[[mop_sadnose_fsub] [drooped m15] m14]......................................

**Comment:** 'Mouth down' is a Frame f_a of mop_sad_gen_Face so these input events are indexed off the frame by the index: 'event'. 'Mouth down's support is increased by 3 but all the other frame face actions have their support cut by 3. This proves disasterous for 'brows in_raised'. It's support is reduced from 3 to 0 causing its expulsion from the Frame and the re-indexing of all events which ever choose this Face_mop and which contain 'brows in_raised' directly below. If there is only one such it would appear in m21 as [[brows] [in_raised m22] m21] but if more than one, mop-formation proceeds in the usual way as shown above, where fsub- and sub_mops are seen in m21 and m22 respectively. Events ev0 and ev1 are indexed below the sub_mop in m23 node while the three events containing just 'mouth down' are indexed directly off the Face_mop Frame by the link 'event'. Had they contained 'brows in_raised', its support would not have fallen to 0, of course; but had they contained any other atypical 'brows' action, the links field of m21 would have contained the action and a reference to the node containing that face action. Frame-indexed events containing just 'brows in_raised' are deleted.

The face actions in the Frame of a Face_mop can be accessed by sending a message to its Object:

mop_sad_gen_Face <- gf_a

reveals that 'brows in_raised' has indeed gone from the list. 'Brows in_raised' now can be re-instated to the Face_mop if more than 2/3 of input to mop_sad_gen_spec contains it at any time in the future beyond the next 4 entries.

### 7.11.2 Promoting a sub_mop face action:

When a sub_mop face action proves to be the rule for a sizeable fraction of the face expressions indexed under the Face_mop, it is promoted to Face_mop Frame status.

This will be traced by changes in the database. In order to keep things simple, the fate of just one face action: 'nose drooped' will be tracked. A situation will be contrived where two identical events ( ev0,ev1 ) are input in sequence containing two face actions not represented in the FRAME of mop_sad_gen_Face. These are input to an untrained Janus whose database representation is in **Listing 7.10** As before, since the changes will affect only mop_sad_gen_Face and its daughter nodes, only this part of the database will be shown. The inputs will be selected to choose mop_sad_gen_Face They are each:

[brows contracted brows centre_raised brows in_raised eyes l_lid_raised mouth down eyes down nose drooped cheeks dropped interp sad]

These contain two face actions ('nose drooped' and 'cheeks dropped') which are not included in the Frame and their serial input to memory produces sub_mops indexed below mop_sad_gen_Face named 'mop_sadnose_drooped_sub' and 'mop_sadcheeks_dropped_sub' and their respective fsub_mops, 'mop_sadnose_fsub' and 'mop_sadcheeks_fsub'. Ev0 and ev1 are the names which represent these input events and these are indexed below each sub_mop by the link: "event". **Listing 7.14** shows the relevant part of the untrained database, and **Listing 7.15**, the changes wrought by these inputs.

**Listing 7.14: Fragment of the untrained database.**

```
[[mop_angry_gen_Face][]m13]
[[mop_afraid_gen_Face][]m12]
[[mop_disgusted_gen_Face][]m11]
[[mop_sad_gen_Face][]m10]
[[mop_happy_gen_Face][]m8]
[[mop_surprised_gen_Face][]m7] .............
```

**Listing 7.15: Changes wrought by inputs ev0 and ev1**

```
[[ev1 ev0] [] m16]
[[ev1 ev0] [] m19]
[[mop_sadcheeks_dropped_sub] [event m19] m18]
[[mop_sadcheeks_fsub] [dropped m18] m17]
[[mop_sadnose_drooped_sub] [event m16] m15]
[[mop_sadnose_fsub] [drooped m15] m14]
[[mop_sad_gen_Face] [cheeks m17 nose m14] m10]...........
```

**Comment:** The two events led to MOP formation in m14, m15, m17 and m18. The two events are indexed in m16 and m19

Further input events are now arranged to demonstrate promotion of the face action in mop_nosedrooped_sub: 'nose drooped', by inputting events which contain that face action as the only one different from the FRAME face actions. After six inputs indexed below this sub_mop, if the number of events indexed below the latter number more than two thirds of all the events indexed below mop_sad_gen_Face, then mop_sadnose_drooped_sub is collapsed; its links are destroyed and 'nose drooped' is included in the FRAME of mop_sad_gen_Face. In order to bring this about, the event:

[brows contracted brows centre_raised brows in_raised eyes l_lid_raised mouth down eyes down nose drooped interp sad]

is input five times. The resulting changes to the database are evident in **Listing 7.16**, and the happening is signalled in the run by:

```
              PROMOTING SUB_MOP GENERALIZATION :
               moving the sub_mop frame content:
**********************  [nose drooped]*********************
*************into the Face_MOP: mop_sad_gen_Face***********

***********************************************************
                       MOVE COMPLETED:
```

**Listing 7.16: database of Listing 7.15 after the collapse of mop_sadnose_drooped_sub**

```
[ [[mop_sad_gen_Face] [event m16 cheeks m17 nose m14] m10]
  [[mop_sadnose_fsub] [] m14]
  [[ev6 ev5 ev4 ev3 ev2 ev1 ev0] [] m16]
  [[ev1 ev0] [] m19]
  [[mop_sadcheeks_dropped_sub] [event m19] m18]
  [[mop_sadcheeks_fsub] [dropped m18] m17] ...........]
```

**Comment:** The support of mop_sad_gen_Face has "nose drooped"' 20 up front:[20 1015 16 1015 1015 985 1015 1015 1000 1000 1000 1000] to match the face action's position in the Face_mop's Frame face actions:

[nose drooped brows contracted brows in_raised brows centre_raised eyes l_lid_raised eyes inlid_raised eyes down mouth down].

Mop_sadnose_drooped_sub has disappeared. Mop_sadnose_fsub remains since there may have been other nose actions indexed below it. The eight events indexed under the former sub_mop have been re-indexed below mop_sad_gen_Face directly. The new FRAME face action in mop_sad_gen_Face is given a support of 20 (a limitation of that of endowed neighbours which have a support of 1000 to imply that it is still provisional). Links to the former sub_mop are erased.

### 7.12 The database and saved images:

New knowledge is acquired by teaching the system: e.g. through humans judging photographs directly. The environment has to be recreated between learning sessions, unless acquisition can be accomplished in one session. All instances of Flavour Classes and records with their bindings will need to be recreated. A saved image in the POPLOG environment will preserve the state of the environment.

If an untrained system is required, the control structure recreates the levels 0-2 of the tree in the same way. An untrained database is used for this purpose(**Listing 7.10**).

This represents each record making up the tree levels as a list. These lists are used to recreate the records and reset the automatic event ID-labelling facility.

As experience adds new nodes to the tree and new records are constructed, they are given unique IDs and represented in the database by lists and updated when required e.g. when MOPs are created and links added as described above.

The tree may be re-created from the database. Some face knowledge is stored explicitly in the database; while some is stored implicitly in chunking representations which are referenced by their unique labels. Such knowledge can not be re-loaded by re-creating the tree - hence the need for saved images. It is therefore not a flat knowledge base, and items retrieved from it may only refer to memory structures which contain face knowledge. The updating is done automatically in the manner as described above.

### 7.13 Psychological considerations:

The main purpose of this section is to say that it is not clear whether Janus emulates human performance in respect to applying Schank's theory to the perception of facial expressions.

The first assumption in doubt is whether 'reminding' is the usual way in which people interpret faces. In contrast to autobiographical life events, it is not so intuitive to us that the face expressions we observe have to remind us of similar expressions before their meaning becomes apparent: that for instance, the smile that does not reach the eyes reminds us of previous experiences of it. It might well do so, depending on our experiences.

A second doubt concerns whether other persons' experience enters into interpretation to the degree that it does in Janus. That we do espouse other persons' mental constructions is well accepted (Kelly 1955), but faced with an expression, do we call to mind a list of alternative labels?.

Another doubt is whether Schank's proposed conceptual structures accommodate face expressions gracefully. Schank conceived a MOP as a conceptual structure which organised a set of scenes with a common goal. The scenes contained knowledge, not the MOP. In extending the theory to face expressions, Face-mops organize feature actions with the goal of communicating an emotion. In the application to Janus, a feature action is such a "scene", and each Face_mop organizes the face actions in the tree below its Frame.

Each face action in life may take part in many expressions and is a complex succession of micro events in time and space, including biochemical, structural and psychological events, the external manifestation of which is a distortion from rest of a feature whose movement has an onset, apex and offset(Ekman & Friesen,1984). The concept: "nose flared" abstracts such events for one such feature action. At the micro level biological scripts (ordered sequences in time) probably organize events. It might however be preferred by some to represent face actions as Scripts.

A Face_mop is a tree in which the nodes are either MOPs or face action components or events and the links are of three kinds: a feature, a face action and "event". A Face_mop is very different from the "Recognition Unit" which has been proposed to explain the recognition of familiarity in faces (Hay & Young 1982), Bruce & Young 1986), or the 'Face Register" in Ellis's model: Ellis,1986). Hay and Young(1982) proposed that there exists a separate Face Recognition Unit for each known person and that another is created for each acquaintance made. The pancultural evidence indicates that whatever is the mechanism in the human for recognising emotions, a core of emotional expressions are familiar to all peoples.

Quite apart from Schank's theory, Janus assumes that interpretations are grounded in a face feature analytic approach, albeit with an implicit configurational approach. The review of the literature in Chapter 2 shows that opinion is not so definitive on the matter (Wells & Hryciw, 1984; Sergent, 1984; Jensen, 1986).

The reason why such questions have not been researched here is that Janus is an expert system not a model of a theory of human memory, and the task was to represent domain knowledge. The task required to scientifically test a theory of human memory is quite different. It might, for instance, involve a statistical analysis of retrieval times for group of experimental subjects exposed to various orders of 'primes'(orienting tasks given before retrieval cues which may 'facilitate' or speed up retrieval depending on whether they address the way in which memories are organized. Such experiments would shed light on the organization of human memory, but they are considered to be beyond the scope of this thesis. The knowledge acquisition stage extracts the experts judgements of instances and reasons for these. Even if the reasons given for these interpretations have little to do with how expressions are interpreted, they have a lot to do with how, when asked for an explanation, they go about the task. Of course agreement between people is never perfect but there was sufficient agreement on configuration of face actions to represent face expressions using these.

Although universally-recognised emotions suggest the probability of innate factors in expression perception, this is not to say that they act in a filtering role in the way described for Janus.

Other questions involve the status of Face_mop frame feature actions. They are not definitional. It is not a necessary condition for an input to match all of them in order to acquire the associated emotion. The one frame might have contradictory feature actions: mouth slightly open and mouth wide (as in surprised). Rather the cluster of feature actions is a pool of "common possibilities" which states what "may" rather than what "must" be signalled for a specific emotion. Also potentially the frame feature actions may be added to by promotion of a feature action from a sub_mop or lose an action by demotion if support is withdrawn so that its credit rating approaches a low level. The frame contents clearly have not definitional status. Experience may confirm these or revise them. If not definitional, a case has been made in this thesis that they could be viewed as the more central features of prototypes.

Frame knowledge can be regarded as expert face knowledge: the knowledge indexed below it is based on the varying personal experience of many people. This would seem to be the same kind of knowledge as in the Frame but the latter was influenced by validated descriptions(Ekman & Friesen 1984). Although it can not claim the validated status of the Ekman & Friesen descriptions the Frame knowledge in Janus is meant to be as faithful to those descriptions as design considerations allow. The dynamic nature of such classification in the light of the strategies for promotion and demotion of frame feature actions introduces the consideration that Frame expert knowledge can lose this resemblance if not validated by user consensus. Conversely the latter is involved in creating Frame knowledge by a process of abstraction and generalisation. Riesbeck(1984) points to the fact that classification is the one area in which expert systems are novices to begin with in the sense that they start with one kind of knowledge and learn another. From another perspective, if this is really how humans organize knowledge, does this process of promotion and demotion reflect the relationship which exists between semantic memory and autobiographical memory; between the notion of concepts as prototypes and prototype fringes?. Semantic memory could contain abstracted facts from personal experience as well as the facts we learn about from other persons' experience.

**Summary:**

In this chapter the dynamic memory which classifies input face expressions has been described. It is based on a theory of human reminding which itself has not been explicitly applied before to the way in which humans interpret face expressions. The application, here, uses the principle of learning from experience. This is facilitated by forming organizational conceptual generalizations in memory which organize identical experiences to facilitate reminding. The implementation takes the form of a tree structure with typical expressions at the root and anomalies in the branches. How successful it is will be examined in the next chapter.

# CHAPTER 8:

# Validation

## 8.0 The Validation of Janus: Introduction:

Validation is done with the aim of crediting the system with attaining the accuracy for which it was built. Janus aims for a level of interpretation of facial expression close to that of the average human, in this case, college personnel.

The goal in this project was to demonstrate that it is possible to transduce an input digitized face representation into an emotional category. The validation was undertaken for the purpose of assessing to what extent the thesis hypothesis was supported by the results. As will be described below, some refinement was necessarily integrated with stages of development e.g. in refining the rule base.

The validation was not undertaken on randomly-selected face photographs on account of the considerable difficulties in obtaining, digitizing and measuring photographs in order to make the geometric input. In the selection of the faces that each person would judge, however, there was no conscious bias.

Both the parts and the whole system were tested. The two principle functional components are the *rule base* which converts geometric representations to syntactic(verbal) representations; and the *memory* which learns and retrieves and automatically re-organizes itself.

If the evaluation standard was to be the judgement of lay-experts, (a name which is roughly equivalent to college folk un-schooled in the domain other than in the wisdom of personal experience), a 'gold standard' was required to rate their performance. To control for bias, "blind" evaluation is appropriate when system performance is being compared with that of human experts. This involves the use of meta-judges to rate all performances unaware that a computer system is involved. There are statistical tests of agreement which can be used in these circumstances where many experts are compared on the same task.

**8.1 Planning the validation of Janus:**

Janus was designed as a research prototype system. The relevant stages in the development of an expert system referred to below are those defined by Shortliffe and Davis (1975).

The overall goal was to interpret a digitized full-face photograph in terms of the emotion portrayed given a geometric configuration of specified 2D facial landmarks. Such social understanding is attempted because computers and robots will require this ability.

This validation was made with the intention of proving that the system has reached an acceptable level of goal attainment and was at stage 4 of the Shortliffe and Davis schema. To be at stage 4, a system prototype should be running well and more formal evaluations are made to see if it can handle any random input. This stage is preceded by stage 3 where informal test cases are run through the system and the system is refined.

A statement of what the measure of system success would be is that an input of x/y co-ordinate values corresponding to specific face positions would produce an output interpretation with which average college personnel would agree. The more explicit statement of how many faces and how many persons would agree was left open, but 17 faces and 55 persons were used in the event. Likewise, the criterion of how much agreement (90% ?, 80% ?, 70% ?) would signify success was left open. The criterion was that the output interpretations would be acceptable to users.

### 8.1.1 The "Gold Standards":

Two "gold Standards" are used:

1: Selected pictures of the primary emotions from 'Pictures of Facial Affect' ('PFA', Ekman, 1976), and

2: Verbal descriptions of typical face actions for the primary emotions extracted from the text of "Unmasking The Face" by Paul Ekman & Wallace V. Friesen(1984), hereafter referred to as the "UTF" standard.

Where human lay-experts' judgements, on the same test cases as were interpreted by Janus, were sought in evaluating Janus, an attempt was made, if possible, to evaluate their own performance in relation to these gold standards.

### 8.1.2 Planning qualitative validation:

#### 8.1.2.1 Planning validation of Face_mop selection:

A form of predictive validation, but also sub-system validation, was planned to test that the input data would "self-select" the correct basic emotion ( i.e. the correct Face_mop contained in the 3rd tier node) by unlocking the right directed links in memory. Descriptions of typical face actions associated with validated common emotions were taken from UTF and the face actions associated with a specific emotion would be entered into Janus in random order and with systematic reduction in number. The system response (an emotion) in each case was compared with the UTF emotion. To defend against a charge of "designer bias", similar descriptions were extracted from UTF by a small group of persons individually and used as input.

#### 8.1.2.2 Planning validation of the learning function:

Inputting face actions with an interpretation sets up a memory store which is a rough equivalent of a database of test cases stored with their expert conclusions in a traditional expert system. A test of this learning capability was planned using UTF faces: one from each of the 6 common emotions: sad, happy, angry, disgusted, surprised and afraid. Experts judge the face actions and emotion present in each case and these are input into Janus in Learn mode. Since they are not limited to choice of the 6 emotions-labels named above they will introduce some idiosyncratic interpretation or face actions into memory and therefore bring about learning. Inputting an identical configuration of face actions should produce a response from the system verifying that learning has taken place.

### 8.1.3 Planning the quantitative validation:

#### 8.1.3.1 Planning quantitative validation of the rule base:

A sub-system validation of the system geometry-to-face-action rule base was planned using independent lay-experts as "gold standard". Face photographs were digitized and represented by 34 xy values as described in the text. Each face so represented was transduced into face actions (Janus face actions) by passage through the rule base. These were compared with the face actions judged to be present by human judges who were given a multiple choice questionnaire divided into the subsections: brows, eyes, nose, mouth, cheeks and jaw. Sub-section total nominal agreement/disagreement scores would be tested for significance by statistical tests. A non-significant result would imply that the system was no worse than the experts, on the number of face actions agreed.

#### 8.1.3.2 Planning quantitative validation of the overall system:

An evaluation was planned with the intention of proving that the system has reached an acceptable level of expertise and is at Stage 4 of the Shortliffe and Davis schema. The system's responses(attributed emotions) and human experts' responses(attributed emotions) to the same photographs would be evaluated by a panel of human experts who are unaware which are which. Their evaluations are tested for significance by statistical tests. Inability to demonstrate a significant difference between system and human evaluations will be taken as the measure of success.

## 8.2 Results:

### 8.2.1 Results of qualitative validation carried out:

Two aspects were considered in validating the system: the dynamic memory in interpret mode and the learning function.

#### 8.2.1.1 Results of qualitative validation of dynamic memory:

The dynamic memory component was tested to confirm that it returns the correct 'basic' emotion category. This was done using data in Ekman & Friesen(1975). To guard against bias several colleagues were asked to select the test data for this purpose. Results were almost always correct and the system was observed to degrade gracefully when systematic reduction of the input face actions returned the basic emotion (see appendix).

The learning capability was investigated to ensure that new input face actions and emotion labels were learned and correctly retrieved in subsequent interpretations. Two experts were asked to view six photographs, one for each basic category and supply lists of face actions together with their own interpretations. These were entered into Janus. Subsequent input of the same face actions did retrieve the correct interpretations. Thus the learning function appeared satisfactory.

Trained and untrained Janus' interpretations were rated by meta-experts and analysed both for the acceptability of interpretation and for whether the basic or learned interpretations were preferred. Discussion of these results is postponed until section 8.8.2.

### 8.2.2 Results of quantitative validation carried out:

#### 8.2.2.1 Results of quantitative validation of the rule base:

The aim here was to obtain a more precise estimate of the measure of agreement between the conclusions of Janus and those of human beings. The rule base was tested using four experts(A-D) (different from those used for the preliminary verification and refinement) and seventeen photographs. The questionnaire for eliciting face actions was divided into six sections corresponding to the six features: brows, eyes, nose, mouth, cheeks and jaw. The number of agreements and disagreements for each feature were computed for all possible pairs involving the experts A-D with and without Janus(J) and compared for each section using the chi-square test. The results were not significant at the 5% level. The number of agreements were in all cases well in excess of the disagreements. The full results of this all-pairs analysis is given in the appendix.

Typical examples exhibiting a range of results were extracted from this material. That for face actions associated with the eyes is shown in **Tables 8.1a & 8.1b** and for the mouth in **Tables 8.1c & 8.1d.** The scores for pairs involving Janus were compared with the others using the chi-square test. **Table 8.1a** compares the agreement(ag)/disagreement(da) scores for Janus-Expert(C) against Expert(B)-Expert(D). **Table 8.1b** makes the same comparison with Expert(A) taking the place of Janus. The values of chi-square obtained in both cases are well below the significant level for p = 0.05(3.84). In **Tables 8.1c & 8.1d,** Expert(C) takes the place of Janus in **Table 8.1d.** There is no sustained trend in the two examples. Janus is appreciably worse for eye actions.

Table 8.1: Validation of the rule base: agreement between Janus and experts tested by chi-square (see text for explanation):

a : eye actions

| EXP | J-C | B-D | TOTAL |
|---|---|---|---|
| ag | 88 | 98 | 186 |
| exp | 93 | 93 | |
| da | 31 | 21.0 | 52 |
| exp | 26 | 26 | |
| total | 119 | 119 | 238 |

$X^2 = 2.46$; d.f. = 1; $p \cong 0.12$.

b : eye actions

| EXP | A-C | B-D | TOTAL |
|---|---|---|---|
| ag | 94 | 98 | 192 |
| exp | 96 | 96 | |
| da | 25 | 21.0 | 46 |
| exp | 23 | 23 | |
| total | 119 | 119 | 238 |

$X^2 = 0.431$; d.f. = 1; $p \cong 0.512$.

c: mouth actions

| EXP | J-A | B-D | TOTAL |
|---|---|---|---|
| ag | 162 | 166 | 328 |
| exp | 164 | 164 | |
| da | 59 | 55 | 114 |
| exp | 57 | 57 | |
| total | 221 | 221 | 442 |

$X^2 = 0.189$; d.f. = 1; $p \cong 0.67$

d: mouth actions

| EXP | A-C | B-D | TOTAL |
|---|---|---|---|
| ag | 162 | 166 | 328 |
| exp | 164 | 164 | |
| da | 59 | 55 | 114 |
| exp | 57 | 57 | |
| total | 221 | 221 | 442 |

$X^2 = 0.189$; d.f. = 1; $p \cong 0.67$

In (b), expert A replaces Janus in (a);
In (d), expert A replaces Janus in (c).

### 8.2.2.1.1 Results of validation of the rule base using 'gold standard' face descriptions:

Although the rule base is thus in fair agreement with lay experts' judgements of face actions present, it would be gratifying to find that both agreed with a standard. It was decided therefore to test the rule base against a published authority. The procedure is described below.

**Test Photographs:** The pictures used for this purpose were taken from 'PFA': nos.84(happy), 91(disgusted), 90(surprise), 92(neutral), 41(neutral), 38(angry), and 37(afraid). No conscious bias dictated this choice except that the expressions seemed very well defined. The pictures featured two persons and therefore two neutral expressions were required.

**Experts:** Five clinical psychologist staff judged the expressions. These were different experts than those above.

**The task:** A separate answer book was completed for each picture by each judge. This was the same version of the forced choice questionnaire which had been refined as described in Chapter 6. It grouped face actions under six features: brows, eyes, nose, mouth, cheeks and jaw, and for many, diagrams and example photoprints of a person displaying the action were shown. For each feature, the subject was asked to tick the face actions present in the picture compared with a neutral expression. The same pictures were digitized, and represented as lists of face point co-ordinate values. These were converted into face actions by passing each list through the rule base, These face actions are referred to as those of Janus in the comparison with human judges.

**Method of analysis of results:** The **Williams** $\hat{I}_n$ **statistic (Williams 1976)** is used to compare the ratings with those of the human judges. This approach is specifically designed for the situation where one specific rater(Janus) is to be singled out for scrutiny. Details of the method are given in the appendix. For purposes of this discussion, we merely state that a statistic $\hat{I}_n$(where n is the number of reference raters, not including Janus) can be derived which is given by:

$$\hat{I}_n = \hat{P}_o/\hat{P}_n$$

$\hat{P}_o$ represents the overall agreement of the isolated rater with the reference raters while $\hat{P}_n$ represents the overall group agreement among raters 1 - n. A value of $\hat{I}_n$ close to 1 would indicate that a rating given by Janus would be in agreement with that of a randomly-selected reference expert about as frequently as the ratings given by two randomly-selected experts agree. It is possible(though tedious) to derive a confidence interval for the corresponding population $\hat{I}n$. This has not been done for this comparison as the small number of cases makes the large sample approximation somewhat dubious. The values of $\hat{I}_n$ quoted in Table 8.2 should therefore be taken as a measure of the plausibility rather than as being strictly validatory.

Most of the face actions show passable agreement. The exception is 'mouth slightly-open'(0.55). Singling this result out for closer scrutiny (**Table 8.2a**), and comparing these results with those of 'mouth wide'(**Table 8.2b**), it is clear that there is complete agreement overall that photograph 2 has neither a slightly-open nor wide-open mouth which is as it should be: face 2 has closed lips. The problem is that the human judges are more varied in their judgment of what is a wide-open as against a slightly-open mouth, sometimes indeed claiming both for

Table 8.2: Validation of the rule base - joint agreement of 5 human judges with Janus tested by Williams $\hat{I}_n$ measure of agreement.

| face action rule | $\hat{I}_5$ | face action rule | $\hat{I}_5$ |
|---|---|---|---|
| brows contracted | 0.83 | brows lowered | 1.1 |
| brows raised | 0.81 | brows in-raised | 1.07 |
| brows centre-raised | 1.00 | | |
| eyes u-lid-lowered | 1.19 | eyes u-lid-raised | 0.95 |
| eyes inlid-raised | 1.11 | eyes l-lid-raised | 0.91 |
| eyes wide | 1.00 | eyes narrowed | 1.05 |
| eyes raised | 0.83 | eyes shut | 1.07 |
| nose flared | 0.72 | nose screwed | 0.85 |
| mouth down | 1.04 | mouth up | 1.46 |
| mouth wide | 1.14 | mouth sl-open | 0.55 |
| mouth shut | 0.80 | mouth pulled | 0.9 |
| mouth compressed | 1.00 | mouth l-lip-lowered | 0.86 |
| mouth l-lip-raised | 0.97 | mouth u-lip-raised | 0.77 |
| mouth. u-lip-tensed | 0.94 | mouth l-lip-everted | 1.33 |
| mouth square | 1.19 | mouth bared | 1.00 |
| cheeks raised | 1.12 | cheeks dropped | 1.15 |
| cheeks n-l-vert | 1.21 | jaw dropped | 0.85 |

the same face. Janus, classifying by numerical thresholds, is more precise.

The problem is that comparing Janus' performance with human judges is not enough. It is necessary to compare the human judges performance with a validated standard. But if one can do this, it seems to render a comparison with human judges unnecessary. A better test of the rule base might then be a comparison of the results of the rule base on the 5 photos above with descriptions of very similar expressions of the same models, pictured and described in some detail in "Unmasking the Face" (hereafter referred to as UTF ) by the authors(P.F.A is a collection of transparencies without details of face actions. 'UTF' is a book which describes photographed expressions in detail).

**Table 8.2a: Judgements of Janus and 5 humans on the presence of 'slightly-open mouth' in five faces: F = false, T = true:**

| Photo | UTF | J | B | M | A | T | S |
|---|---|---|---|---|---|---|---|
| | | Janus & Judges | | | | | |
| photo1 | 84 | F | T | T | T | T | F |
| photo2 | 91 | F | F | F | F | F | F |
| photo3 | 90 | F | T | T | T | T | F |
| photo7 | 38 | F | T | F | T | F | F |
| photo8 | 37 | F | T | T | T | T | T |

**Table 8.2b: Judgements of Janus and 5 humans on the presence of 'wide-open mouth' in five faces: F = false, T = true:**

| Photo | UTF | J | B | M | A | T | S |
|---|---|---|---|---|---|---|---|
| | | Janus & Judges | | | | | |
| photo1 | 84 | F | F | F | F | T | T |
| photo2 | 91 | F | F | F | F | F | F |
| photo3 | 90 | T | F | F | F | F | T |
| photo7 | 38 | T | F | T | F | T | T |
| photo8 | 37 | F | F | F | F | T | F |

It is however, advantageous to retain the human comparison: it is in line with Janus' multi-user knowledge source. The correspondence between UTF description and Janus rule terminology is not straightforward and the following table of comparisons is approximate. The comparisons between Janus and the experts are detailed below in **Table 8.3.** To the extent that the descriptions in UTF are a "gold standard", Janus performed slightly better than the human judges as reflected in the total number of f-a achieved, but varied in performance slightly from face to face as was true of the human judges.

**Table 8.3: Comparison of Janus and 5 lay-experts using the writer's interpretation of a "gold standard"(UTF) description of face actions present in 5 'PFA' photographs.** (Only those face actions for which there was a question in the writer's questionnaire completed by lay-experts: B, M, A, T & S are compared).

1 = present, 0 = absent.

| **photo 1, UTF description:** | **J** | **B** | **M** | **A** | **T** | **S** |
|---|---|---|---|---|---|---|
| lip corners back & up | 1 | 1 | 1 | 0 | 1 | 1 |
| mouth sl.open / open | 1 | 1 | 1 | 1 | 1 | 1 |
| n-l grooves <> not Λ | 1 | 1 | 1 | 0 | 0 | 1 |
| cheeks raised | 1 | 1 | 1 | 1 | 1 | 1 |
| mouth widened back | 0 | 1 | 1 | 1 | 1 | 1 |
| **FACE TOTAL: 5** | 4 | 5 | 5 | 3 | 4 | 5 |

| Photo 2. UTF description: | J | B | M | A | T | S |
|---|---|---|---|---|---|---|
| mouth upper lip raised | 1 | 0 | 0 | 1 | 0 | 0 |
| nose wrinkled up | 1 | 1 | 1 | 1 | 1 | 1 |
| lower lip raised | 1 | 0 | 0 | 0 | 1 | 0 |
| lower lip protruding | 0 | 0 | 1 | 0 | 0 | 0 |
| cheeks raised | 1 | 1 | 1 | 1 | 1 | 1 |
| lower eyelid raised | 1 | 1 | 1 | 1 | 1 | 0 |
| FACE TOTAL: 6 | 5 | 3 | 4 | 4 | 4 | 2 |

| Photo 3. UTF description: | J | B | M | A | T | S |
|---|---|---|---|---|---|---|
| brows raised | 1 | 1 | 1 | 1 | 1 | 1 |
| eyes wide | 1 | 1 | 1 | 1 | 1 | 1 |
| upper eyelid raised | 1 | 0 | 1 | 1 | 1 | 1 |
| mouth open/slightly open | 1 | 1 | 1 | 1 | 1 | 1 |
| teeth showing | 1 | 1 | 1 | 1 | 1 | 1 |
| jaw dropped | 1 | 1 | 1 | 1 | 1 | 1 |
| FACE TOTAL: 6 | 6 | 5 | 6 | 6 | 6 | 6 |

| Photo 7 UTF description | J | B | M | A | T | S |
|---|---|---|---|---|---|---|
| brows drawn together | 1 | 0 | 1 | 1 | 1 | 1 |
| brows lowered | 1 | 1 | 1 | 1 | 1 | 0 |
| lower lids raised | 1 | 1 | 1 | 0 | 1 | 1 |
| narrowed eyes | 0 | 1 | 0 | 0 | 0 | 1 |
| mouth sl.open / open | 1 | 1 | 1 | 1 | 1 | 1 |
| teeth showing | 1 | 1 | 1 | 1 | 1 | 1 |
| jaw drop | 1 | 1 | 1 | 1 | 1 | 1 |
| FACE TOTAL: 7 | 6 | 6 | 6 | 5 | 6 | 6 |

| Photo 8 UTF description: | J | B | M | A | T | S |
|---|---|---|---|---|---|---|
| brows drawn up & together | 1 | 1 | 1 | 1 | 0 | 0 |
| eyes wide open | 1 | 1 | 1 | 1 | 1 | 0 |
| lower lid raised | 1 | 1 | 0 | 0 | 1 | 1 |
| mouth sl.open / open | 1 | 1 | 1 | 1 | 1 | 1 |
| upper lip tensed | 1 | 0 | 0 | 0 | 1 | 0 |
| FACE TOTAL: 5 | 5 | 4 | 3 | 3 | 4 | 2 |
| GRAND TOTAL: 29 | 26 | 23 | 24 | 21 | 24 | 21 |

### 8.2.2.2 Results of quantitative validation of the dynamic memory:

**8.2.2.2.1 Interpretation of basic emotion category:** The results of 8.2.2. showed that Janus was almost always correct in identifying the basic emotions using descriptions in Ekman & Friesen(1975). This is hardly surprising, as the associated face actions built into Janus were based on the same work. A more stringent test would be whether Janus returns the same basic emotions as human "experts" do when presented with an arbitrary set of face photographs.
Four experts(A-D) were presented with 17 photographs and asked to list the emotion category. The face actions obtained by passing the geometric descriptions of these same photographs through the rule base were input to Janus and the returned basic emotion was noted in each case. These are presented in **Table 8.4.**. Photograph 2 was used as the "neutral" for comparison and has been omitted from the analysis.

The **Kappa statistic**('K', Cohen 1960, 1968) was used to test for agreement among Janus and the four human experts. This is a measure of agreement which is applicable to categorical variables(e.g. the six categories of emotion) and where assignments are made by an arbitrary number of raters. Again detailed discussion of the method is postponed to the appendix and summarise the main points here. 'K is a measure of agreement among raters. A value of 1 implies complete agreement while 'K' = 0 suggests that there is no agreement other than what would be expected to occur by chance. The significance of 'K' can be tested (assuming large samples) by computing a Z-score. A high value of Z (small p) would imply that the observed level of agreement is not likely to occur by chance.

The following results were obtained when the basic emotions output by Janus were compared with those assigned by the human experts A-D: K = 0.467, Var(K) = 0.0013, Z = 13.13. This value of Z exceeds that for the 1%

Table 8.4: Interpretation of basic emotion category

| Photo | Janus | Expert A | Expert B | Expert C | Expert D | Photo |
|---|---|---|---|---|---|---|
| 1 | Happy | Sad | Happy | Disgusted | Sad | 1 |
| 3 | Afraid | Afraid | Surprised | Surprised | Surprised | 3 |
| 4 | Afraid | Sad | Surprised | Happy | Happy | 4 |
| 5 | Happy | Happy | Happy | Happy | Happy | 5 |
| 6 | Angry | Surprised | Surprised | Angry | Sad | 6 |
| 7 | Disgusted | Disgusted | Disgusted | Sad | Disgusted | 7 |
| 8 | Afraid | Angry | Angry | Angry | Disgusted | 8 |
| 9 | Afraid | Afraid | Afraid | Angry | Afraid | 9 |
| 10 | Disgusted | Sad | Disgusted | Happy | Disgusted | 10 |
| 11 | Surprised | Happy | Surprised | Happy | Happy | 11 |
| 12 | Happy | Happy | Happy | Happy | Happy | 12 |
| 13 | Sad | Sad | Sad | Sad | Sad | 13 |
| 14 | Afraid | Afraid | Surprised | Angry | Surprised | 14 |
| 15 | Surprised | Surprised | Surprised | Angry | Surprised | 15 |
| 16 | Sad | Sad | Sad | Sad | Sad | 16 |
| 17 | Angry | Angry | Angry | Angry | Angry | 17 |
| 18 | Disgusted | Disgusted | Disgusted | Surprised | Disgusted | 18 |

significance (Z = 2.33) and we conclude that the five raters, including Janus, exhibit significant agreement. Similar results are obtained when Janus is omitted from the analysis: 'K" = 0.450, Var (k) = 0.0023, Z = 9.463.

A better test of the ratings of Janus vis-a-vis the human experts may be obtained using the Williams $\hat{I}_n$ statistic introduced earlier in 8.5.1 This test was applied to the data of **Table 8.4**. Five sets of calculations were done, with Janus and the four human experts being selected in turn for scrutiny. Two additional cases were considered:

(a) The ratings of Janus were all replaced by the fixed emotion 'disgusted' and
(b) The ratings of Janus were all replaced by randomly-selected emotions.

These contrived situations, suggested by Max Bramer, were used to assess the sensitivity of the test.

In addition to calculating $\hat{I}_n$ for each case, it is important to be able to estimate the error limits for the In. Following Williams (1976), upper bounds were calculated for the population $\hat{I}n$, at the 5% significance level. The results are presented in **Table 8.4a**. A value $\hat{I}n$ close to 1 would suggest that the ratings of the test judge are as consistent with those of the reference judges as the ratings of the latter are mutually consistent. An upper bound of 1 or more would confirm this at the 5% confidence level. A value of $\hat{I}n$ significantly less than 1 (and a corresponding upper bound of less than 1 at the chosen confidence level) would imply that the ratings of the test judge are not consistent with those of the reference judges.

The first case in **Table 8.4a**(Janus as the test judge) shows that Janus is as consistent with the human experts as they are with each other. The following four cases (experts A-D selected as test judges) are roughly similar, though the results with expert C as the test judge are slightly anomalous. This is in marked contrast with the last two (contrived) cases, where Janus (with tailored ratings) is clearly inconsistent with the human experts.

Table 8.4a: $\hat{I}_4$ comparisons in test cases of Table 8.4.

| Case | $\hat{I}_4$ | upper bound on $\hat{I}_4$ at the 5% sign. level |
|---|---|---|
| Using Janus as the focused expert: | 1.04 | 1.21 |
| Using Expert A as the focused expert: | 1.09 | 1.27 |
| Using Expert B as the focused expert: | 1.19 | 1.37 |
| Using Expert C as the focused expert: | 0.63 | 0.90 |
| Using Expert D as the focused expert: | 1.14 | 1.34 |
| Contrived fixed response from Janus: | 0.27 | 0.52 |
| Contrived random responses from Janus: | 0.29 | 0.51 |

Details of the calculations are presented in the appendix.

Another approach used 'meta-judges' to rate the interpretations in **Table 8.4** in a 'blind' comparison. The meta-judges were asked to indicate whether each interpretation was (a) correct("What I would have said" = 'A' grade), (b) plausible("Not what I would have said but I can see its there" = 'B' grade) or (c) incorrect("I can not see it" = 'C' grade). They assigned these grades on the basis of their own appraisal of the same photographs. The number of 'A' or 'B' ranks assigned to each face by the combined 'meta-judges' were compared using the Friedman Anova test (Friedman 1937).

The Friedman test is appropriate where the same group of subjects are studied under different treatments and the outcomes are to be compared. In this case, the outcomes are the number of A/B grades given to the interpretations made by four human experts and Janus to the same set of photographs. We wish to find out whether or not the grades obtained by the five experts(including Janus) differ significantly. The calculation is in the appendix and the main points are summarized here.

The data is prepared as a two-way table of 5 columns(experts and Janus) and 17 rows, in which each row contains the rank positions across the row of the number of 'As' and 'Bs' accredited to each expert(Janus included) by the three meta-judges for the same face-photograph. A column tabulates these ranks over 17 face photographs.

The test statistic is $X_r$ which reflects whether the rank totals of the columns differ significantly(a condition which would argue against all the samples coming from the same population). $X_r$ is distributed approximately as the chi-square distribution with degrees of freedom equal to the number of columns minus 1. A value equal to, or larger than that in the chi-square table for a 5% level of significance at d.f.= 4 (9.49) implies that the hypothesis that all the samples came from the same population may be rejected.

The result of this test argues for accepting the five samples as coming from the same population at the 5% level($X_r$ = 2.5694, d.f. = 4, p ≅ 0.5). The calculation is shown in the appendix.

Demonstrating that no significant difference divides the meta-judges gives no idea of how accurate these were in their judgements. What is needed is a 'gold standard' consensus of judgment with which the accuracy of the meta-judges can be compared. Such a 'gold standard was found in the Pictures of Facial Affect (Ekman 1976), which sets out the percentage of consensus for each of the six primary emotions for each of the 'pictures'. Each picture is produced as a slide transparency and shows the full-face expression. Fifty slides were shown in succession to the three meta-judges who were asked to judge which of the emotions: happy, sad, angry, afraid, surprised and disgusted was depicted. These judgments were later compared with the appropriate classification given in the Pictures of Facial Affect brochure.

The three meta - judges scored, respectively, 40 (80%), 42 (84%), and 45 (90%) out of 50 (100%) in agreement with the 'gold standard'. On average then, by projection, the accuracy of their ratings of Janus and the four lay-experts might each be awry of the gold standard by 15% or by 1 judgement in six or seven. The absence of significant differences found in the Friedman test above gives no ground for believing that such errors are unevenly distributed among the judgements.

### 8.2.2.2.2 Validation of the Learning and Recall Functions:

The learning and recall functions of the dynamic memory were tackled next. A set of face photographs different from those used in previous tests were presented to a group of 30 people and a total of 50 event descriptions(face actions and non - standard emotion labels) were obtained from them. These were entered into Janus in learn mode. This resulted in an experienced('trained') memory. The question addressed is: how well are these learned labels rated by other judges?. Geometric descriptions of the 17 photographs used in the 'untrained' validation(but not in the training session) were converted to face actions by the rule base and input to the dynamic memory. The interpretations returned are given in **Table 8.5**. Fifty-five independent judges were asked to rate these as 'good', 'fair' or 'poor'. They were not told that the interpretations came from a computer. Each judge rated up to three photographs each both for the basic emotion category and the new 'learned' emotion labels yielding 149 basic and 336 learned ratings. The results are given in **Table 8.6**.

**Table 8.5: Interpretations output from trained Janus for garryphotos:**

| Photo | Basic Label | Alternative Learned Labels |
|---|---|---|
| 1 | Happy | Cheerful, Anticipating pleasure |
| 3 | Afraid | Puzzled, Fearful. Uncomprehending |
| 4 | Afraid | Puzzled, Fearful, Uncomprehending |
| 5 | Happy | Cheerful, Anticipating pleasure |
| 6 | Angry | Disliking, Glowering, Hostile |
| 7 | Disgusted | Disliking, Displeased |
| 8 | Afraid | Puzzled, Fearful, Uncomprehending |
| 9 | Afraid | Puzzled, Fearful, Uncomprehending |
| 10 | Disgusted | Disliking, Displeased |
| 11 | Surprised | Receptive to argument, interested |
| 12 | Happy | Cheerful, anticipating pleasure |
| 13 | sad | Depressed, Unhappy, Having distaste |
| 14 | Afraid | Puzzled, Fearful, Uncomprehending |
| 15 | Surprised | Receptive to argument, interested |
| 16 | Sad | Depressed, Unhappy, Having distaste |
| 17 | Angry | Disliking, Displeased |
| 18 | Disgusted | Disliking, Displeased |

**Table 8.6: Ratings on learned interpretations and basic emotion label:**

55 meta-judges each rated up to 3 sets of interpretations.
Each contained a basic emotion and its sub-grouped learned emotions which vary in number. In total 149 sets were judged.

| rating | Learned emotion. | | Basic emotion label | |
|---|---|---|---|---|
| | no | % | no | % |
| good | 103 | 30.66 | 78 | 52.35 |
| fair | 117 | 34.82 | 47 | 31.54 |
| poor | 116 | 34.52 | 24 | 16.11 |
| total | 336 | 100.0 | 149 | 100.0 |

There are several learned emotions but only one basic emotion for each face.

**8.3 Comment:**

The results show a clear preference for the basic emotion label (the FaceMOP emotion label), against the learned terms. This is in line with the view of emotions from a prototype perspective rather than treating them as classical concepts (Fehr & Russell, 1984) and the finding that there is in such prototypes a core of agreed meanings and a fuzzy perimeter. By analogy the Face_mop descriptions could communicate the prototype emotion and the anomalous f-a, the fuzzy margin.

It would be wrong to make such a preference for the basic emotion the reason for devaluing Janus for the following reasons:

**a)** The "learned" labels are valid human judgements of face expressions and even though people may disagree with the interpretations of others, there is no quarrel that peoples' experiences in life differ so that their constructions of reality differ.

**b)** The analysis is not a proper evaluation of the basic choice for the user between the basic or a learned interpretation. Both the basic and the learned emotion labels are output to the user as interpretations for an input face description in search of an interpretation. An analysis which does reflect this run-time choice is presented in **Table 8.7**. A comparison of ratings for 'basic' emotions categories and the 'learned' emotion labels *per face* shows that in about 47% of the cases, the ratings are similar. The basic emotion labels are more acceptable in about 29% of the cases, and the learned labels in about 25% of the cases. It is possible, however, that this may change as the system gains more experience.

This places the analysis within the contexts of single face photographs. On the basis of the above figures,

| Table 8.7: Ratings on learned interpretations and basic emotion labels | | |
|---|---|---|
| Relation between basic and learned emotion labels rated as good, fair or poor interpretations for individual faces: 55 human judges rated up to 3 selections each from Janus' interpretations of 17 faces: | | |
| | No. | % |
| The highest basic emotion grade (HBEG) | | |
| The highest learned emotion grade (HLEG) | | |
| HBEG = HLEG | 71.0 | 47.65 |
| HBEG > HLEG | 42.0 | 28.19 |
| HBEG < HLEG | 36.0 | 24.16 |
| total | 149.0 | 100.00 |

Table 8.8: Validation of Janus output interpretations:
Highest Grade assigned per face (Basic & Learned combined)
149 face - interpretations graded by lay - experts

| Grade | No. | % |
|---|---|---|
| Good | 105 | 70.5 |
| Fair | 35 | 23.5 |
| Poor | 9 | 6.0 |
| Good & Fair | 140 | 94.0 |

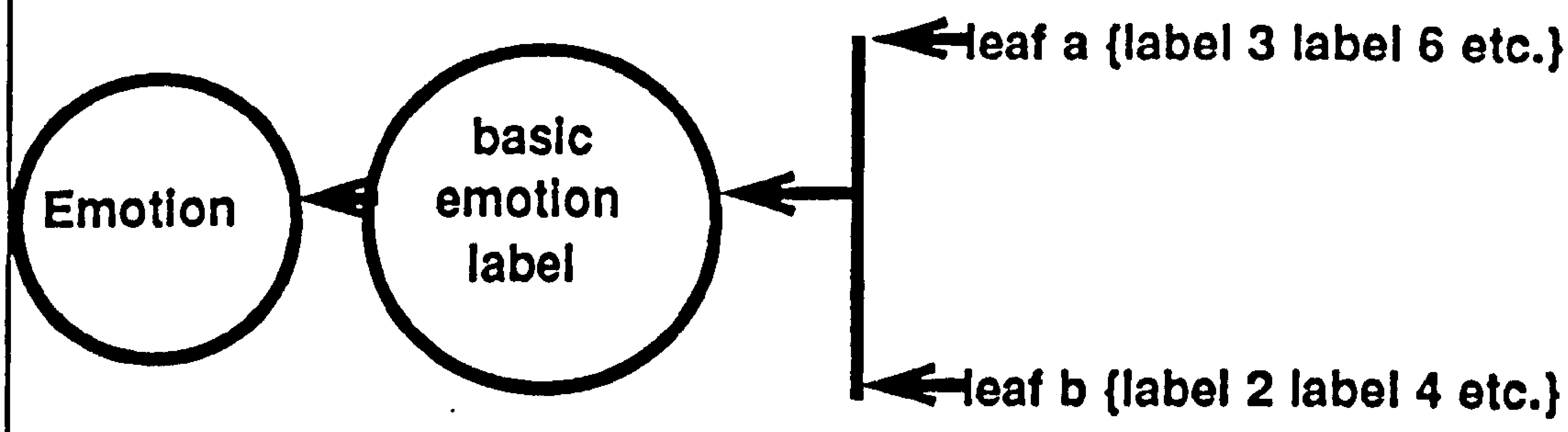
Janus is presently producing learned emotion interpretations rated as containing at least one which is graded equal to or greater than the corresponding basic emotion rating in 71.81% of cases'. There is the possibility always that a user's interpretation will be identical to the basic emotion. Such user interpretations were omitted from the 50 events used to train the memory. Had they been included the effect would be to augment cases where the HBEG = HLEG at the expense of cases where HBEG > HLEG. This would raise the above percentage.

c) A more practical validation of Janus' interpretative power is seen if the choice given to the user is addressed directly. Remembering that the Face_mop basic emotion is output as well as the leaf interpretations for each face submitted, validation may proceed with reference to the highest grade obtained per face, regardless of the category in which it falls. **Table** 8.8 lists the highest grade obtained for 149 faces from the above data. From **Table** 8.8 it is seen that 94% of the interpretations were approved in some measure.

As in the case of the meta - experts described earlier, it would be pertinent to test these new meta-experts for their judgment of facial expression in relation to the 'PFA' gold standard. Unfortunately there occurred a sizeable drop-out on re-calling these meta-experts to test their judgments in comparison with this gold standard, with 39 of 55(70.9%) being ultimately tested. Each of these meta-experts viewed and interpreted the same 24 selections from the Pictures of Facial Affect (Ekman & Friesen 1976) covering the six basic emotions. These Pictures were randomly selected within emotion categories. The task was to judge which basic emotion: happy, sad, angry, afraid, surprised or disgusted, each Picture depicted. The responses for each of the meta-experts were compared with the classification of these Pictures in the Brochure accompanying these published Pictures. The

Figure 8.0. Emotions From A Prototypical Viewpoint

a. Organization of emotion labels in Janus




b. Prototype view of the concept 'Emotion'

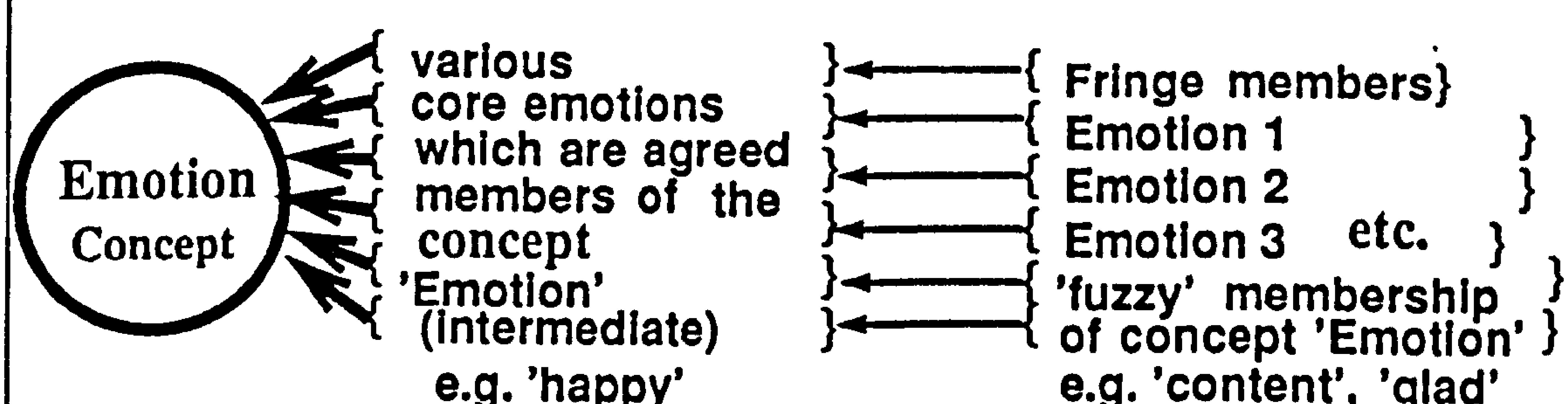



Table 8.9: 39 Meta-experts' Performance in Classifying Standard Affect Pictures

| | of possible total = 24 | % |
|---|---|---|
| Average per expert correct | 20.5 | 85.42 |
| Range | 13 - 24 | 54.0 - 100 |

results were totalled for each and the average and range of 39 meta-experts calculated. These are displayed in **Table 8.9.** The accuracy of the meta-experts appears satisfactory: on average each judging about 20 out of 24 pictures in accordance with the classification of the standard.

## 8.4 Discussion:

Summarizing the implications of the system validation, the overall choice for the user contains an acceptable interpretation in a high percentage of cases. There is good agreement with the judges if the Face_mop basic emotion and the learned leaf interpretations are taken together. In operation these are, of course, both ~~output~~ and therefore the overall interpretation is approved of in a good percentage of cases. The basic emotion was preferred over the learned emotions in roughly 1 in 4 cases and vice versa. No such preference existed in about half of the cases. It would seem then that there is an equal chance of either being acceptable. In 9 of 149 faces neither category of interpretation satisfied the judge, but 6 of these judges found interpretations for other faces satisfactory.

The question arises: are the average college persons' interpretations too idiosyncratic on which to base a knowledge system ?. The results of validation do not suggest so: seven times out of ten a learned interpretation was acceptable. This suggests that the use of many lay-experts would achieve a consensus.

Dissatisfaction with other person's interpretations could be inherent to the approach or due to a characteristic of human prototype-formation.

Seen from the perspective of the approach adopted for the knowledge acquisition for this system, the issue has to do with persons' interpretations and the reasons they give for them. In the system, the interpretations are just treated as tags to a face description so they cannot influence the learning and retrieval in any way. The same is not true for the face actions used to describe a face because these determine where in the tree the interpretation comes to rest, and hence under which basic emotion it is indexed. Validation of the leaf interpretations is addressing the question of whether the same anomalous face actions are producing nuances of meaning acceptable to most subjects. It is therefore testing also, the accuracy of the lay-experts' perceptions of the face actions signalling the emotion. Another factor possibly influencing satisfaction with leaf interpretations is the number of alternative interpretations offered. Repetition of labels from many trainers can lead to few alternatives from which to choose, for retrieval purposes This was the case in the present training.

On the other hand, dissatisfaction with the interpretations of others could be inherent to human concept-appraisal. Fehr & Russell,(1984) present evidence that there is an internal structuring in the concept of 'Emotion', with a central 'core' membership and a 'fuzzy' periphery the membership of which is more uncertain. This is to be expected from a prototypical viewpoint of the concept of 'Emotion'. It is, under this viewpoint, reasonable to expect more controversy in assessing the membership of fringe emotion labels to the emotion concept and this may be the reason for the disparity between acceptance of leaf emotion and FaceMOP emotion labels in **Table 8.6**. If agreement within the population varies as to whether a concept is or is not an 'emotion' then people holding the one view are unlikely to rate highly the opposing view.

If it is assumed that FaceMOP basic emotion labels have an intermediate 'core' membership of the concept 'Emotion' and leaf learned emotion labels are a 'fuzzy' fringe, then **Fig.8.0** conveys the correspondence which can be drawn between Janus's organization of emotion labels and the writers interpretation of the Fehr & Russell hypothesis.

**Summary:**

Validation of the rule base and the learning and retrieval functions of Memory was undertaken in comparison with human experts whose accuracy was monitored on a 'gold standard'.

The techniques used included blind study ratings and statistical tests of inter-rater agreement. In most of these tests the system was shown to reach good agreement with the human lay-experts whose accuracy on the 'gold standard' was considered satisfactory. It is evident that the basic emotion are about as equally acceptable as the learned emotion labels at least to the college raters.

In overall performance, however, combining the basic emotion and the learned emotion in each case, produced an acceptable interpretation in a high percentage of cases.

In the next chapter some theoretical and design issues are discussed before the thesis hypothesis is reviewed in the light of what has been achieved.

# Chapter 9:

# Discussion of results

## 9.0 Introduction:

The chapter presents an assessment of the achievements and limitations of the Janus system vis-a-vis the project aims.

## 9.1 An appraisal of Janus in terms of the aims of the thesis:

### 9.1.1 Achievements:

What has been achieved is a small research prototype which produces an emotional label or a choice of several of these from an input description of the x/y coordinates of 34 standard face locations. The validation suggests a passable agreement with the 'gold standards' adopted. A further procedure applies the prototype to the interpretation of the facial expressions of a human dyad and suggests causes for these in terms of cognitive dimensions of emotion. These are considered to be too general to be of practical use without contextual knowledge.

The significance of the work is that:

1. It is the first implementation, to the writer's knowledge, which uses its experience of face expressions to transform face geometry into concepts of emotion;

2. It opens a potential for a video camera equipped with a frame grabber, and an automatic measuring capability, to interpret full-face expressions directly.

3. It is an extension of memory-based methodology into a new order of knowledge - that of micro-events e.g. face actions. The method seems to deal with these very well, and the method could perhaps be extended to deal with geometric data in other image processing; for example to

form a learning and classifying back-end for vision systems which undertake scene analysis. As an example of how lower order of data can be so used, it is conceivable that face actions could be replaced by their equivalent geometry in the dynamic memory, thus removing the intermediate representation level. Then generalizations (sub_mops) would form when juxtaposition occurred between two identical face measures or within their tolerance interval.

The dynamic organization of memory employed here might find an application in an understanding facility for machine vision. The various geometric models, against which the end-products of machine vision scene analysis are matched in order to identify them, could be classified in a way so that any anomalous features could be indexed below these typical forms. The typical form of a chair, for instance, could classify the branch of the tree which contains armchairs, deckchairs, barber's chairs, swing chairs et cetera. Such a memory back-up would need to be taught these anomalies both as regards the anomalous features and the labels to be attached. The reorganization of the memory however would be automatic.

4. In the writer's opinion the most exciting prospect for a machine which interprets face signals of emotion is that it could represent a small step forward in hypothesizing the psychological preoccupations of the person viewed, especially the reasons for their emotions. Emotions are becoming of great interest to cognitive scientists, and Sloman and Croucher(1981a) have gone so far as to consider their possession by robots a necessity (see Chapter 2). Whether emotions are viewed as 'global messengers' and 'interrupts' or as signalling 'significant junctures in plans'( Oatley & Johnson-Laird 1985), they are a valuable clue to the motivations of the signaller. As such they have a promising part to play in man-machine interaction.

5. Janus is significant because it does explicitly indicate a representation for the way in which memories of face expressions may be classified in an intelligent system. The treatment has not included all the psycho-physical experimentation which would be required to validate a theory that this is the way in which humans interpret facial expressions, because the project was not addressed to that issue, but the field is lacking even the suggestion which Janus puts forward.

The way faces are represented in human memory is still unknown. Semantic net and connectionist representations predominate in the literature. Chapter 2 describes the current scene. Bower and his colleagues proposed a node and link representation for emotion but this does not specifically deal with facial expressions. The nature of the 'traces' which represent episodic memory have been represented by both semantic net and connectionist networks.

The connectionist model of McClelland & Rumelhart (1985) can extract the central tendency of the patterns of multiple instances and create prototypes from repeated training sessions with instances of a concept. Both multiple prototypes and instances can be stored and retrieved in these composite traces - indeed multiple prototypes can be produced in the same trace. Such a model is suggested by its builders as an alternative to the face recognition unit concept. It possibly is also an alternative to the Face_mop functions in Janus. In contradistinction to Face Recognition Units, Face_mops endorse the appearance of all persons with a basic emotion, not the appearance of the one familiar person. This reflects the universality of basic emotions in contrast to the uniqueness of the individual.

Although the formula for the adjustments made in connectionist models may be known, absence of discrete, easily-manipulable internal states(symbols) in these models could detract from their usefulness in explaining cognition which is held to require a symbolic level analysis - but such judgment may be premature (see Clark 1987 on this issue and on the relationship of connectionism and cognitive science).

### 9.1.2 Limitations:

1. If basic emotion labels achieve more consensus of agreement in the population than more idiosyncratic labels(as they have done in these validation studies), a point of view might be advanced with some justification that the classification of facial expressions could be achieved more efficiently with a 'look-up' table composed of the six basic emotions indexed by their typical expressions. This is quite reasonable and feasible. There needs only to be incorporated some arrangement for dealing with incomplete and anomalous face descriptions since any departure from the typical expressions of these would fail to access them. Unfortunately, such a system would never benefit from learning new expressions nor any emotion other than the basic six. It would lose out on the possibility of an inductive learning of all facial descriptions and interpretations since, in comparison, Janus has an infinite potential capacity to organize all the world's experience in this domain - if 'all the world' were users.

2. The need to compare a face action with a neutral or rest position in order to define a face action within a rule puts a severe constraint on the use of this methodology in a situation where the input would come from a video camera scanning unfamiliar faces. Its use would be limited to faces already measured in neutral pose. One possibility might be to average the face measures required over a system's experience, and use these measures as

default values in place of the neutral values of any one individual. Another possibility is the averaging of values of random time-elapsed frames of the target face as 'neutral' values.

3. There should also be concern at the arbitrariness of the measures on the face used to define face actions. In respect of this, the statistical method for determining the measure with the highest correlation with the face action, adopted by Pilowski and his group(1986, 1985) for a smile, is a positive advance. Such correlations would however need to be drawn over large samples of representative faces.

4. Accuracy of face measurements is important since at least one of the definitional changes in position measured on the face (that of a rise in the lower eyelid) is of the order of a minimum of one pixel at a scale of 75 pixels to the inch. It is doubtful whether automatic face measuring programs can be so accurate. They will need a tolerance margin of many more pixels than this and may be very prone to error as a result.

5. It might also be held that the context-free classification of facial expression is irrelevant to everyday life - that people do not find themselves normally in a situation where they are called upon to judge photographs or even real faces isolated from context. This is true, and one could say much the same for a physician diagnosing jaundice from a colour photograph of a face without any history or other findings to hand. The physician is able to assert jaundice but not the cause. Notwithstanding that there is a place for spot diagnosis in the physician's diagnostic acumen, (s)he is not often (save perhaps in tests of his powers), called upon to act on the spot diagnosis alone. Normally (s)he will seek the context of the patient to give meaning to the symptom or sign. It seems the same in regard to facial

expression. It is quite permissible to read John's expression as one of sadness. Prior knowledge of John's customary expression will dictate whether this sadness is unusual(which is one use of context) but if we are unaware of what might be the reason for John to look so, we will seek a context. We will ask him. If for some reason we cannot ask we might review what we know of John's recent activities and hypothesize a cause. These illustrations are to make the point that it is commonplace knowledge that we can read faces with only the context of what signals emotion in the face, but that for an understanding of the motivation behind the face, a wider context is needed. Janus reads face actions in the context of face actions and the physician reads jaundice in the face without having to know at that precise moment the wider context necessary for a deeper, causal understanding in order to say that it is jaundice.

6. The question of uncertainty poses a problem. Careful consideration was given to including uncertainty measures in the acquired face Knowledge. It was decided not to implement these for the following reasons: While there is no denying that attributing emotion is a fuzzy area of human judgement, the inference chain which the user implies by associating face actions with emotion is not altered by traversing the MOP. The emotion would be output from the leaf at the same certainty it was judged to have at input. The emotion term attributed to the same face by five judges could be the same or all different and if they came to cluster together in the same leaf node, retrieval from that leaf would offer the user the five emotion labels at the same measures of confidence as the five judges gave to each on input. In this research prototype where retrieval offers a choice, the retriever would choose anyway on the basis of his/her own judgement of the expression and it is difficult to believe they would be influenced in their choice by the certainty of other users.

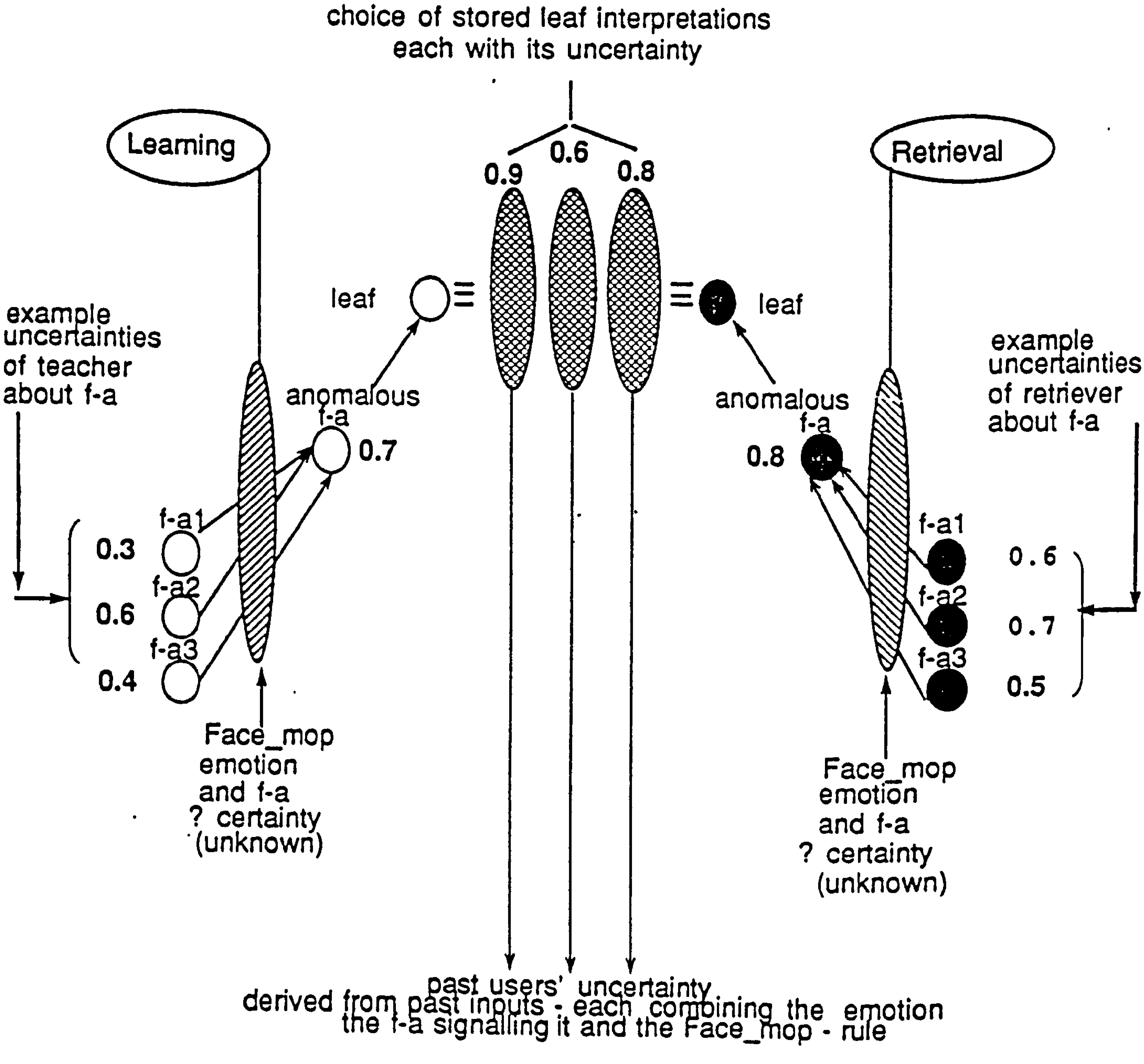
There would be problems also if certainty measures were restricted to individual face actions. If the judges attributed certainty measures to each face action they deemed to play a causal role then whatever combination rules applied, they would play no part in selecting a Face_mop to traverse since this, at present, is decided with reference to the greatest overlap between the input event and a pool of face actions within the Face_mop Frame. This is not to say that a threshold of certainty could not be imposed on each face-action in the pool.

If there were thresholds of certainty to overcome before a face action was fired such thresholds would have little theoretical basis in terms of the underlying Ekman & Friesen theory, since it is not known what value those authors would place on each face action in the face which signals emotion. Fuzzy limits would need also to apply to the matching of face action components which trigger the instantiation of fsub_mops and sub_mops.

The signalling of an emotion depends on interaction between face actions which, as it stands in Janus, must be there or not. The selection of the Face_mop is one level where face actions influence the interpreted emotion; the other level is that of the leaf node of anomalous face action branches. To impose thresholds for these face actions would be equivalent to forcing one user's uncertainty on another.

Arriving at a leaf in search of an interpretation for the particular face actions input, with only, say, a 0.7 certainty that the correct Face_mop has been traversed (because of doubt whether the face actions described in the face were really there), entails that the leaf interpretations (which would have degrees of certainty attributed by prior users) would be offered at a certainty level dictated by a rule combining 0.7 with each. On

**Figure 9.1: Problems for uncertainty in Retrieval**




f-a = face action

all uncertainties are hypothetical

retrieval by another user this uncertainty will be combined with the latter's uncertainty about what face actions are present since a face-action description has to be input in order to effect retrieval. This seems intuitively wrong and it is for these reasons that uncertainty has not been implemented in Janus. The writer has doubt whether it would be justifiable. **Fig.9.1** shows the teach/retrieve cycle. It shows, on the left hand side, the various uncertainties (hypothetical only) which would need to be combined to produce leaf node uncertainties if the teachers' uncertainty were incorporated. On the right hand side are those which would express the retriever's certainty about the face description that is input in search of an interpretation.

A consistent treatment of uncertainty would require that the latter should be combined with the former to represent the uncertainty in the offered interpretations but this does not appear intuitively right.

7. The generalizability of the rule base has been established only in relation to sets of three different faces. Measurement of points on digitized face photographs is by no means exact. Although some research algorithms for automatic measuring of face dimensions report good results in measuring some feature distances working on line images produced by the Laplacian operator, the published images would seem to lack the definition required to measure all the face points Janus requires. This is because the feature movement distances required for discriminative emotion detection are both different and much finer than the parameters which are used to narrow the search for a face in a face database. A related student final year project in this college used many different operators without producing a satisfactory line image of the required definition. Doubtlessly, such problems can be overcome with state-of-the-art technology, and automatic measuring of useful parameters for

expression will improve so that video camera frames can be measured. Even so there is still the problem of profile and less than full-face images which this rule base can not handle.

8. Is the representation adequate? The domain is the domain of face expressions - the vehicles of the signalled emotions. Face expression here has been constrained to face actions involving the brows, eyes, nose, mouth, and jaw. While this constraint reduces the information given by some people for their reason for attributing the interpretation they make, it appears to have sufficient power of representation so that rules for converting face geometry into face actions predict fairly successfully both the face actions which people judge to be present in new faces and the emotions judged to be present as a result. The validation studies confirm this.

9. It is not clear that Janus models human face expression processing because it is not known how humans process expressions. One might argue that one does not intuitively respond to an expression by scanning a mental list of emotion labels from memory - a response on which some decision is then made. That people are able to do so, if they put their mind to it, is different from accrediting it as the method of brief impression formation.

It has been beyond the scope of this project to design and run the psychological experiments which would add support to the theory that one's interpretation of others' expressions are ordinarily derived in the manner described here. In particular, the device of tagging an expression with an emotion is clearly a gross simplification of the way in which an emotional context is represented. Principally the idea of the input 'reminding' the system of previous identical traces without context other than that of other face components was implemented because of its learning, organizational and retrieving potential. It allows the system to benefit from experience.

With regard to the organization of its memory, also, it is not clear whether Janus accurately models human memory.

The classification of an anomalous expression under a particular Face_mop expression implies both similarity with and some difference to the latter. This is plausible enough. That expression-similarity implies similarity of *meaning* between the leaf interpretation produced by this anomalous input expression and a basic emotion is, perhaps, less intuitive. As noted in the previous chapter this is in no way endorsed by all, in that 25% roughly of leaf emotions are judged less apt with regard to the corresponding basic emotion. It may be that the more true relationship is that both the basic and leaf emotions are core and fuzzy fringe members of a subcategory of the concept: 'emotion'. Human perception, of course, involves pre-processing of brightness intensities, distance, depth and colour cues etc. before a face is recognised as a face much less as an expression. Memory functions alone are very incomplete theories of perception.

10. A further limitation lies in the use of still photographs and of posed expressions for knowledge acquisition and validation. Their relation to real-life spontaneous emotional display is not clear. The dependence of the Janus rule base on full-face viewing is unrealistic if natural camera scanning were to supply the image.

## 9.2 Conclusion:

The thesis examined whether it is possible to quantify face expressions and produce machine-made interpretations of emotional states from such quantities alone, such that these interpretations are rated no worse than those provided from humans beings judging the same facial expressions. While the work done contains grounds for optimism in this direction, caution is needed

. in view of the limitations of the method discussed above.

In the next chapter some extensions to the thesis are proposed.

# Chapter 10:

# Extensions

## 10.0 Two emphases for further research:

Further experimentation might be useful in the following areas:

1. Recognition of the face emotion ,
2. Inferring cognitive associations of the face emotion.

These will be discussed in turn.

### 10.1 Recognition of the face emotion:

Automatic measurement of face points was not addressed in Janus. It would be a natural extension of this work to apply state of the art techniques to the problem of extracting required face points from greyscale images from photographs and video frames. The work done so far has been focused on the extraction of parameters useful in classifying faces for identification purposes(Wong, 1989; Craw et al., 1987; Buhr 1986; Bromley, 1977; Kaya & Kobayashi, 1972; Sakai & al.(1972); Bisson, 1965), the detection of blinks(Turk & Pentland, 1989) and head movements(Sheehy, 1989), and automatic lip-reading(Petajan, 1985). The combination of acoustic and visual recognition in this last study has produced superior recognition accuracy than acoustic recognition alone for digits, letters and words. The study also provides useful ideas to facilitate its real-time implementation such as hardware edge parameterizing and parallel frame processing - pertinent for computer recognition of user body language, including that of emotion.

Further work could be directed towards the following areas:

It might be possible to reduce the number of face points that are used to represent the primary emotions by a statistical comparison of the parameters along the lines of Pilowski et al., 1985.

The rule base may be dispensed with by coding Memory in terms of geometric parameters only.

Even Memory could be dispensed with in an alternative system option; replacing it with a look-up table of emotions equated to geometric face parameters. This might be suitable for some applications say simple human-computer discourse.

Eigenface-classification (Turk & Pentland, 1989) may provide an alternative method for classifying expressions. An eigenface extracts the information in a face by capturing the variation in a collection of face images, independently of any judgement of features. It does this in the same way that a set of eigenvectors represents the variance in a covariance matrix. An individual face can be represented as a linear combination of the eigenfaces.

The work done by Turk & Pentland is on identification of persons and involves the formation of eigenface classes based on a set of images of several people. Each person shows some variation of expression within his/her images. It remains to be researched whether this technique could be adapted to classify emotional expressions. It could be arranged so that the set is of one person showing variants of emotional expressions. Eigenvectors with the highest associated eigenvalues would be extracted and combined to form the eigenfaces (representing the combinations of the training set of images). A target face showing a specific emotional expression would be converted to its eigenface components and its pattern vector (representing the contribution of each eigenface in representing the input face image) is used to calculate the nearest face emotion class.

If an alternative to Janus is required - one which uses mathematics only and obviates the need for both rule base and memory and which delivers a 'nearest' match to standard classes of emotion, one can use the concept of 'Expression space'. Instead of representing an expression as, say, 34 points in two dimensions, one could represent it as one point in a 68-dimensional space. Then the typical face expressions of the six primary emotions will have static positions in this space represented as points. A target expression will be also a point at a fixed location and its position relative to the six will be analogous to its closeness to them in expression. An ordering of these measures would correspond to the ordering of the target emotion in terms of the primary six. It would be interesting to compare the results of this technique with that of Janus for the same input face expression.

Distributed memory models such as that of McClelland & Rumelhart, 1985 and Kohonen, 1981, give very thought-provoking analogies to prototype formation and face recognition respectively. There would seem to be every reason to believe that the former model has the ability to form face expression prototypes. This, however, has yet to be claimed.

Along the same lines, it would be interesting to allow a prototype Janus, but without the six primary emotions in the nodes 7-12, to form its own generalizations over time. Serial input of face expressions would create or follow existing paths making matches and MOPs as they go. Frequently recurring face actions would, in time, tend to build up sets as they were 'promoted'. These may be found to resemble those of the six primary emotions.

Further work on Janus as it stands would be well directed to the possibility of introducing an acceptable form of certainty rating and also a parallel processing approach so that the one input face expression is matched to the six Face_mop Frames at the same time and traverses those MOPs with which a criterion match obtains.

This chapter is concluded with a description of an implemented procedure which uses the basic emotion output by Janus to address general cognitive dimensions which are believed to be associated with specific emotions.

## 10.2 Inferring cognitive associations of emotions..

One could describe face expressions as 'the royal road to motives'. Human expressions should be of some communicatory value to computers and robots. To rely on expression alone to communicate junctures of plans from person to person is not the usual way of going about this. Speech is more commonly employed. Humans make use of multiple channels of communication. Body language is one channel only. 'Hal', the knowing computer in the film '2001' could also lip-read and hear. Patejan(1985) combines acoustic recognition with automatic lip-reading and found that the latter always improves the recognition rate of digits, letters and words compared to acoustic recognition alone.

In Chapter 2 the work of Sloman & Croucher and of Oatley & Johnson-Laird emphasized the crucial role which emotions may play in intelligent systems. In this Chapter, is described a procedure, DYAD, which uses the emotion label produced by a Janus run to extend the inference into the domain of the motives associated with the basic emotions. This procedure has been implemented as a separate procedure which 'calls' Janus to infer the basic emotion from separate input of two faces described in terms of face actions. The procedure sends the output basic emotions in sequence to a Flavour object which infers the cognitive dimensions of each basic emotion in general terms according to the writer's understanding of the specific pattern these dimensions take to each emotion as pioneered by Roseman(1982).

### 10.2.1 Roseman's cognitive dimensions of discrete emotions:

The five cognitive dimensions identified by Roseman and found to differentiate 13 discrete emotions are:

1. Motivational State,
2. Situational State,
3. Probability,
4. Legitimacy, and
5. Agency.

These are described briefly below, drawing on Roseman's explanation:

**Motivational State:** whether the priority goal is a **desirable** state to attain or an **undesirable** state to avoid in the present situation, (options: **d** or **u**).
**Situational State:** whether the motivational state is **realised** or **not**, (options: **p** or **a**)
**Probability:** whether the realisation of the goal is **certain** or **uncertain**, (options: **c** or **unc**).
**Legitimacy: whether or not** the person feels s/he **deserves** the outcome, (options **des** or **undes**).
**Agency:** whether the outcome came about by the agency of **self, another** or **neither**, (options: **s**, **o** or **l** ).

The basic emotions are then defined in terms of the differentiating options of these five dimensions thus:

**Happy is defined as:**

[
[motivational state: d ]
[[situational state p ]
[probability c]
[legitimacy d or u]
[agency s, o, or l]
]

**or**

[
[motivational state: u]
[situational state a]
[probability c]
[legitimacy d or u]
[agency s, o, or l]
]

On this analysis we could say that a person who by his/her expression is judged to be happy has some desired goal realised for certain, whether deserved or not and irrespective of whose agency brought it about - with the rider that if it came about through his/her own agency he or she might as well feel pride, if by another: gratitude and liking, or if by impersonal means, perhaps wondrous of his or her luck,

OR

The person is happy because an outcome s/he wished to avoid will not materialize for certain. It is not clear whether this result was thought to be deserved. If the outcome results from self effort s/he may also feel pride; if due to others, some affection may be felt towards them. Relief or guilt may also colour the picture especially if the outcome was undeserved.

The cognitive dimensions of the other basic emotions can be expressed in such general terms as these also. Clearly without information about the value of these dimensions, such inferences can only be given in such generalities, but if it were possible to add some contextual information about the person viewed it would be possible to venture more informed guesses about what was producing the emotion and within a MMI context, ask questions of the user to implement a user model.

In order to illustrate the link with the output from

Janus, on the following page an abridged 'print out' from the run of a procedure named "DYAD" is shown. This file infers the values of Roseman's cognitive dimensions (as interpreted by the writer) from primary emotions output by JANUS in response to an input of two consecutive syntactic face descriptions in search of interpretation. A dyad is an interacting couple. By specifying a dyad it is assumed that the cause of each expression is in the other person.

### 10.3 The 'DYAD' procedure:

Dyad inputs a face description in search of interpretation for each of the dyad in turn into a trained Janus. Janus interprets the facial expression of, first, John and then Mary - a couple who are interacting with each other. DYAD theorizes on the cognitive dimensions of each depicted emotion along the lines of the Roseman (1982) thesis. JANUS attributes these cognitive dimension-values to the respective member of the dyad.
The input takes the form:

**dyad(John, Mary,[mouth raised eyes l-lid-raised cheeks raised mouth bared],[mouth compressed brows low eyes l-lid-raised eyes l-lid-tensed mouth u-lip-raised]),**

First the expression of John is input into JANUS as a list of face actions in search of an interpretation: Only the basic-emotion interpretation is printed in the following run. The input retrieved learned emotions also which are not shown for clarity.

*****************TARGET IDENTIFIED*********************

A SIMILAR TARGET FACE-DESCRIPTION HAS BEEN IDENTIFIED LABELLED [ev1] indexed under the general conceptual category mop_happy_gen_Face with THE GENERAL INTERPRETATION :

"happy"

Now the expression of Mary is input as a list of face actions in search of an interpretation: the interpretation is output:

****************TARGET IDENTIFIED********************

A SIMILAR TARGET FACE-DESCRIPTION HAS BEEN IDENTIFIED LABELLED [ev2] indexed under the general conceptual category mop_disgusted_gen_Face with THE GENERAL INTERPRETATION :

"disgusted"

So John is showing a happy expression and Mary is showing a disgusted expression. The attributions that follow are a direct consequence of these emotions. They are produced by DYAD. Had the depicted emotions been otherwise, these attributions would be different. The following inferences (after Roseman 1982) can be drawn from the discrete emotions displayed by John and by Mary:

John considers the presence, information or attention of Mary to fit in with his plans and Mary is present. Hence Mary is responsible for the happiness of John who might feel he has deserved this success or not.
Alternatively, John being happy may not be connected with Mary, herself, but due to his belief that she can help him achieve a desired goal.
Mary is confronted by an undeserved situation which she did not desire and tries to avoid but cannot and, as a result, is disgusted. The cause is connected with John - either his presence, appearance, behaviour or something he is saying.

There is a conflict in the motivation of the dyad: John has attained some step in a current goal that centres on Mary and this is something he wanted but Mary has clearly indicated that there is something about it she finds distasteful and would wish to avoid.

Similar scenarios would be obtained for any combination of emotional expressions in the dyad, One could focus the interpersonal causes also by adding interactive context, e.g. by seeking causes in connection with something held by one of the dyad, e.g. a gun, or the computer process being edited et cetera. One could constrain the interpretation so that its focus is on an external context (Tiberghien 1986, Baddeley 1982) by including a context which by its mere presence and nature(e.g. a tiger) would interrupt any prior motives. It is possible to include symbols relating to context within the input to JANUS - in fact a "sit" slot already exists which can take a value e.g. "tiger". It is another matter to represent the relevance and salience of such to humans in the scene. It is another matter again to identify the tiger from a grabbed image but this is a problem for machine vision.

Even though a visual scanner front end might conceivably get so far in trying to interpret emotional expressions there are still enormous problems to be overcome in determining what the motives of John and Mary are as they affect this situation. Unless the context can be constrained to a scenario where the goals can be made explicit and the person's actions within a restricted set all construed as motivated by desire to attain that goal, emotional expression alone will not be sufficient to specify the motives. Human-computer interaction is a context which might be so constrained.

Context is the word given to the total situation, internal and external which make up an event. The search for user motives in the user's interaction with his computer must be constrained first and foremost to the programming situation. The computer must develop a model of the user's top goal, and its decomposition into sub-goals and plans. Once the computer has modelled the goal tree, the user's expression might reveal the junctures of plans, and prompt the computer to query the user to confirm if there is a problem. There is opportunity for the user's idiosyncratic expressions to be associated in the computer's face memory with particular goal junctures and so producing a better incremental understanding of the user's expressions. This is possible if face memory, like Janus, is able to reorganize itself automatically. This knowledge of the user can then become part of the context for that user in any future discourse. This appears to the writer to be an important extension of the present work.

# Bibliography


Abdi, H.(1986),
"A generalized approach for connectionist auto -associatve memories: Interpretation, implication and illustration for face processing", in "Artificial Intelligence and Cognitive Sciences", Demongeot, J.et al. (eds.,1988),Manchester University Press.

Aleksander, I., Thomas, W.V., & Bowden, P.A.(1984),
"WISARD - a radical step forward in image recognition",
Sensor Rev., 4, 120-4.

Aleksander, I., & Burnett, P.(1983),
"Reinventing Man",
Kogan Page, london.

Aleksander, I., & Wilson, J.D.,
"Prospects for Adaptive Window Architectures", available from Dept. of Computing, Imperial College, 180 Queensgate, London SW7 2BZ

Anderson, J.R.(1989),
"Features of advisory and expert systems" : chapter 8 in "Intelligent Systems in a Human Context (Murray, L. & Richardson, J., eds.), Oxford University Press.

Anderson, J.R., & Bower, G.(1972),
"A propositional theory of recognition memory"
Memory & Cognition, 2,3, pp.406-412

Anderson, J.R., & Bower, G.(1972),
"Recognition and Retrieval Processes in Free Recall", Psychological Review, 79, 2, 97-129.

Anderson, J.(1976),
"Language, Memory And Thought"
Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Arnold, M.B.(1960)
"Emotion and Personality", Vol.1. Psychological aspects
N.Y: Columbia University Press.
2ndary reference from Izard, 1971.

Arnold, M.B.(1960)
"Emotion and Personality", Vol. 11. Neurological and physiological aspects.
N.Y: Columbia University Press,
2ndary reference from Izard, 1971.

Averill, J.R., Opton, E.M.,Jr. & Lazarus, R.S.(1969)
"Cross-cultural studies of psychophysiological responses during stress and emotion". Int. J. Psychol. 4, 83-102, 2ndary reference from Izard, 1971.

Baddeley, A.(1982),
"Domains of recollection",
Psychological Review, 89, 708-729.

Baddeley, A.(1979),
"Applied Cognitive and Cognitive Applied Psychology:
The case of Face Recognition", in "Perspectives On Memory Research: Essays in Honor of Uppsala University's 500th Anniversary" ed. Nilsson,L. Lawrence Erlbaum Assoc., Hillsdale, N.J.

Bain, B.(1984),
"Assignment of responsibility in Ethical Judgements", Paper presented at First Annual Workshop on Theoretical Issues in Conceptual Information Processing, Atlanta,GA, March 1984.

Bartlett, F.C.(1932),
"Remembering: A study in Experimental and Social Psychology", The University Press, Cambridge.

Bisson, C.L.(1965a),
"Measurement by computer of the distances on and about the eyes" (Report No. PRI-18). Panoramic Research, Inc., Palo Alto, California. 2ndary reference from

Laughery et al.(1981).

Bisson, C.L.(1965b),
"Location of some facial features by computer", (Report No. PRI-20). Panoramic Research, Inc., Palo Alto, California.2ndary reference from Laughery et al.(1981).

Bower, G.H.,(1981),
"Mood and Memory", American Psychologist, 36, 129-148.

Bower, G., Gilligan, S., Monteiro, K.(1981),
"Selectivity of Learning Caused by Affective States",
J. Exper. Psychol. (General), 110, 4, 451-473

Bower, G., & Karlin, M.(1974),
"Depth of processing pictures of faces and recognition memory", J. Exper. Psychol., 103, 4, 751-57.

Brennan, S.E.(1985),
"Caricature Generator:
The Dynamic Exaggeration of Faces by Computer"
Leonardo, 18, 3, 170-178.

Bromley, L.K.(1977),
"Computer-aided processing techniques for usage in real-time image evaluation", Master's Thesis, University of Houston. 2ndary reference from Laughery et al.(1981).

Bruce, V. & Young, A.(1986),
"Understanding face recognition", British J. Psychol. 77, 305-327.

Bruce, V.(1988),
"Recognizing Faces", Lawrence Erlbaum Associates, London.

Bruce, V.(1979),
"Searching for politicians: An information-processing approach to face recognition.", Quarterly Journal of Experimental Psychology, **31**, 373-395.

Bruce, V(1977),
"Searching for politicians: An information-processing approach to face recognition.", Paper presented to the Experimental Psychology Society, Sheffield, March 23 1977.

Buhr, R.(1986),
"Front Face Analyses & Classification (pictures)", NTZ ARCH(Germany), 8, pp.245-256.

Bull, N.(1951)
"The attitude theory of emotion" N.Y: Nervous and Mental Disease Monographs, No. 81. 2ndary reference from Izard(1971).

Clancey, W.J.(1983),
"The Epistemology of a Rule-Based Expert System - a Framework for explanation", Artificial Intelligence, 20, 215-251

Clancey, W.J.(1985),
"Heuristic Classification", Stanford Knowledge Systems Laboratory Report No. KSL 85 - 5.

Clark, A.(1987),
"Connectionism and Cognitive Science", pp. 3-15 in "Advances in Artificial Intelligence", eds. Hallam, J. & Mellish, C., John Wiley & Sons, Chichester, England.

Cohen, J.(1960),
"A coefficient of agreement for nominal scales", Educational and Psychological Measurement, 20, 37-46.

Cohen, J.(1968),
"Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit", Psychological Bulletin, 70, 4, 213-220.

Conway, M.A., & Bekerian, D. A.(1987a),
"Situational Knowledge and Emotions", Cognition and Emotion, 1 (2) 145-191.

Conway, M.A., & Bekerian,D.A.(1987b),

"Organization in autobiographical memory", Memory & Cognition, 15 (2), 119-132

Courtois, M.R. & Mueller, J.H.(1979), "Processing multiple physical features in facial recognition", Bull. Psychonom. Soc. 14 (2)., 74-76

Craik, F.I.M.(1973) "A levels of analysis view of memory" in "Communication and affect; Language and Thought" (P.Pliner, L.Krames & T.Alloway,Eds., Academic Press, N.Y..

Craik, F.I.M. & Lockhart, R.S.(1972), "Levels of processing: A framework for memory research", J. Verb. Learning and Verb.Behav., 11, 671-684.

Craw, I., Ellis, H. & Lishman, J.R.(1987), "Automatic extraction of face features", Pattern Recognition Letters, Netherlands, 5, 2,- 183-187.

Darwin, C.(1872), "The Expression of the Emotions in Man and Animals", Murray, London.

Davies, G.(1986), "The recall and reconstruction of faces Implications for theory and practice", in "Aspects of Face Processing"(Ellis et al eds.), Martinus Nijhoff, Lancaster.

Davies, G.M., Ellis, H.D. & Shepherd, J.W.(1977). "Cue saliency in faces as assessed by the Photofit technique", Perception, 6, 263-269.

Dittmann, A.T.(1972), "Interpersonal messages of emotion", Springer, N.Y. 2ndary reference from Salzen, 1981.

Duchenne, B. (1862), "Mecanisme de la physionomie humaine, ou analyse electro-physiologique de l'expression des passions", Bailliere, Paris.2ndary reference

(Ekman & friesen, 1984.

Dyer, M.G.(1983),
"In-Depth Understanding - A Computer Model of Integrated Processing for Narrative Comprehension", M.I.T. Press, Cambridge Mass.,

Eibl-Eibesfeldt, I.(1972),
"Similarities and differences between cultures in expressive movements. In "Non-verbal Communication", R.A. Hinde(Ed), C.U.P., Cambridge.2ndary reference from Saltzen, 1981.

Eibl-Eibesfeldt, I.(1970),
"Ethology, the Biology of Behaviour", New York, Holt, Rhinehart & Winston.(2ndary ref.Ekman & Friesen, 1984).

Ekman, P.(1982),.
"Methods for measuring facial action", in "Handbook of methods in nonverbal behavior research", K.R.Scherer & P.Ekman(eds.), Cambridge University Press.

Ekman, P.(1977),
"Emotion in the Human Face"(2nd. edition), Cambridge University Press.

Ekman, P., Oster, H. 1979,
"Facial expressions of emotion", Ann. Rev. Psychol, 30: 527-554.

Ekman, P., Friesen, W.(1984),
"Unmasking the Face (reprint). A guide to recognizing emotions from facial cues", Prentice Hall, Englewood Cliffs, New Jersey.

Ekman, P. & Friesen, W.(1978),
"The Facial Action Coding System: A technique for the measurement of facial movement". Palo Alto, Calif.: Consulting Psychologists Press.

Ekman, P. & Friesen, W.(1976),
"Pictures of facial affect", Consulting Psychologists Press, Inc., 577 College Ave., Palo Alto, CA., 94306.

Ekman, P. & Friesen, W.(1975), "Unmasking the Face. A guide to recognizing emotions from facial cues", Prentice Hall, Englewood Cliffs, New Jersey.

Ekman, P. & Friesen, W.(1976a), "Measuring facial movement", Journal of Psychology and Nonverbal Behavior, 1, 56-57

Ekman, P., Friesen, W. & Tomkins, S.1971, "Facial Affect Scoring Technique; A First Validity Study", Semiotica, 3, 37-58.

Ekman, P., Friesen, W. & Ellsworth, P.(1972), "Emotion in the human face: Guidelines for research and an integration of findings.", New York: Pergamon Press, 1972. Reprinted 1982, Cambridge University Press.

Ekman, P. & Friesen, W.(1982), "Measuring facial movement with the Facial Action Coding System" in 'Emotion in the human face', (second edition: ed. Ekman, P.,) Cambridge University Press, pp 178-211

Ekman, P. & Friesen, W.(1969), "The repertoire of non-verbal behavior", Semiotica, 1, 49-98.

Ekman, P., Sorenson, E. & Friesen, W.(1969), "Pan-cultural Elements in Facial Displays of Emotion", Science, 164, 86-88.

Ekman, P. & Friesen, W.(1971), "Constants across cultures in the face and emotion", Journal of personal and social psychology, 17, 124-129.

Ellis, H.D.(1986), "Processes underlying face recognition", in "The neuropsychology of face perception and facial expression", (R. Bruyer, ed.), Erlbaum Associates Inc., Hillsdale, New Jersey

Ellis, H.D., Jeeves, F., Newcombe, A. & Young, A., Eds.(1986), "Aspects of Face Processing", Dordrecht: Martinus Nijhoff.

Ellis, H.D., Shepherd, J. & Davies, G.M.(1975), "An investigation of the use of the Photo-fit technique for recalling faces", British Journal of Psychology, 1975, 66, 29-37.

Ellis, H.D., Deregowski, J.B. & Shepherd, J.W.(1975), "Descriptions of white and black faces by white and black subjects" International Journal of Psychology 10 119-123.

Ermiane, R. & Gergerian, E.(1978) "Atlas of facial expressions; Album des expressions du visage, Paris": La Pensee Universale. 2ndary reference (Ekman 1982).

Fehr, B. & Russell, J.A.(1984), "Concept of Emotion Viewed From a Prototype Perspective". Journal of Experimental Psychology: General, 113, 3, 464-486.

Friedman, M.(1937), "The use of ranks to avoid the assumption of normality implicit in the analysis of variance", Journal of the American Statistical Assoc., 32 675-701.

Frois-Wittman, J.(1930), "The Judgment of facial expression", J. Exper. Psychol., 13, 113-151. 2ndary. reference from Ekman(1982).

Fulcher, J.S.(1942), "Voluntary' facial expressions in blind and seeing children", Archives of Psychology, 38, 272, 1-49. 2ndary. reference from Ekman, 1982

Galper, R.E. & Hochberg, J.(1971). "Recognition memory for photographs of faces" American of Pschology, 84, 351-359.

Gashnig, J., Klahr, P., Pople, H., Shortliffe, E., & Terry, A.(1983), "Evaluation of expert systems: issues and case studies", Chapter 8 in "Building expert systems" eds. Hayes-Roth, Waterman & Lenat, Addison Wesley, Reading, Mass. U.S.A..

Gillenson, M.L.(1974),
"The interactive generation of facial images on a CRT using a heuristic strategy", Ph.D. thesis Ohio State University, Columbus, Ohio. 2ndary reference from Gillenson & Chandrasekaran,1975.

Gillenson, M.L. & Chandrasekaran, B.(1975),
"A heuristic strategy for developing human facial images on a CRT", Pattern Recognition, 7, 187 - 196.

Goldstein, A.G. & Mackenberg, E. J.(1966),
"Recognition of human faces from isolated facial features: A developmental study", Psychonomic Science, 6 149-150.

Hay, D.C. & Young, A.W.(1982),
"The human face", in "Normality and pathology in cognitive functions" (A.W. Ellis, ed.), Academic Press, London.

Haig, N.(1986),
"Investigating Face Recognition With AN Image Processing Computer" in "Aspects of Face Processing"(Ellis et al: eds.), Martinus Nijhoff, Lancaster.

Hjortsjo, C.(1970),
"Man's Face and Mimic Language", Student-literature, Lund, Sweden. 2ndary. reference from Ekman & Friesen, 1984

Hochberg, J. & Galper, R.E.(1967),
"Recognition of faces in photographic negative", Psychonomic Science 9: 619-620.

Izard, C.E.(1959),
"Positive affect and behavioral effectiveness", unpublished ms., Vanderbilt University. 2ndary reference from Izard, 1971.

Izard, C.E.(1960),
"Personality similarity and friendship", J. of abnormal and social psychology, 61(1), 47-51. 2ndary reference from Izard, 1971.

Izard, C.E.(1971),
"The Face of Emotion". Appleton-Century Crofts,New York.

Izard, C.E.(1979),
"Emotions as motivations: an evolutionary-developmental perspective". In "Nebraska Symposium on Motivation,1978" (R.A.Dienstbier, ed.), Univ. of Nebraska Press, Lincoln (2ndary reference from Saltzen, 1981).

Izard, C.E. & Tomkins, S. (1966),
"Affect and behavior: Anxiety as a negative affect". In C.D. Spielberger(ed.), "Anxiety and Behavior", N.Y.: Academic Press, pp. 81-125. 2ndary reference from Izard, 1971.

Jensen, D.G.(1986),
"Facial Perception: Holistic or Feature-Analytic?", Proceedings of the human factors society: 30th Annual Meeting, 30, 1, pp 729-733.

Jensen, D.G.(1987),
"Facial Perception studies using the Macintosh", Behavior Research Methods, Instruments, & Computers 1987, 19 (2), 252-256.

Johnson-Laird, P.N.(1989),
"Human experts and expert systems" in "Intelligent systems in a human context", Oxford University Press.

Jones, L.E. et al.(1976),
"The Face Atlas: Anthropometric,Cosmetic and Physiognomic Measurements of 200 Male Faces. Technical Report No. 1, NSF GS-42801. Champaign, Illinois: Dept. of Psychology, University of Illinois, 1976. 2ndary reference from Shepherd(1986).

Kaya, Y. & Kobayashi, K.(1972),
"A basic study on human face recognition" in 'Frontiers of Pattern Recognition' (S. Watanabe ed.), Academic Press, 1972

Kelly, G. A.(1955),
"The psychology of personal constructs" Norton, N.Y.

Kohonen, T.(1977),
"Associative Memory - A system-theoretical approach", Berlin: Springer-Verlag.

Kohonen, T., Oja, E. & Lehtio, P.(1981),
"Storage and processing of information in distributed associative memory systems", in "Parallel models of associative memory", G. Hinton & J.A. Anderson(Eds.), Hillsdale, New Jersey: Lawrence Erlbaum Associates Inc..

Kolodner, J.L.(1984),
"Retrieval and Organizational Strategies in Conceptual Memory: A Computer Model", Lawrence Erlbaum Assoc. Hillsdale, N.J.

Landis, C.(1924),
"Studies of emotional reactions: II. General behavior and facial expression", J. Comparative Psychology, 4, 447-509. 2ndary. reference from Ekman(1982).

Laughery, K., Duval, C., & Wogalter, M.(1986),
"Dynamics of Facial Recall", in "Aspects of Face Processing"(Ellis et al.:eds.), Martinus Nijhoff, Lancaster.

Laughery, K., Rhodes, B.,(Jr) & Batten, G.W.,(Jr) (1981),
"Computer-guided Recognition and Retrieval of Facial Images" in "Perceiving and Remembering Faces", Davies,G.M. et al (eds.), Academic Press, London.

Lebowitz, M.(1980),
"Generalization and memory in an integrated

understanding system". Ph.D. Thesis, Dept.of Computer Science, Yale University, New Haven, Conn., U.S.A.

Leeper, R., W. (1965), "Some needed developments in the motivational theory of emotions", in D. Levine(Ed.), Nebraska symposium on motivation. Lincoln: Univ. of Nebraska Press, pp 25-122. 2ndary reference from Izard, 1971.

Liebowitz, J. (1986), "Useful approach for evaluating expert systems", Expert Systems, 3, 2

Marr, D. (1982), "Vision", W.H.Freeman & Company, San Francisco.

Mase, K, Suenaga, Y. & Akimoto, T. (1987), "Head Reader - a head motion understanding system for better man-machine interaction", Proc. 1989 IEEE Int. Conf. on Systems, Man, and Cybernetics.

McClelland, J.L. & Rummelhart, D.E. (1985), "Distributed memory and the representation of general and specific information", Journal of Experimental Psychology: General, 114, 159-188

McDermott, J. (1986), "Face to Face : It's the Expression that bears the message", Smithsonian, March, 112-123.

Minsky, M. (1975), "A framework for representing knowledge". In "The Psychology of Computer Vision", Ed. Winston, P., McGraw-Hill, N.Y., pp 211-277.

Morton, J. (1969), "Interaction of Information in Word Recognition", Psychol. Rev., 76, 165-178.

Morton, J. (1979), "Facilitation in Word Recognition: Experiments causing change in the Logogen model", in "Processing of Visible Language", P.A.Kolers et al.( eds. ), Plenum, New York.

Mueller, J.H.,Carlomusto, M., and Goldstein, A.G.(1978), "Orienting task and study time in facial recognition", Bull. Psychonom. Soc., 11(5),313-316.

Norman, D,A.(1979 ), "Perception, Memory, and Mental processes" in "From Perspectives on Memory Research", ed. L. Nilsonn, Lawrence Erlbaum Assoc., Hillsdale, N.J.

Oatley, K. & Johnson-Laird, P.N.(1985), "Sketch for a Cognitive Theory of the Emotions", University of Sussex Cognitive Science Research Paper CSRP.045.

O'Keefe, R.M., Balci O, Smith, E.P.(1987), "Validating Expert System Performance" IEEE expert 2, 4, winter pp. 81-89.

O'Toole, A.J., Millward, R.B. & Anderson, J. A.(1988), "A Physical System Approach to Recognition Memory for Spatially Transformed Faces", Neural Networks, 1, pp. 179-199, 1988.

Patterson, K. E., Baddeley, A.(1977), "When Face Recognition Fails", J. Exper. Psychol. Human Learning and Memory 3, 4, 406-417.

Penry, J.(1971), Photo-fit Kit. Leeds, England: John Waddington of Kirkstall Ltd..

Penry, J.(1971), "Looking at faces and remembering them: A guide to facial identification". London: Elek Books, 1971.

Petajan, E.(1985), "Automatic lipreading to enhance speech recognition", Proc. IEEE Conference on C V P R, July, pp. 40-47.

Pilowski, I., Thornton, M., & Stokes, B.(1985), "A Microcomputer Based Approach To The Quantification Of Facial Expressions", Australasian Physical & Engineering Sciences in Medicine 8, 2.

Pilowski, I., Thornton, M., & Stokes, B.(1986), "Towards The Quantification Of Facial Expressions With The Use of a mathematical Model of the Face" in "Aspects of Face Processing", Ellis et al.(eds.), Martinus Nijhoff Publishers, Lancaster.

Plutchik, R.(1962), "The emotions: Facts, theories, and a new model". N.Y., Random House.

Reiser, B., Black,J.B. & Abelson, R.P.(1985),, "Knowledge Structures in the organization and retrieval of autobiographical memories", Cognitive Psychology, 17, 89-137.

Robinson,J.A.(1976), "Sampling autobiographical memory", Cognitive Psychology, 8, 578-595.

Rhodes, B. & Bargainer, J.(1976), "A mini-computer system for retrieval of look-alikes from a mug file", Proc.of the 1976 Systems Engineering Conf. AIIE, Atlanta, Georgia, U.S.A.

Rhodes, M. & Klinger, A.(1977), "Conversational text input for modifying graphics facial images", in "Fuzzy Reasoning and its applications", eds. Mamdani & Gaines, Academic Press, London.

Riesbeck, C.K.(1981), "Failure-driven reminding for incremental learning", in "Proc. of the seventh int. joint confer. on artificial intelligence", pp115-120.

Riesbeck, C.K.(1984 ), "Knowledge Reorganization and Reasoning Style", Inter. J. Man-Machine Studies,20, pp45-61.

Ringland, G.A. & Duce, D.A.(1988),
"Approaches to knowledge representation: An introduction", Research Studies Press Ltd. Letchford, Herts.

Roseman, I.(1982),
"Cognitive aspects of discrete emotions", Ph.D. Thesis, Yale University.

Sakai, T., Ngao, M., Fujibayashi, S.(1969),
"Line extraction and pattern recognition in a photograph". Pattern Recognition, 1, 233-248.

Sakai, T., Ngao, M. & Kanade, T.(1972),
"Computer analysis and classifications of photographs of human faces". Proceedings of the First USA-JAPAN Computer conference, pp. 55-62.

Salberg, C.(1983),
"Generating Hypotheses to Explain Prediction Failures", Paper presented at AAAI-83, Washington,DC, August 1983.

Salzen, E.A.(1981),
"Perception of Emotion in Faces": Chapter. 7 in "Perceiving and Remembering Faces", Davies, G.M. et al (eds.), Academic Press, London.

Schacter,S. & Singer,J.E.(1962),
"Cognitive, social, and physiological determinants of emotional state", Psychological Review, 69, 5, pp 379--399. 2ndary reference from Izard(1971).

Schank, R.C.(1986),
"Explanation Patterns; understanding mechanically and creatively", Lawrence Erlbaum ssociates, Inc., Hillsdale, New Jersey.

Schank, R.C.(1984),
"Learning, Explanation, and a Little History", Invited talk at First Annual Workshop on Theoretical Issues in Conceptual Information Processing, Atlanta,GA, March, 1984.

Schank, R.C.(1982),
"Dynamic Memory: A theory of reminding and learning in computers and people", C.U.P. Cambridge.

Schank,R.C. & Abelson,R.(1977),
"Scripts, Plans, Goals and Understanding", Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Scragg, G.(1976),
"Semantic Nets as Memory Models" in "Computational Semantics" (Charniac, E. & Wilks, Y.,eds.1976), North-Holland Publishing Company.

Sergent, J.(1984),
"An investigation into component and configural processes underlying face perception", Brit. J. of Psychology, 75, 221-242.

Sheehy,N.(1989),
"Non-verbal behaviour in the demonstrator", in "Communication Failure in Dialogue: techniques for detection and repair", Deliverable 9, Implementation of the dialogue system, 4 April, ESPRIT Project 527 (CFID), University of Leeds, Leeds, Yorks., England.

Shepherd, J.W.(1986),
"An Interactive Computer System For Retrieving Faces" in "Aspects of Face Processing",(Ellis et al:eds.), Martinus Nijhoff, Lancaster.

Shepherd, J.W., Davies,G.M. & Ellis H.D.(1981),
"Studies of cue saliency" in "Perceiving and Remembering Faces" ,Davies, G.M., Ellis, H.D., Shepherd, J.W.(eds.), Academic Press, London.

Shortliffe, E. & Davis, R.(1975),
"Some considerations for the implementation of knowledge-based systems", SIGART Newsletter, no. 55, December, pp. 9-12.

Simon, H.A. (1967),
"Motivational and emotion controls of

cognition", Psychological Review, 74, 29-39.

Simonov, P. V. (1966), "The Informational Theory of Emotion", Part of a report presented at the Congress of Psychology, Moscow, 1966.

Sloman, A.(1986), "Motives, Mechanisms and Emotions", University of Sussex Cognitive Science Research Reports, Serial No. CSRP 062., School of Social Studies, Falmer, Brighton, BN1 9QN, England.

Sloman, A. & Croucher, M.(1981a), "Why robots will have emotions", International Joint Conference on Artificial Intelligence,7.

Sloman, A. & Croucher, M.(1981b), "You don't need a soft skin to have a warm heart", Cognitive Science Research Paper serial No. csrp 004 University of Sussex Cognitive Studies Programme, School of Social Sciences, Falmer, Brighton, U.K. BN1 9QN.

Sloman, A.(1978), "The Computer Revolution in Philosophy, Science and Models of the Mind", Harvester Press.

Sorce, J. & Campos, J.J,(1974), "The role of expression in the recognition of a face", Amer. J. Psychol., 87, pp.71-82.

Stonham, T.J.(1986), "Practical Face Recognition And Verification With Wisard" in "Aspects of Face Processing", (Ellis et al:eds.), Martinus Nijhoff, Lancaster.

Strnad, B., & Mueller, J.H.(1977), "Levels of processing in facial recognition", memory., Bull. Psychonom. Soc., 9, 17-18.

Thornton, M. & Pilowski, I.(1982), "Facial expressions can be modelled mathematically", Brit. J. Psychiat., 140, 61-63.

Tiberghien, G.(1986),
"Context effects in recognition memory of faces: some theoretical problems", in "Aspects of face processing", eds. Ellis, H. et al., Martinus Nijhoff Publishers, Lancaster.

Tomkins,S.S.(1962),
"Affect, imagery, consciousness. Vol. 1 The positive affects". N.Y: Springer. 2ndary reference from Izard(1971).

Tomkins,S.S.(1963),
"Affect, imagery, consciousness. Vol. 2 The negative affects". N.Y: Springer.2ndary reference from Izard(1971).

Townes, J.R.(1976),
"A computer algorithm for Mug Shot Identification". EAI Symposium on Automatic Imagery Pattern Recognition, College Park, Maryland, April, 1976. - 2ndary reference from Rhodes & Bargainer.

Tulving, E.(1983),
"Elements of Episodic Memory". Oxford University Press.

Tulving, E.(1972),
"Episodic and semantic memory", In "Organization and Memory" (E. Tulving & W. Donaldson,eds), Academic Press, New York.

Turk, M. & Pentland, A.(1990),
"Face Processing Models For Recognition", Proc. SPIE INT. SOC. OPT. ENG. U.S.A.. Philadelphia, Pa., vol 1192 Intelligent Robots and Computer Vision V111: Algorithms and Techniques(1989), Part 1, pp. 22-32.

Watkins. M.J., Ho, E. & Tulving, E.(1976),
"Context Effects in Recognition Memory for Faces", Journal of Verbal Learning and Verbal Behaviour, 15, 505-517.

Warrington, E., & Ackroyd, C.(1975),
"The effect of orienting tasks on recognition

memory", Memory and Cognition 3(2), 140-142.

Weiss, S.M. & Kulikowski, C.A.(1983), "A practical guide to designing expert systems", Chapman and Hall, London.

Wells, G.L. & Hryciw, B. A.(1984), "Memory for faces: Encoding and retrieval operations", Memory and Cognition, 12, 338-344.

Williams, G.W.(1976), "Comparing the Joint Agreement of Several Raters with Another Rater", Biometrics, 32, 619-627, September.

Winograd, E.(1976), "Recognition memory for faces following nine different judgments", Bull. Psychonom. Soc., 8, 419-421.

Winograd, E. & Rivers-Bulkeley, N. T.(1977), "Context Effects On Remembering Faces" Journal of Experimental Psychology: Human Learning and Memory, 3, 4, 397-407.

Winograd, E.(1978), "Encoding operations which facilitate memory for faces across the life span." in M.M. Gruneberg, P.E.Morris & R.N. Sykes(eds.) "Practical Aspects of Memory", London, Academic Press,1978.

Winograd, E.(1981), "Elaboration and distinctiveness in memory for faces", Journal of Experimental Psychology: Human Learning and Memory, 7, 181 - 190.

Wong, K.H. et al.,(1989), "A system for recognising human faces", Proc: ICASSP-89 INT. CONF. On Acoustics: Speech and Signal Processing. IEEE 3 pp. 1638-42.

Wyatt, J., & Spiegelhalter, D.(1990), "Evaluating Medical Expert Systems: what to test and how", Medical Informatics, 15, pp 205 - 217.

Yin, R.K. (1969), "Looking at upside-down faces", Journal of Experimental Psychology 81, 141-145.

Zajonc, R. B. (1980), "Feeling and Thinking", American Psychologist, 35, 2, 151-175.

# Appendix:

# Contents:

# Appendix A:

## A.1 POP11 code for the Rule Base:

/*....*/ encloses a comment; ';;;' prefixes a comment.

```
;;;rules with norm alone : these define face_attributes in terms of a 2_D
;;;Cartesian graph on which 34 xy-points are specified corresponding to
;;;standard face components.



load normgarrylist.p;

;;;normgarrylist.p contains a list of lists assigned to the variable "lookslist".
;;;Each of the 18 embedded lists contains the 34 xy-positions
;;;of a face-photo in a set order of pairs. As required, normgarrylist(1..18) is
;;;assigned to the variable "mug", thus defining which face is referred to.
;;;The points to be measured are then indexed in the embedded list(= mug).


vars mug2 mug3 mug6 mug7 mug9 mug10 l norm2 norm3 norm6 norm7 ;
vars brows_contracted w norm  brows_low brows_raised brows_horseshoe;
vars brows_in_raised eyes_inlid_raised eyes_u_lid_lowered eyes_closed;
vars mouth_bared nose_screwed nose_flared mouth_down mouth_up;
vars mouth_square mouth_u_lip_tensed mouth_l_lip_everted mouth;
vars mouth_shut mouth_l_lip_lowered mouth_open;
vars mouth_wide  mouth_pulled cheeks_ n_l_vert cheeks_raised;
vars mouth_u_lip_everted eyes_u_lid_tensed;
vars mouth_slightly_open mouth_down mouth_compressed ;
vars eyes_wide mouth_l_lip_raised mouth_u_lip_raised mouth_smile;
vars eyes_down eyes_wide eyes_l_lid_raised eyes_u_lid_raised;
vars  yy xx n jaw_drop cheeks_dropped mouth_l_lip_tensed eyes;
vars l_lid_tensed eyes_narrowed norm9 norm10;


normgarrylist(2)  -> norm2;
vars r_brow_i_nx = norm2(34)(1);
vars r_brow_i_ny = norm2(34)(2);
vars r_brow_o_nx = norm2(12)(1);
vars r_brow_o_ny = norm2(12)(2);
vars l_brow_i_nx = norm2(33)(1);
vars l_brow_i_ny = norm2(33)(2);
vars l_brow_o_nx = norm2(11)(1);
vars l_brow_o_ny = norm2(11)(2);
vars l_u_lid_nx = norm2(9)(1);
vars l_u_lid_ny = norm2(9)(2);
vars r_u_lid_nx = norm2(31)(1);
vars r_u_lid_ny = norm2(31)(2);
```

```
vars l_l_lid_nx = norm2(10)(1);
vars l_l_lid_ny = norm2(10)(2);
vars r_l_lid_nx = norm2(32)(1);
vars r_l_lid_ny = norm2(32)(2);
vars l_eye_i_nx = norm2(2)(1);
vars l_eye_o_nx = norm2(1)(1);
vars l_eye_o_ny = norm2(1)(2);
vars r_eye_i_nx = norm2(4)(1);
vars r_eye_i_ny = norm2(4)(2);
vars l_eye_i_ny = norm2(2)(2);
vars r_eye_o_nx = norm2(3)(1);
vars r_eye_o_ny = norm2(3)(2);
vars l_pupil_nx = norm2(7)(1);
vars l_pupil_ny = norm2(7)(2);
vars r_pupil_nx = norm2(8)(1);
vars r_pupil_ny = norm2(8)(2);

vars l_nose_nx = norm2(13)(1);
vars l_nose_ny = norm2(13)(2);
vars r_nose_nx = norm2(14)(1);
vars r_nose_ny = norm2(14)(2);
vars c_nose_nx = norm2(15)(1);
vars c_nose_ny = norm2(15)(2);
vars l_lip_nx = norm2(16)(1);
vars l_lip_ny = norm2(16)(2);
vars botlip_ct_nx = norm2(21)(1);
vars botlip_ct_ny = norm2(21)(2);
vars botlip_cb_nx = norm2(20)(1);
vars botlip_cb_ny = norm2(20)(2);
vars r_lip_nx = norm2(17)(1);
vars r_lip_ny = norm2(17)(2);
vars toplip_ct_nx = norm2(18)(1);
vars toplip_ct_ny = norm2(18)(2);
vars toplip_cb_nx = norm2(19)(1);
vars toplip_cb_ny = norm2(19)(2);
vars cheek_l_nx = norm2(28)(1);
vars cheek_l_ny = norm2(28)(2);
vars cheek_r_nx = norm2(27)(1);
vars cheek_r_ny = norm2(27)(2);
vars jaw_nx = norm2(22)(1);
vars jaw_ny = norm2(22)(2);
vars pro_l_ny  = norm2(23)(2);
vars pro_r_ny = norm2(24)(2);
vars teeth_t_ny = norm2(25)(2);
vars teeth_b_ny = norm2(26)(2);
```

```
define brows_contracted( mug ) ->i;
   ;;;The distance between innermost points on the eyebrows is less than
   ;;; that of norm2 The brow is contracted towards the midline.
vars r_brow_i_x = mug(34)(1), l_brow_i_x = mug(33)(1), weight = -0.2;
(
abs((r_brow_i_x) - (l_brow_i_x)) <
abs((r_brow_i_nx) - (l_brow_i_nx) + weight)
 ) -> i;
enddefine;

define brows_low(mug)-> i;
   ;;;the brow is lowered onto the eye. Y-distances between top of
   ;;;eyebrow and inner angle of the eye are less that of norm2.
vars l_brow_o_y = mug(11)(2), l_eye_i_y = mug(2)(2),
r_brow_o_y = mug(12)(2), r_eye_i_y = mug(4)(2), weight = 0.7;

(
  (
     mug /= norm2 and
    (abs(l_brow_o_y - l_eye_i_y) < abs(l_brow_o_ny - l_eye_i_ny -
weight) or
    abs(r_brow_o_y - r_eye_i_y) < abs(r_brow_o_ny - r_eye_i_ny -
weight))

  )
) ->i;

enddefine;

define brows_in_raised(mug)-> i;
   ;;;The vertical difference between the medial end of the eyebrow and
   ;;;the uppermost point of it above centre-pupil is less than the norm.
    vars l_brow_o_y = mug(11)(2);
    vars l_brow_i_y = mug(33)(2);

(
(abs(l_brow_o_y - l_brow_i_y) < abs(l_brow_o_ny - l_brow_i_ny))
) -> i;
enddefine;

define brows_raised( mug ) -> i;
    ;;;the right eye-brow is raised compared to norm2.
    vars l_brow_o_y = mug(11)(2), weight = 1.4;

(
```

```
    l_brow_o_y < (l_brow_o_ny - weight )
) ->i;
enddefine;

define brows_centre_raised(mug) -> i;
    ;;;The medial end of the right eyebrow is higher than that of the norm
    ;;;and the brow is contracted.
vars l_brow_i_y = mug(33)(2), weight = 1.4;
(
  l_brow_i_y < (l_brow_i_ny) - weight;
  and brows_contracted(mug)
)-> i;
enddefine;

define eyes_u_lid_raised(mug) -> i;
    ;;;The vertical distance between the centre of the pupil and the upper
    ;;;eyelid is greater than the norm.
vars l_u_lid_y = mug(9)(2), l_pupil_y = mug(7)(2), weight = 0.18;
(
  (l_u_lid_y - l_pupil_y) < (l_u_lid_ny - l_pupil_ny - weight)
) -> i;
enddefine;

define eyes_inlid_raised(mug) -> i;

    ;;;centre-raised brows  or in-raised brows
    ;;;and contracted brows have to be true.

(
(brows_centre_raised(mug)  or brows_in_raised(mug))
and brows_contracted(mug)
) -> i;
enddefine;

define eyes_u_lid_lowered(mug) -> i;
    ;;;The right upper lid margin mid point is lower than that of the norm
vars l_u_lid_y = mug(9)(2);
(
 l_u_lid_y  > l_u_lid_ny
 ) -> i;
enddefine;

define eyes_l_lid_raised(mug) -> i;
    ;;;The vertical distance between the inner canthus of the right eye
    ;;;and the point on the lower eyelid below centre of pupil is less than
    ;;;that of the norm.
```

```
vars l_l_lid_y = mug(10)(2);
vars l_eye_i_y = mug(2)(2);
(
 (l_l_lid_y - l_eye_i_y) < (l_l_lid_ny - l_eye_i_ny)
) -> i;
enddefine;

define eyes_l_lid_lowered(mug) -> i;
   ;;;The vertical distance between the inner canthus of the right eye
   ;;;and the point on the lower eyelid below centre of pupil is greater
   ;;;than that of the norm.
vars l_l_lid_y = mug(10)(2);
vars l_eye_i_y = mug(2)(2);
(
 (l_l_lid_y - l_eye_i_y) > (l_l_lid_ny - l_eye_i_ny)
) -> i;
enddefine;

define eyes_l_lid_tensed(mug) -> i;
   ;;;The lower lid is raised and the cheeks are not raised.
(eyes_l_lid_raised(mug) and not(cheeks_raised(mug))) -> i;
enddefine;

define eyes_u_lid_tensed(mug) -> i;
   ;;;The upper lid is lowered and the upper lid is medially raised putting
   ;;;strain on the eyelid.
(eyes_u_lid_lowered(mug) and eyes_inlid_raised(mug) and
eyes_l_lid_tensed(mug))-> i;
enddefine;

define eyes_down mug ->i;
   ;;;the upper lids are lowered as in reading The upper lid margin
   ;;;mid point is lower than the inner canthus of the
   ;;;right eye
vars l_u_lid_y = mug(9)(2);
vars l_eye_i_y = mug(2)(2);

 l_u_lid_y > l_eye_i_y -> i;
enddefine;

define open_eyed( mug, entry )-> i;
     ;;;defined by the gap between upper and
     ;;;lower lids(y-values) arbitrary
     ;;;cut-off y-values.

vars gap norm,
```

```
l_u_lid_y = mug(9)(2),
l_l_lid_y = mug(10)(2), weight = 0.3333;
    abs(l_u_lid_y - l_l_lid_y) ->gap;
    abs((l_u_lid_ny) - (l_l_lid_ny)) - weight -> norm;

    switchon entry
    case = [wide] then (gap > norm and (mug /= norm2))
    case = [narrowed] then 0 < gap and gap < norm and mug /= norm2
    case = [shut] then gap = 0  and mug /= norm2
    else false
    endswitchon -> i
enddefine;
open_eyed(% [wide] %) -> eyes_wide;
open_eyed(% [narrowed] %) -> eyes_narrowed;
open_eyed(% [shut] %) -> eyes_closed;

define nose_flared mug -> i;
   ;;;the width of the nose measured along the x axis
   ;;;at its widest point is greater
   ;;;than a norm.
vars r_nose_x = mug(14)(1), l_nose_x = mug(13)(1),
weight = 0.2;

( mug /= norm2 and
 (abs(r_nose_x) + abs(l_nose_x)) >
 (abs(r_nose_nx) + abs(l_nose_nx - weight))
) ->i;

enddefine;

define mouth_up(mug) -> i;
   ;;;the y-level of the corners must be above
   ;;;that of mid upper - lip lower boundary and their vertical
   ;;;distance from the inner canthi of the eyes must be less than the norm.
vars r_lip_x = mug(17)(1),
r_lip_y = mug(17)(2),
l_lip_x = mug(16)(1),
l_lip_y = mug(16)(2),
toplip_cb_x = mug(19)(1),
toplip_cb_y = mug(19)(2),
r_eye_i_y = mug(4)(2),
l_eye_i_y = mug(2)(2), weight = 0.1;

(  mug /= norm2 and
(l_lip_y < toplip_cb_y ) and
(r_lip_y < toplip_cb_y )  and
```

```
(abs(l_lip_y - l_eye_i_y - 0.1)  < abs((l_lip_ny) - (l_eye_i_ny))) and
(abs(r_lip_y - r_eye_i_y - 0.1)  < abs((r_lip_ny) - (r_eye_i_ny)) )
) -> i;
enddefine;

define nose_screwed(mug)  ->i;

    ;;;The brows are lowered and the cheeks are raised
    ;;;this indicates horizontal puckering in the skin of the nose
    ;;;and lowered brow which frequently accompanies it.
(
       brows_low(mug)
       and cheeks_raised(mug)
)       -> i;
enddefine;

define open_mouth( mug, entry )-> i;
   ;;;defined by the gap between upper and lower lips(y-values) arbitrary
   ;;;cut-off y-values.
vars botlip_ct_y = mug(21)(2),
toplip_cb_y = mug(19)(2),
l_eye_i_x = mug(2)(1),
toplip_ct_y = mug(18)(2),
teeth_t_y = mug(25)(2),
teeth_b_y = mug(26)(2),
gap, weight = 0.5, gapmargin = 1.0;
      botlip_ct_y - toplip_cb_y ->gap;
      switchon entry
      case = [square] then ((gap > 4.0 * gapmargin) and
            (abs(teeth_t_y - teeth_b_y)
            > weight)
            and (toplip_ct_y < toplip_ct_ny)
           and ((toplip_cb_y - toplip_ct_y)
            >
            (toplip_cb_ny - toplip_ct_ny + 0.092 * gapmargin)))
      case = [wide] then gap > 3.0 * gapmargin
      ;;;case = [bared] then 1.35 * gapmargin < gap and gap < 3.0 * gapmargin
      case = [bared] then abs(teeth_t_y - teeth_b_y) < (gap - 0.4 * gapmargin)
      case = [slight] then 0.76552 * gapmargin < gap and gap < 1.35 * gap-
margin
      case = [open] then gap > 0.76552 * gapmargin
      case = [shut] then gap < 0.76552 * gapmargin or gap = 0.76552 * gap-
margin
      else false
      endswitchon -> i
```

```
enddefine;
open_mouth(% [wide] %) -> mouth_wide;
open_mouth(% [bared] %) -> mouth_bared;
open_mouth(% [slight] %) -> mouth_slightly_open;
open_mouth(% [open] %) -> mouth_open;
open_mouth(% [square] %) -> mouth_square;
open_mouth(% [shut] %) -> mouth_shut;



define mouth_compressed(mug) -> i;
;;;The mouth must be shut and the vertical measure of the centre upper lip
;;;red part must be less than the norm. Alternatively, the mouth must be
;;;bared, the mouth open, and the upper and lower teeth together and the
;;;mouth corners must not be turned upwards.
vars botlip_ct_y = mug(21)(2),
toplip_cb_y = mug(19)(2),
l_eye_i_x = mug(2)(1),
toplip_ct_y = mug(18)(2),
teeth_t_y = mug(25)(2), weight = 0.092,
teeth_b_y = mug(26)(2);
(
 (mouth_shut(mug) and
 ((toplip_cb_y - toplip_ct_y) < (toplip_cb_ny - toplip_ct_ny + weight )))
  or (mouth_bared(mug) and mouth_open(mug) and (teeth_t_y = teeth_b_y)
and
     not(mouth_up(mug)))
) -> i;
enddefine;



define mouth_down(mug) -> i;
   ;;;One mouth angle  must be lower than the centre of the lower
   ;;;margin of the upper lip and the y-distance between the inner canthi of
the
   ;;;eyes and the respective corner must be greater than the norm.
vars toplip_cb_y = mug(19)(2),
l_eye_i_y = mug(2)(2),
r_eye_i_y = mug(4)(2),
r_lip_y = mug(17)(2), weight = 0.1,
l_lip_y = mug(16)(2);
(
not(mouth_open(mug)) and
(  mug /= norm2) and
(l_lip_y > toplip_cb_y  or
```

```
r_lip_y > toplip_cb_y )  and
abs(l_lip_y - l_eye_i_y - weight)  > abs(l_lip_ny - l_eye_i_ny) and
abs(r_lip_y - r_eye_i_y - weight)  > abs(r_lip_ny - r_eye_i_ny)
) -> i;
enddefine;

define mouth_pulled(mug) -> i;
   ;;;The x-value of each corner must be displaced laterally compared
   ;;;to the norm and the corners must not be "up".
vars r_lip_x = mug(17)(1),
l_lip_x = mug(16)(1);

(
(l_lip_x < l_lip_nx) and (r_lip_x > r_lip_nx)
and not(mouth_up(mug))
) -> i;
enddefine;

define mouth_u_lip_raised(mug) -> i;
   ;;;The centre point of the lower margin of the upper lip is displaced up-
wards
   ;;;compared to the norm.
vars toplip_cb_y = mug(19)(2);
(
 toplip_cb_y < toplip_cb_ny
) -> i;
enddefine;

define mouth_u_lip_tensed(mug) -> i;
   ;;;The y-distance in the midline between the margins of the red
   ;;;part of the lip is greater than the norm. The corners of the mouth
   ;;;must not be turned upwards. Alternatively, the mouth is compressed.
vars toplip_ct_y = mug(18)(2), toplip_cb_y = mug(19)(2);
 (
  (mug /= norm2 and
(toplip_cb_y - toplip_ct_y) > (toplip_cb_ny - toplip_ct_ny)
and
 not(mouth_up(mug))) or mouth_compressed(mug)
) -> i;
enddefine;


define mouth_l_lip_lowered(mug) -> i;
   ;;; the y-value of the centrepoint on the bottom lip upper margin
   ;;; is lowered in comparison to the neutral value.
vars botlip_ct_y = mug(21)(2), weight = 0.3;
```

```
(
 (botlip_ct_y - weight) > botlip_ct_ny
) -> i;
enddefine;

define mouth_l_lip_tensed(mug) -> i;
   ;;;the height of the lower lip is greater than the neutral lip.
   ;;;and the mouth is not 'up' nor 'compressed'.
vars botlip_ct_y = mug(21)(2),
botlip_cb_y = mug(20)(2);
 (  mug /= norm2 and
(botlip_cb_y - botlip_ct_y) > (botlip_cb_ny - botlip_ct_ny)
and not(mouth_up(mug)) or mouth_compressed(mug)
) -> i;
enddefine;

define mouth_l_lip_raised(mug) -> i;
   ;;;The centre y-value of the upper margin of the lower lip is
   ;;;displaced upwards compared to the norm.
vars botlip_ct_y = mug(21)(2), weight = 0.5;
(
 (botlip_ct_y + weight) < botlip_ct_ny
) -> i;
enddefine;

define mouth_l_lip_everted(mug) -> i;
   ;;;The centre y-value of the upper margin of the lower lip is
   ;;;displaced upwards and the lower lip red part is greater vertically com-
pared
   ;;;to the norm. The mouth is not 'up'.
vars botlip_ct_y = mug(21)(2);
vars botlip_cb_y = mug(20)(2);
(
 (botlip_ct_y < botlip_ct_ny) and not(mouth_up(mug))
 and ((botlip_cb_y - botlip_ct_y) >
    ( botlip_cb_ny - botlip_ct_ny))
) -> i;
enddefine;

define mouth_u_lip_everted(mug) -> i;
   ;;;The centre y-value of the upper margin of the upper lip is
   ;;;displaced upwards and the upper lip red part is greater vertically com-
pared
   ;;;to the norm.
toplip_cb_y = mug(19)(2);
toplip_ct_y = mug(18)(1);
```

```
(
 (toplip_ct_y < toplip_ct_ny) and not(mouth_up(mug))
 and ((toplip_cb_y - toplip_ct_y) >
    (toplip_cb_ny - toplip_ct_ny))
) -> i;
enddefine;

define  cheeks_n_l_vert(mug)->l;
   ;;;The nose is flared and the point at which
   ;;;two projected lines from top cheek and mouth line meet is below a point
close
   ;;;to the origen on the y-axis.
vars pro_l_y = mug(23)(2), weight = 0.5;
   (
    nose_flared(mug)
    and (pro_l_y > - weight)
   )-> i;
enddefine;

define cheeks_raised mug ->l;
   ;;;The points at which
   ;;;two projected lines from top cheek and mouth line meet are above a
point close
   ;;;to the origen on the y-axis and the lower eyelid is raised. The jaw is not
   ;;;dropped.
vars pro_r_y = mug(24)(2),
pro_l_y = mug(23)(2), weight = 1.7;
(
not(jaw_drop(mug)) and
eyes_l_lid_raised(mug) and ((pro_r_y < weight)  or
(pro_l_y < weight))
   )->i;
enddefine;

define cheeks_dropped mug ->l;
   ;;;The points at which two projected lines from top cheek and mouth line
   ;;;meet are lower in y-value than the norm. The lower eyelid is not raised.
vars pro_r_y = mug(24)(2);
vars pro_l_y = mug(23)(2);

   (  not(eyes_l_lid_raised(mug)) and
    (pro_r_y > pro_r_ny)    and
    (pro_l_y > pro_l_ny)
   )->i;

enddefine;
```

```
define jaw_drop(mug) -> i;
  ;;;The mouth is wide open.
(
mouth_wide(mug)
) -> i;
enddefine;
```

## A.2 POP11 code for the Janus flavour hierarchy and cognitive dimensions:

```
uses flavours;
vars efsub make_sense explain affect info links ;
vars mop_surprised_gen_Face potsublist;
vars mop_happy_gen_Face;
vars mop_sad_gen_Face;
vars mop_disgusted_gen_Face;
vars mop_afraid_gen_Face;
vars mop_angry_gen_Face decompose compose road;
vars nod t_add wait spec latestmop facet links;
vars calc event_count event_list;
flavour episode;
ivars input = [], fcon = [] , time = undef ;
endflavour;

flavour interpret;
ivars x y emx emy name;
defmethod after initialise;
;;;matches the two emotions in a dyad to their goal state
vars d = 'desired', u= 'undesired', p = 'present', a = 'absent', c = 'certain';
vars uc = 'uncertain', d = 'deserved', ud = 'undeserved', you = 'sit/or/oth';
vars sg = 'satisfied_goal',ug = 'unsatisfied_goal',usg = 'undes_satis_goal';
vars motive sit prob leg agency g cogdimens item parlist;
vars xsit ysit conflict ;
;;; each Face_mop emotion term is associated with 6 cognitive dimensions
;;;which can take theory-based values(after Roseman 1982). see text for de-
tails
[[happy ^x  motive d sit p prob c leg undef agency you g sg ]
 [sad ^x  motive d sit a prob c leg ud agency you g ug]
 [afraid ^x motive u sit p prob uc leg ud agency you g ug]
 [angry ^x motive u sit p prob c leg ud agency you g ug]
 [disgusted ^x motive u sit p prob c leg ud agency you g usg]
 [surprised ^x  motive undef sit p prob c leg undef agency you g undef]
]-> cogdimens; ;;;associates each emotion term with dimension values.
These ;;;are printed out in text in the following code and procedure
'make_sense' draws ;;;upon them to infer the cause for the expressed emo-
tion in each member of the ;;;dyad.
foreach [^emx == ] in cogdimens do
it -> xsit endforeach;
foreach [^emy == ] in cogdimens do
it -> ysit endforeach;
y ->ysit(2);
[%for item from 1 by 2 to 13 do
[%xsit(item),xsit(item + 1),ysit(item),ysit(item + 1)%];
```

```
endfor%]->parlist;
make_sense(parlist(1)(1),parlist(1)(2),parlist(1)(3),parlist(1)(4));
[% [%
parlist(1)(1),parlist(1)(2),parlist(1)(3),parlist(1)(4)%];
for item from 2 by 1 to 7 do
if parlist(item)(2) /= parlist(item)(4) then
[%(parlist(item)(1),parlist(item)(2)),(parlist(item)(4))%] endif endfor
  %] -> conflict;
^explains(conflict);
pr(conflict)
enddefmethod;

defmethod printself;
   pr('<interpret_' >< name >< '>\n')
enddefmethod;
endflavour;

define make_sense(em1,one,em2,Other);
for affect in [^em1 ^em2] do
switchon affect
case = "happy"
      then
    pr(one);
    pr(' considers the presence, information or attention or motivations of ');
    pr(Other);
    pr(' to be desirable and ');
    pr(Other);
    pr(' is present. \n');
    pr('Hence ');
    pr(Other);
    pr(' is responsible for the happiness of \n');
    pr(one);
    pr(' who feels deserving of the interest of \n');
    pr(one);
    pr('  .Alternatively the happiness of ');
    pr(one)  ;
    pr(' may be quite unconnected with \n');
    pr(Other);
    pr(', being due to the satisfaction of a desired goal\n');
    pr('outside the present situation.');
 case = "sad"
      then
    pr(one);
  pr(' considers an outcome involving ');
    pr(Other);
    pr(' to be ');
```

```
    pr('desirable but to be unattainable; or some thing or person or\n');
    pr('desireable state is absent, or has news to that effect\n');
    pr('Hence ');
    pr(one);
    pr(' may feel unhappy and if deserved, may feel guilt. \n');
    pr('or if not, anger.\n');

case = "afraid"
      then
    pr(one);
    pr(' considers the presence of ');
    pr(Other);
    pr('or what has passed between them to be\n');
    pr('undesirable and to hold some threat. \n');
    pr(one);
    pr(' may feel he has deserved some retribution \n');
    pr('or anger.  Alternatively his fear may be \n');
    pr('baseless or unconnected with ');
    pr(Other);
    pr( ',being related to another\n');
    pr(' cause in the situation.');

case = "angry"
      then
    pr(one);
    pr(' considers the presence of ');
    pr(Other);
    pr('to be undesirable. \n');
    pr(one);
    pr(' alternatively feels that ');
    pr(Other);
    pr(' was responsible\n');
    pr(' for some negative outcome which he did not deserve, and\n');
    pr(' feels angry.  Again, the anger may be');
    pr(' unconnected with');
    pr(Other);
    pr(' and to do with different people or \n');
    pr('undirected, being due to the thwarting of a goal.\n')

case = "disgusted"
       then
    pr(one);
    pr(' is confronted by a situation which he tries \n');
    pr('to avoid but is confronted with and reacts with disgust and\n');
    pr(' distress.  The cause may be connected with ');
    pr(Other);
```

```
    pr(' - either his/her appearance, ');
    pr('habits or some event he is reporting; or may be ');
    pr('to do with different people or \n');
    pr('animals in the situation.\n')


case = "surprised"
       then
    pr(one);
    pr('is confronted by a situation for which he has \n');
    pr('no current goal motive and reacts with surprise \n');
    pr('The emotion is fleeting and is overwritten by the emotion\n');
    pr('relevant to what caused the surprise.  The cause may be \n');
    pr('connected with');
    pr(Other);
    pr(' either her appearance, habits or some event she is reporting\n');
    pr('or may be');
    pr(' unconnected, and to do with different people or \n');
    pr('animals in the situation\n')

endswitchon
endfor
enddefine;


define explains(dimension);
vars item;
for item in dimension do
  switchon(item(1))
  case = "motive" then
      if item(2) = d
      then
        pr('There is a disproportion in the motivation of the dyad.  \n');
        pr(parlist(1)(2));
        pr(' wishes to attain a current goal more than does \n');
        pr(parlist(1)(4));
      if parlist(1)(1) = "happy"
   then
          pr(' A desired outcome has been attained or an undesired');
          pr(' outcome has been avoided.');
        else
          pr(' A undesired outcome has been attained or an desired');
          pr(' outcome has not been avoided.')
        endif
        endif;
  case = "sit" then
```

```
      if item(2) = p
      then
        if parlist(1)(1) = "happy"
        then
          pr('The presence of a desirable state or the absence of an \n');
          pr('undesirable state can be inferred for ');
          pr(parlist(1)(2));
          pr(' but not for \n');
          pr(parlist(1)(4));
        else
          pr('The presence of an undesirable state or the absence of \n');
          pr('a desirable state can be inferred for ');
          pr(parlist(1)(2));
          pr(' but not for ');
          pr(parlist(1)(4));
        endif
        endif
   case = "prob" then
      if item(2)       = "c" and
         parlist(1)(2) = "d" and
         parlist(2)(2) = "p" and
         parlist(1)(1) = "happy"

      then
          pr('The presence of \n');
          pr(parlist(1)(4));
          pr(' is a desirable state for \n');
          pr(parlist(1)(2));
          pr(' but not vica versa.');
      else
          pr(parlist(1)(2));
          pr(' is hopeful for the desirable outcome.');
      endif


   case = "leg" then
      if item(2) = "d" and
         member((parlist(1)(2)),[disgusted sad afraid])
      then
         pr(parlist(1)(2));
         pr(' feels that he deserved this outcome whereas ');
         pr(parlist(1)(4));
         pr(' does not.');
      endif

  case = "agency" then
```

```
if item(2) = "you"
then
  pr(parlist(1)(2));
  pr(' feels that the outcome is caused by \n');
  pr(item(4));
  pr(' or some agency in this situation');
switchon(parlist(1)(1))
 case = "happy"
        then
          pr('Assume a positive outcome due to ');
          pr((parlist(1)(4)));
          pr(' or to circumstances and that ');
          pr((parlist(1)(1)));
          pr(' likes ');
          pr((parlist(1)(4)));
          pr('.');
 case = "angry"
        then
         if
           parlist(7)(2) = "d"
         then
           pr('Suggests that a positive outcome ');
           pr('was felt to be deserved and did not');
           pr(' materialize');
         endif;
 case = "surprise"
        then
          pr('Suspect that the outcome was unexpected and\n');
          pr(' expect ');
          pr((parlist(1)(2)));
          pr(' is to change emotion abruptly');

 case = "guilt"
        then
          pr('Suggests that a positive outcome was not deserved or\n');
          pr(' that a negative outcome was deserved. ');


endswitchon
endswitchon
endif
endfor
enddefine;
```

```
flavour Face_mop;

ivars name   Face_event_count = 0, Face_event_list = [];
ivars node = false, support = [1000 1000 1000 1000 1000 1000 1000 1000
1000
1000 1000 1000 1000 1000
1000 1000 1000 1000 1000 1000],
ivars
Face_sub_list = [],  Face_fsub_list = [];
ivars gf_a = [], ginterp = [],  brows  eyes nose interps path;
ivars common = [];
ivars interp event ;
defmethod printself;
   pr('<Face_mop_' >< name >< '>\n' )
enddefmethod;

defmethod support_attribute(attribute,update);
/*keeps check on frame f_a support: if support for a
frame f_a falls from 1000 to 0, the f_a is removed
and events indexed directly off the frame which
contain this f_a are re-indexed below the frame
using that f_a's labels as indexes.*/
lvars comlis lis tail each subindexlists;
vars attrib count a c q m fevlist subindexes submo;
vars frameindexes evlist reindex ev;

0 -> count;
[] -> evlist;
[] -> frameindexes;
[] -> reindex;
decompose(^gf_a)-> lis; ;;;the Face_MOP's f_a made a list of lists
decompose(^common)-> comlis; ;;;the Face_MOP's "common" made a list
of lists
 for attrib on lis do ;;;support is a number in a list in order of the f_a
   count + 1 -> count; ;;;advances the index in the list
   if
    attrib(1) = attribute ;;;finds the f_a to be updated
   then
    support(count) + update -> support(count); ;;;updates it
    if support(count) = 0 ;;; the check for removal
    then                ;;;removal
      delete(0,support)->support;
      count -1 -> count ;
      delete(attribute,lis)-> lis; ;;;from gf_a in MOP
      delete(attribute,comlis)-> comlis; ;;; from "common"
      compose(lis)-> self <- gf_a;
```

```
compose(comlis)-> self <- common;
;;;finds those events indexed directly off frame
;;; - these are the only ones possibly "at risk"
;;;They may contain the demoted f_a
;;;if they do, they are re-indexed by the demoted f-a and their direct
;;;indexing off the Frame is eliminated
;;;contains all events under this MOP
^Face_event_list -> fevlist;
[%
 for ev in fevlist do
   allbutfirst(3,ev)->ev;
   valof(ev) <- input -> evlist;
   decompose(evlist)->evlist; ;;; makes it a list of lists
  ;;;retrieves the f_a from each event
 if
  member(attribute,evlist)
 then
  ev
 endif
 endfor
%] -> reindex; ;;;l_o_lists those containing the "at risk" f_a
attribute -> road;
if null(reindex) then ;;;the branch is still made with no events
    t_add([^(hd(attribute))],[^(hd(attribute))],node);
    t_add(attribute,attribute,node);
else

      for ev in reindex do
      ;;; indexes them below the frame by the demoted labels
       t_add([^(hd(attribute))],[^(hd(attribute))],node);
       t_add(attribute,attribute,node);
       t_add([^^attribute event],[^ev],node);
endfor
endif;
   ;;;remove those directly off Frame only containing attribute
   ;;;these will be indexed by 'event' in the links field of the record
   if links(valof(node)) matches [ == event ?m == ]
   ;;;'m' is the node (record) where these are stored in the 'info' field

then               [into remove]=>
  for ev in info(valof(m)) do ;;;for each event thus found
   if
     valof(ev) <- input matches [^^attribute interp ?o]
     ;;;attribute alone
   then       ;;;delete this ev from the record
     delete(ev,(info(valof(m)))) -> info(valof(m));
```

```
            if present([?a [] ^m])
            then remove(it);
            delete(ev,a) ->a; ;;; remove it from the database
            add([^a [] ^m])
            endif
          endif
        endfor;
      endif;
      endif
    endif
  endfor  ;;;all done
enddefmethod;
endflavour;


;;; ----------------------------------------
flavour fsub_mop isa Face_mop;
ivars name Spec  fsub_sub_list  = [], path  fsub_event_count = 0;
ivars fsub_event_list = [], node = false,  common;
defmethod printself;
   pr('<fsub_mop_' >< name >< '>\n')
enddefmethod; ;;;a procedure foroutputting the name of an instance

defmethod after initialise;
;;;a  demon for updating the FaceMOP's list of fsubMOPs is automatically
;;;triggered when an instance is instantiated
unless Spec = false or Spec = undef
then
 ^name :: (valof(Spec) <- Face_fsub_list) -> valof(Spec) <- Face_fsub_list;
endunless
enddefmethod;
endflavour;


;;; ----------------------------------


flavour sub_mop isa fsub_mop;
ivars name Spec fSub path  sub_event_count = 0;
ivars sub_event_list = [], node = false,
ivars gf_a = [], ginterp = [],common = [] ;
defmethod printself;
   pr('<sub_mop_' >< name >< '>\n')
enddefmethod; ;;;a procedure foroutputting the name of an instance
```

```
defmethod subb;
;;;a new subMOP's name is added to its FaceMOP's list of subMOPs
;;;and its fsubMOP's list of subMOPs :
lvars v;
unless Spec = undef or Spec = false   then

 (^name :: (valof(Spec) <- Face_sub_list)) -> (valof(Spec) <-
Face_sub_list);
endunless;
unless fSub = undef or fSub = false then
 (^name :: (valof(fSub) <- fsub_sub_list)) -> (valof(fSub) <- fsub_sub_list);
endunless;
enddefmethod;
endflavour;

;;; ——————————————————————————————
flavour evvent isa sub_mop ;
ivars name Spec = false, fSub = false,  Sub = false, path node Face_match
= false,
sub_match = false;
ivars  Face_support = [20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20
20 20 20 20 20 20], paths
sub_support =[20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20
20];

defmethod printself;
   pr('<evvent_' >< name >< '>\n')
enddefmethod; ;;;a procedure foroutputting the name of an instance

defmethod before initialise( i ) -> i ;
;;;'i' is a list of ivars instantiations for the particular input event.
;;;prints details of the storage of the input face event
pr('*****************************************************************');
pr('               EVENT REGISTERED IN LONG TERM MEMORY            \n');
pr('*****************************************************************\pr(\n');
pr('This face description has been registered under the label : \n');
pr(i(2));
pr(' at node ');
i --> [ == node ?nod == ];
pr(nod);
pr('\n\n');
pr('Its total description is as follows :\n');
pr(i);
pr('\n');
pr('This same description will be registered in memory in several different');
pr(' places which are accessed by the different items in it. These are : \n');
```

```
pr(i(6));
pr('\n ***************************************************************\n');
enddefmethod ;

defmethod after initialise;
;;;the number of events indexed under a submop is monitored : if it exceeds
;;;2/3 of those under the parent Facemop the submop is moved into the
 ;;;content frame of the Facemop. This does not apply to the first 6 events .
vars mopp aspec  collapse_sub  remove_gen  a h;
vars nn mm p q Lo content submo subs contents;
if ((( Sub = false) or  (Sub = undef)) and ((Spec /= false) and (Spec /= un-
def)))
then
valof(Spec) -> aspec;
aspec <- Face_event_list ->Lo;
   unless member(name,Lo) then ;;;name of event added to FaceMOP's list
        (name :: Lo) -> (aspec <- Face_event_list);
        ;;;remove any duplications of events.by procedure 'calc'
        calc(aspec <- Face_event_list) -> aspec <- Face_event_list;
        endunless; ;;;and the count updated
        length(aspec <- Face_event_list) -> (aspec <- Face_event_count);

elseif (Spec /= false and Spec /= undef) and ((Sub /= false)
and (Sub /= undef))
;;;else if indexed below a subMOP, the list & count of it are updated
then valof(Sub) -> mopp;
   valof(Spec) -> aspec;
(mopp <- sub_event_list) ->Lo;
   unless member(name,Lo) then
        (name :: Lo) -> mopp <- sub_event_list;
        ;;;remove any duplications of events by procedure 'calc'
        calc(mopp<- sub_event_list) -> mopp <-sub_event_list;
        length(mopp <- sub_event_list) -> mopp <- sub_event_count;
        aspec <- Face_event_list ->Lo;
        unless member(^name,Lo) then
        ;;;event's FaceMOP list & count updated
        (name :: Lo) -> aspec <- Face_event_list;
        (length(aspec <- Face_event_list)) -> (aspec <- Face_event_count);
   endunless;
unless ^fSub = undef  or ^fSub = false
then
  ;;;event's fsub_MOP list & count updated
        valof(fSub)->efsub;
        (efsub <- Face_event_list) ->Lo;
        unless member(name,Lo) then
        (name :: Lo) -> (efsub  <- fsub_event_list);
```

```
        (length(efsub <- fsub_event_list)) -> (efsub <- fsub_event_count);
    endunless;
    endunless;
    endunless;
length(info(valof(node))) -> event_count;
info(valof(node)) -> event_list;
calc(aspec("Face_event_list")) -> aspec("Face_event_list");
(length(aspec <- Face_event_list)) -> (aspec <- Face_event_count);
calc(event_list) -> mopp("sub_event_list");
 length(mopp <- sub_event_list) -> mopp <- sub_event_count
;;;for promotion, the events indexed below a subMOP(in event_count) must
;;;be more than 6 in number and also account for more than 2/3rds of those
;;;indexed under the FaceMOP;
if (event_count  > 6 and aspec("Face_event_count") /= 0
  and  ((event_count / aspec("Face_event_count"))
   > (2/3) ))
then
   pr('            PROMOTING SUBMOP GENERALISATION :\n');
   pr('\n');
   pr('            moving the sub_mop frame content:\n');
   pr('*********************** ');
   pr(mopp <- common);
   pr('*********************** \n');
   pr('***********************into the SPEC_MOP : ');
   pr(Spec);
   pr(' ********************\n');
   pr('                                                  \n');
   pr('*************************************************************\n');
;;;the promoted face action is added to the Frame contents
   [^^(mopp <- common) ^^(aspec <- common)] -> ivalof(aspec,"common") ;
   pr('                   MOVE COMPLETED:\n\n');
   pr('*************************************************************\n');
;;;the promoted face action is given a support of '20'
   [20  ^^(aspec <- support)] -> ivalof(aspec,"support") ;
   [0] -> sub_support;
   [^^(mopp <- sub_event_list) ^^(aspec <- Face_event_list)]
                                -> aspec <- Face_event_list;
   calc(aspec("Face_event_list")) -> aspec("Face_event_list");
   (length(aspec <- Face_event_list)) -> (aspec <- Face_event_count);
   pr('*************************************************************\n');
   pr('   SUBMOPS EVENTS TRANSFERRED TO THE PARENT
Face_MOP\n');
   pr('*************************************************************\n'
 );
   [^^(mopp <- gf_a) ^^(aspec <- gf_a)]      -> aspec <- gf_a ;
```

```
collapse_sub([^Spec],[^Sub],aspec,mopp);
;;;call to procedure for collapsing the subMOP
undef -> latestmop; ;;;the subMOP is not now the latestmop
undef->mopp;
endif;
endif;

false -> facet;
enddefmethod;
endflavour;


define collapse_sub(jj,kk,j,k);
pr('********************** COLLAPSING SUBMOP **********************');
;;;removes the subnode when its gen is moved to Face_mop
vars  efmop lspec lfsub lsub liz;
vars  ss dd newspec subnode aa bb ind nod slot newspec newinst fsubn-
ode;
vars Facenode;

pr('\n COLLAPSING  SUBMOP : ');
pr(k);  ;;;names the subMOP
pr('\n');
        k <- fSub -> efmop;  ;;;the fsubMOP
        valof(efmop) <- node -> fsubnode;  ;;;fsubMOP's node
        j <- node -> Facenode;  ;;;FaceMOP's node
        if present([[^efmop] ?ss ^fsubnode])  ;;;if fsubMOP record in db
        ;;;removes it temp to alter link to sub
            then remove(it);  ;;;'ss' is a list of links to subnodes
        endif;
   ivalof(k,"node") -> subnode; ;;;submops node is found
      [] -> info(valof(subnode)); ;;;record's info field made an empty list
      if ss matches [??aa ?a ^subnode ??bb ] ;;;subnode's  index is cancelled
      then    [^^aa ^^bb] -> ss;  ;;;revised index list is replaced in db
      add([[^efmop] ^ss ^fsubnode]);
      endif;
   if present([?c ?h ^subnode])  ;;;subnode's db entry goes
   then
     remove( it ); endif;

   if present([^jj ?liz ^Facenode])
   then
      ;;;deleted subnode's daughters added to FaceMOP's db record's list of
      ;;;daughters
            remove(it);
```

```
      add([^jj ^liz ^Facenode]);
   endif;
  links(valof(Facenode))-> lspec;
  links(valof(subnode))-> lsub;
  links(valof(fsubnode))-> lfsub;
 [^^lsub ^^lspec] -> links(valof(Facenode));
        ss -> links(valof(fsubnode));
        [] -> links( valof(subnode));
        undef -> subnode;
        undef ->k;

enddefine;
```

## A.3 POP11 code for the learn and retrieve functions:

This codes the dynamic memory and processes inputs face knowledge in two modes: 'Learn' and 'Retrieve'. Input is entered in 'Learn' mode as a parameter to procedure 'insert' if it is to be learned; and to procedure 'traverse', if an interpretation is to be retrieved. This version of Janus in 'Retrieve' mode, accepts as input a syntactic representation of a face expression and interprets it as an emotional label. It uses a primed database which learned the interpretations of 50 persons judging face photographs. The input is restricted to brow, eyes, nose, mouth, cheeks, jaw, and their values. It tries to locate the most likely FaceMOP by counting the number of face actions which the content frames of candidate MOPs have in common with the input expressing this as a proportion of the total in the frame, and traversing the mop with the greatest proportion. Should a tie occur, a procedure "findbestmop" applies a heuristic to the input to find the best mop and should this also produce a tie, a random choice is made from the finalists.

/*....*/ encloses a comment

';;;' prefixes a comment.

```
uses file; ;;;a library function
load oflav.p; ;;;loads the FLAVOUR hierarchy
;;;declared variables
vars face_affect_display feature act z zz isnode links facelist facts sub cn-
ode entry lproc o lastmop facet road support success = false, latestmop
mopp fmopp, Face latermop routes remop target fa emot traverse_mop gg
res insert recap = false, arc listp frame_grab mop_handle com-
pare_gentarg icn compare_list greatest t_add mopem t_isin entr inf emotion
nodename learn_from_mistakes traverse_mop mop_disgusted_gen_face
mop_afraid_gen_face top1 m2 mop_angry_gen_face
mop_surprised_gen_face m3 m4 m5 mop_sad_gen_face vars
mop_happy_gen_face facemop m6 m7 search_mop new_indexes
new_mopnode bestmop vent facenode face show_tree face path sub a
vars b sub1 list1 list2 e1 d2 m1;
;;;the 6 following numbers = the number of face actions in the Face_mop
Frames:
vars happysum = 9.0, vars sadsum = 7.0; vars disgustsum = 17.0; vars
afraidsum = 14.0; vars angrysum = 17.0; vars surprisesum = 8.0;
;;;Face_mops and their nodes in a list:
[%
[m13 mop_angry_gen_Face] ,
[m12 mop_afraid_gen_Face] ,
[m11 mop_disgusted_gen_Face],
[m10 mop_sad_gen_Face],
[m8 mop_happy_gen_Face],
[m7 mop_surprised_gen_Face]
```

%] -> facemop;

/*Below the.6 instances of the "inborn" FaceMOPs are made. The parameter passed in the 'make_instance' command is a list made up of the MOP class i.e. Face_MOP, followed by the name, followed by a number of slots(in bold type in the first instance) and values and each instance is assigned to the name. The name of the instance is a word; strings are converted thus: */

```
consword('mop_happy_gen_face') -> mop_happy_gen_face;
make_instance( [Face_mop name mop_happy_gen_face gf_a [cheeks
raised  mouth bared mouth up  mouth open mouth u_lip_raised mouth slight-
ly_open mouth wide mouth pulled eyes l_lid_raised ] interp happy common
[cheeks raised  mouth bared mouth up  mouth open mouth slightly_open
mouth wide mouth pulled eyes l_lid_raised interp happy] subevlist [] subev-
count 0 Face_sub_list [] Face_fsub_list [] Face_event_list [] fSub undef
Sub undef Face_event_count 0 support [1000 1000 1000 1000 1000 1000
1000 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000
1000] interps [pleased cheerful calm non_aggressive loving mesmerized
captivated good_humoured pleasant_to_talk_to interested not_angry
friendly good_humoured  amused open relaxed pleasant not_serious
weighing_up thoughtful] node m8]
) -> mop_happy_gen_face;

consword('mop_sad_gen_face') -> mop_sad_gen_face;
make_instance( [Face_mop name mop_sad_gen_face gf_a [brows con-
tracted brows in_raised brows centre_raised eyes l_lid_raised eyes in-
lid_raised eyes down mouth down] interp sad common [brows contracted
brows in_raised eyes l_lid_raised  brows centre_raised eyes inlid_raised
eyes down mouth down interp sad] Face_event_count 0 Face_sub_list []
Face_fsub_list [] Face_event_list [] fSub undef Sub undef subevlist [] sub-
evcount 0 support [1000 1 1000 1000 1000 1000 1000 1000 1000 1000
1000 1000 1000 1000 1000 1000 1000 1000 1000 1000] interps
[disinterested sad not_wanting_anyone_to_be_here bored assessing
bit_fed_up angry anguished]    node m10 ]
) -> mop_sad_gen_face;
consword('mop_disgusted_gen_face')
-> mop_disgusted_gen_face;

make_instance( [Face_mop name mop_disgusted_gen_face gf_a [brows
low nose screwed mouth l_lip_everted mouth u_lip_everted mouth bared
mouth u_lip_raised mouth l_lip_lowered nose flared mouth l_lip_raised
cheeks n_l_vert cheeks raised eyes l_lid_raised eyes narrowed mouth
open mouth compressed mouth l_lip_tensed mouth u_lip_tensed] interp dis-
gusted common [brows low nose screwed mouth l_lip_everted mouth
u_lip_everted mouth bared mouth u_lip_raised mouth l_lip_lowered nose
```

```
flared mouth l_lip_raised cheeks n_l_vert cheeks raised eyes l_lid_raised
eyes narrowed mouth open mouth u_lip_tensed mouth l_lip_tensed interp
disgusted] Face_sub_list [] Face_fsub_list [] Face_event_list [] fSub undef
Sub undef Face_event_count 0 subevlist [] subevcount 0 interps [bored
disbelieving depressed mesmerized deliberating displeased disdainful con-
temptuous scornful] support [1000 1000 1000 1000 1000 1000 1000 1000
1000 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000] node
m11]
) -> mop_disgusted_gen_face;

consword('mop_afraid_gen_face') -> mop_afraid_gen_face;
make_instance( [Face_mop name mop_afraid_gen_face gf_a [brows raised
brows in_raised brows contracted eyes l_lid_tensed eyes wide eyes
u_lid_raised eyes l_lid_raised eyes inlid_raised mouth wide mouth open
mouth pulled mouth u_lip_tensed mouth l_lip_tensed mouth bared] interp
afraid common [brows raised brows in_raised brows contracted eyes
l_lid_tensed eyes u_lid_raised eyes l_lid_raised eyes inlid_raised mouth
wide mouth open mouth pulled mouth u_lip_tensed mouth l_lip_tensed
mouth bared interp afraid] Face_sub_list [] Face_fsub_list []
Face_event_list [] fSub undef Sub undef Face_event_count 0 subevlist []
subevcount 0 support [1000 1000 1000 1000 1000 1000 1000 1000 1000
1000 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000] interps
[fearful] node m12]
) -> mop_afraid_gen_face;
consword('mop_angry_gen_face') -> mop_angry_gen_face;
make_instance( [Face_mop name mop_angry_gen_face gf_a [brows low
brows contracted eyes narrowed eyes inlid_raised eyes l_lid_raised eyes
l_lid_tensed eyes u_lid_lowered eyes u_lid_tensed nose flared mouth com-
pressed mouth l_lip_tensed mouth u_lip_tensed mouth bared mouth open
mouth wide mouth square cheeks n_l_vert ] interp angry common [brows
low  brows contracted eyes narrowed eyes inlid_raised nose flared eyes
l_lid_raised eyes l_lid_tensed eyes u_lid_lowered eyes u_lid_tensed
mouth compressed mouth bared mouth open mouth wide mouth square in-
terp angry] Face_sub_list [] Face_fsub_list [] Face_event_list [] fSub undef
Sub undef Face_event_count 0 subevlist [] subevcount 0 support [1000
1000 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000
1000 1000 1000 1000 1000 1000] interps [angry strong_feeling
expressing_an_opinion committed self_absorbed disgusted retiring_to_bed
happy amused angry] node m13 ]
) -> mop_angry_gen_face;
consword('mop_surprised_gen_face') -> mop_surprised_gen_face;
make_instance( [Face_mop name mop_surprised_gen_face gf_a [brows
raised eyes wide eyes u_lid_raised eyes l_lid_lowered mouth open jaw
drop mouth slightly_open mouth wide] interp surprised common [brows
raised eyes wide eyes u_lid_raised eyes l_lid_lowered mouth open jaw
drop mouth slightly_open mouth wide interp surprised] Face_sub_list []
```

```
Face_fsub_list [] Face_event_list [] fSub undef Sub undef
Face_event_count 0 subevlist [] subevcount 0 support [1000 1000 1000
1000 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000
1000 1000 1000 1000] interps [puzzled not_showing_boredom some-
thing_catches_his_attention questioning disbelieving] node m7]
) -> mop_surprised_gen_face;
[brows eyes nose mouth cheeks jaw]-> face;
false -> facet;
[] -> facelist;
[] ->routes;
[] ->path;
[] ->facts;
recordclass node info links ;  ;;; makes a record 'node' ;;;with fields

define maketree1;   ;;; converts db into node-records
load /usr/u3/garry/demo_db.p; ;;;loads the db.
vars line inf links e x y f ans;
for line in database do
;;;db format: [...[[info][links][node]]...]

if line matches [?inf ?links ?e]
then
;;;form the record
consnode( inf,links ) ->valof(e);
endif;
endfor; 14 -> gensym( "m" );
;;;on each call makes m14 m15 m16 m17 m18 etc
0 -> gensym("inst");  ;;;makes inst1, inst2 etc.
1000 -> gensym("eev"); ;;;makes eev1000, eev1001 etc.
0 ->gensym("ev");     ;;;makes ev0, ev1 etc.
enddefine ;

define nomop(item);
;;;'item' must not begin with 'mo' e.g.'mop'
substring(1,2,item) /= 'mo';
enddefine;

define evtec(i); ;;; to detect 'ev0', 'ev1' etc.
substring(1,2,i) = 'ev';
enddefine;

define nevtec(i);
;;; to exclude 'ev0', 'ev1' etc.
substring(1,2,i) /= 'ev';
enddefine;
define twoplus(L); length(L) > 2; enddefine;
```

```
define feat(g) -> result;
 member(g,face) -> result;
enddefine;

define decompose(rlist) -> rList_of_lists;
/*a list is converted to a list of lists of consecutive pairs*/
lvars L;
[%
for L on rlist do
if length( L ) mod 2 = 0 and length( L ) > 0
then
[% hd( L ) ;  hd(tl( L )) ; %]
endif
   endfor
%] -> rList_of_lists;
enddefine;

define compose(rlist_of_lists) -> rList;
 /*a list of lists is converted to a list*/
lvars L rlist_of_lists;

 [%
unless null(rlist_of_lists)
then
for L in  rlist_of_lists do
explode(L)
endfor;
endunless
%] -> rList;
 enddefine;

define find_indices_from_target(listx,listy) -> targ_indexes;
/*finds anomalies twixt an input and a MOP Frame in terms of face actions;
listx contains the matches and listy, a copy of the input */
 lvars listx listy component;
 for component in listx do
;;;each matched f-a removed from input-list
delete(component,listy) -> listy; ;;;anomalies are left. endfor;
listy -> targ_indexes;
enddefine;

define bullseye(fringenode,Inf,Facenode);
      /*The result of traversing a MOP in RETRIEVE mode, a leafn-
ode(fringenode) has been located. Its face-event contents are event la-
bel(s) in 'Inf'. The interpretations in the input corresponding to these labels
is extracted and output to the user.*/
lvars fringenode Inf Facenode howmany termss;
```

```
vars view pp idea scan ans gen lastgen box newinterp q;
vars capital x y label tax mood nam term anterms evv;
if (length(Inf) = 1 or [false] = Inf)              then
"one"
else
"many"                                    endif -> howmany;
pr('********************TARGET   pr('IDENTIFIED************************');
pr('\n');     unless
Inf = [false]
then
if
howmany = "one"
then
pr('A SIMILAR TARGET FACE-DESCRIPTION HAS BEEN');
pr('IDENTIFIED ');
else
pr('SIMILAR TARGET FACE-DESCRIPTIONS HAVE BEEN');
pr('IDENTIFIED ');
endif;
pr('LABELLED \n ');
pr('               ');
pr(Inf);
pr('\n');
pr('\n indexed under the general conceptual category ');
pr(             lastmop); ;;;(the FaceMOP traversed)
pr('\n WITH  \n');
pr('           THE GENERAL INTERPRETATION : \n' );
pr('                 ');
valof(lastmop) <- interp -> idea;
;;;(the FaceMOP's interpretation)
pr(idea);
pr('\n');
pr(' The interpretation for this input is therefore\n');
pr('\n');
[%
for evv in Inf do
;;; for each stored event in the leaf, the
valof(evv)<-input --> [ == ?mood];
;;;interpretation is extracted.
mood; endfor           ;;;(may be more than one)
%] -> termss;
[%
until   ;;; removes duplications of terms
null(termss) do
hd(termss)-> nam;
      delete(nam,termss)->termss;
```

```
nam
enduntil
%] -> anterms;
for tax in anterms do
pr('                 ');
pr(tax);  ;;;prints them.
pr('\n');
endfor;
else ;;;if the leaf is empty the FaceMOP Frame
    ;;;interpretation is printed
pr('***********************************************\n');
pr('THE INTERPRETATION OFFERED FOR THIS EXPRESSION');
pr(' IS:\n');
valof(lastmop) <- interp -> idea;
pr('                        ');
pr(idea);
pr('                  \n');
pr('\n');
pr('              \n');
pr('SESSION COMPLETED');
pr('              \n');
pr('***********************************************\n');
true -> success;
return();
endunless
enddefine;
```

**define target_print(Inf,recentnode,entry,Facenode);**

```
/*having traversed a MOP branch, a face description in search of an interpretation
the system has to discover what content-value the last node encountered has.
Depending on the content, the information is either sent off to the 'Bullseye' pro-
cedure or a deeper traversal is made.*/
lvars Inf nodename jay;
vars lot res1 item newres c xinf y f leafnode goal v; unless
(Inf = [false]) or null(Inf) then
if
length(Inf) = 2 ;;;(the node content is a face action)
then
t_isin([event],recentnode) -> Inf -> leafnode;
;;;onto the leaf
unless
Inf = [false] ;;;(not there)
then
bullseye(leafnode,Inf,Facenode);
endunless
elseif
```

```
(length(Inf) = 1 and (substring(1,2,(hd(Inf))) /=
 'ev')) ;;;a MOP-label
then
t_isin([event],recentnode) -> Inf -> leafnode;
;;;on to the leaf
if
substring(1,2,(hd(Inf))) = 'ev'
then
bullseye(leafnode,Inf,Facenode);
endif
elseif
(length(Inf) = 1 and (substring(1,2,(hd(Inf))) =
 'ev')) ;;; a leaf
then
bullseye(recentnode,Inf,Facenode);
endif
endunless
enddefine;

define findbestmop(listing) -> bestmop ;
/*It tries to locate the most likely FaceMOP in retrieval by a heuristic: match-
ing the face actions (input in search of an interpretation) to some salient
configurations for the 6 primary emotions the input to find the best mop and
should this also produce a tie, a random choice is made from the finalists.
'listing' is the input face description as a list of face actions. */
 undef -> bestmop;
if
(
(listing matches [ == eyes l_lid_raised == ] = false)
and
(listing matches [ == brows raised == ])
)then
"mop_surprised_gen_Face" -> bestmop
elseif
(
(listing matches [ == eyes l_lid_tensed == ] and
listing matches [ == eyes l_lid_raised == ]) and
(listing matches [ == brows raised == ])
)then
"mop_afraid_gen_Face" -> bestmop

elseif
(
(listing matches [ == mouth up == ]) and not
( listing matches [ == nose screwed == ])
)then
```

```
"mop_happy_gen_Face" -> bestmop
elseif
(
(
listing matches [ == brows low == ] and
listing matches [ == brows contracted == ]
) and
listing matches [ == eyes inlid_raised == ] and
(
(
listing matches [ == mouth compressed == ] or
listing matches [ == mouth wide == ]
) and
not(listing matches [ == mouth u_lip_raised == ]) = true
)
)then
"mop_angry_gen_Face" -> bestmop
elseif
(
(
listing matches [ == mouth u_lip_raised == ] and
listing matches [ == mouth u_lip_tensed == ]
) and
(
listing matches [ == mouth l_lip_raised == ] or
listing matches [ == mouth l_lip_lowered == ]
) or
(
listing matches [ == nose screwed == ] and
listing matches [ == cheeks raised == ] and
listing matches [ == eyes l_lid_raised == ] and
listing matches [ == brows low == ]
) or
(
listing matches [ == nose screwed == ] and
listing matches [ == cheeks raised == ] and
listing matches [ == eyes l_lid_raised == ] and
listing matches [ == mouth u_lip_raised == ]
) or
(
listing matches [ == mouth u_lip_raised == ] and
listing matches [ == mouth u_lip_everted == ] and
listing matches [ == cheeks raised == ] and
listing matches [ == nose screwed == ]
) or
(
```

```
listing matches [ == mouth u_lip_raised == ] and
(
listing matches [ == mouth l_lip_raised == ] or
listing matches [ == mouth l_lip_lowered == ]
) and
(
listing matches [ == nose screwed == ] or
listing matches [ == cheeks n_l_vert ==]
)
)
)then
"mop_disgusted_gen_Face" -> bestmop
elseif
(
(
listing matches [ == brows centre_raised == ] and
listing matches [ == eyes inlid_raised == ] and
listing matches [ == eyes l_lid_raised == ]
) or
(listing matches [ == mouth down == ]) or
(listing matches [ == brows centre_raised  == ])
)then "mop_sad_gen_Face" -> bestmop
endif
enddefine;

define traverse(entr,nodename);
/*enters a list of user-supplied face actions(in entr) in search of interpreta-
tion and controls the search procedure. 'nodename' is the root node in the
memory tree from which the search starts*/
lvars  answer len moppa matchers okinterp;
vars monad Inf  a b c  result dubb search Inf_list each ss tt pubb listp node
Inf j jj;
vars  item2 trash1 trash2; false ->success;
/*sets a flag that the interpretation has not been found*/
"trav" -> lproc;
[] -> len;
[] -> res;
[undef] -> Inf;
gensym("eev") -> vent; ;;;given a unique ID
pr('The face-description entered has been given the ');
pr('EVENT NUMBER : ');
pr(vent);
pr('\n');
[^^entr event ^vent] -> entry; ;;;combines ID with input. The ID is made
 /*the name of an instance of a FLAVOUR from which the input is retrieva-
ble.*/ make_instance([episode  input ^entr ])  -> valof(vent);
```

```
[]-> Inf_list;
[^^entr] -> target;
[%
for tt on target do
if
  length(tt) mod 2 = 0 and length(tt) > 0
then
  [% hd(tt),hd(tl(tt))%]
endif
endfor %] -> listp;  /*makes the input into a l-o-l of face action(f-a)
pairs*/
[] -> res;
for each in facemop do  /*a list of FaceMOPs and their node names*/
frame_grab(each(1),each(2),target) -> res; /*extracts MOP f-a and com-
pares them with input f-a to find matching f-a.*/
 endfor;
[% for moppa in res do
 ;;;res contains Mop matches for the six FaceMOPs
explode(moppa(1)) endfor; %] -> matchers;
unless
null(matchers) then
mop_handle(res,o,target);
;;;finds the MOP to traverse and controls the retrieval search
if success = true then
pr('\n                ORACLE: ');
pr(EMOTION);  ;;;prints the interpretation
pr('                or possibly  ');
pr('you have another interpretation? - if so type');
pr(' it now in one word:');
readline()->EMOTION;
unless null(EMOTION) do
learn_from_mistakes(EMOTION) -> okinterp;
pr('Thank you. I have learned that now. Goodbye');
endunless
endif; false -> recap;
else
pr('There are insufficient face-actions input to ');
pr('come to a conclusion in this case. It is thus ');
pr(' interpreted as "undifferentiated" or "neutral"');
endunless;
popval([file database in 'mydata1.p';]);
enddefine;

define trace_index(indexes,mopnode);
/*traverses branches of the memory tree with supplied indexes.*/
vars path  endnode nextnode lastnode;
```

```
indexes -> new_indexes;
mopnode -> new_mopnode;
if indexes = [] ;;;all input comsumed by FaceMOP Frame
then
t_isin([event],mopnode) -> Inf ->endnode;
/*returns leafnode content (directly off Frame) using index 'event'*/
 bullseye(endnode,Inf,lastmop)
 ;;; sends it to procedure 'bullseye'.
else
for path in indexes do
;;;using supplied indexes in turn
unless
success
then
t_isin([^(path(1))],mopnode) -> Inf ->endnode;
/*traverses branches of memory and returns the fourth
 node-content.*/
unless
Inf = [false] or Inf = [] ;;;unless end node is empty
then
;;;according to content, retrieves the node content
;;;and sends it to procedure 'target_print'.
if
substring(1,3,(hd(Inf))) = 'mop'
then
search_mop((hd(Inf)),endnode,path)-> Inf -> nextnode;
target_print(Inf,nextnode,entry,mopnode);
else
t_isin(([^(path(2))]),endnode)-> Inf -> lastnode;
target_print(Inf,lastnode,entry,mopnode)
endif
endunless
endunless
endfor
endif
enddefine;

define traverse_mop(mop,listp,target,route,mopnode);
vars index path findindices vent p lentry ;
lvars mop target node o ppp route result endnode i dtarget gen component
res2 res1 re re1;
mop -> lastmop route(1) -> road;
pr('Traversing the MOP by way of indices in search of');
pr('an event identical to the input');
find_indices_from_target(route,listp) -> indexes;
/*input face actions anomalous to those in the chosen Facemop frame will
```

```
be used to open existing paths through the branches:*/
trace_index(indexes,mopnode);
if success = false ;;;if no interpretation found, try 'event' as an index.
then
t_isin([event],mopnode) -> Inf ->endnode;
;;;retrieves a leafnode content
 bullseye(endnode,Inf,lastmop)
;;;sends it to bullseye
endif;
[] -> new_indexes;
enddefine;

define tester(str,wd);/*tests word has a certain ending*/
vars d;
wd  >< ''-> d;
isendstring(str,d);
enddefine;

define search_mop( content,node,path ) -> Inf -> endnode;
 /*tests a discovered MOP for its class so to set path variables */
unless tester('fsub',content) = false
;;;unless its not a fsubMOP
then
  content -> latermop;
  t_isin([^(path(2))],node) -> Inf -> endnode;
  ;;;retrieves 5th level node-content
  if null(Inf) or Inf = [false]
  then return();
  else
 unless tester('sub',Inf(1)) = false
 ;;;unless its not a subMOP
 then
  hd(Inf) -> latestmop;
 endunless
  endif;
endunless
enddefine;

define mop_choice(mop,listp,target,o,route,node) -> Inf - >nodename;
/*controls the storage of an input face description in LEARN mode once the fittest
FaceMOP is chosen, the anomalous face actions either open existing links in the
tree below the Frame or create them.*/
vars index path nodename findindices vent r road matchers
lvars   result listt;
entry --> [ == ?vent];
listp -> matchers;
```

```
find_indices_from_target(route,matchers) -> indexes;
indexes -> matchers;
valof(mop) <- gf_a  -> listt;  ;;;Frame f-a are made into a l-o-lists.
[%
for tt on listt do
if length(tt) mod 2 = 0 and length(tt) > 0
then
[% hd(tt),hd(tl(tt))%]
endif
endfor
%] -> listt;
unless null(indexes)
then
for path in indexes do
;;;node by node descends the tree to leaf
unless ((member(path,listt)) or null(route))
then   path -> sub1;
[^^(hd(route)) ^(hd(path))] -> road;
t_add([%(hd(path))%],[%(hd(path))%],node);
[^^(hd(route)) ^^path] -> road;
t_add(path,path,node);
[^^(hd(route)) ^^path event] -> road;
t_add([^^path event],[^vent],node);
endunless endfor
else    ;;;if there are no indexes index input directly off Frame
[^^(hd(route))  event] -> road;
t_add([event],[^vent],node);
endunless;
true ->success;
return();
enddefine;

define frame_grab(nodum,mop,target) -> res;
  /*retrieves face(f-a) actions from the Frame of a FaceMOP*/
vars gen rest;
valof(mop) <- gf_a ->gen;
compare_list(gen,target) -> rest;
;;;extracts f-a matches in the Frame and input
[%rest, nodum, mop %] :: res -> res;
enddefine;

define mop_handle(LISt,lisT,target);
/*decides which FaceMOP the input will select and determines its further
processing according to whether the system is to learn or retrieve.*/
vars c LIS result each;
false -> remop;
unless  success = true then greatest(LISt) -> result;
```

```
;;;procedure to select a FaceMOP
(valof(result(3)) <- interp) -> EMOTION;
undef -> Face; undef -> fmopp; undef-> mopp;
undef ->latermop;
undef ->latestmop;
result(3) -> lastmop;
if lproc = "inser"  ;;;if input is to be learned:
then
mop_choice(result(3),listp,target,o,result(1),result(2))
-> Inf -> nodename;
else  ;;;is input is to be interpreted:
traverse_mop(result(3),listp,target,result(1),result(2));
endif;
endunless
enddefine;

define insert(entr,nodename);
/*procedure for inputting a face description
and interpretation in LEARN mode*/
lvars  nodename moppa matchers;
vars monad Inf v ss a b c result len each;
vars ss tt dubb listp node Inf res facs m;
gensym("ev") -> v; ;;;input given a unique ID
[^^entr event ^v] -> entry;
false -> success;
"inser" -> lproc;    ;;;denotes Learn mode
[] -> len;
[]->res;
[undef] -> Inf;
entry --> [??ss event ^v];
;;;'ss' contains all input list content
make_instance([episode input ^ss ]) -> valof(v);
ss --> [??facs interp ?m]; ;;;'facs' contains all face actions
facs -> target;  ;;;assigns them to target
[%
for tt on entry do
;;;puts every second member of the input in a list
if
length(tt) mod 2 = 1 and length(tt) > 0
then
  hd(tt);
endif
endfor %] -> o;
[%
for tt on target do
```

```
if
length(tt) mod 2 = 0 and length(tt) > 0
then
  [% hd(tt),hd(tl(tt))%]
endif
endfor
%] -> listp; ;;;makes each face action pair a list in a list
[] -> res;
for each in facemop do ;;;
each = [node FaceMOP]
frame_grab(each(1),each(2),target) -> res;
endfor;
[%
for moppa in res  do
  ;;;moppa = [[matched f-a][node][FaceMOP]]
explode(moppa(1))
endfor;
%] -> matchers; ;;;makes a list of all Frame-matched face actions
unless null(matchers) then mop_handle(res,o,target
);
endunless;
popval([file database in 'mydata1.p';]);
sysprwarning -> prwarning;

enddefine;

define olist(g) -> resy;
member(g,o) -> resy; ;;;o is a list of input actions(face)
enddefine;

define mophead(mops) -> mopp -> fmopp ->Face;
 /*sorts a list of branch-MOPs by their class into variables*/
lvars mops title for mop in mops do
allbutfirst(length(mop >< '') - 4, mop >< '') -> title;
switchon title
case = 'Face' then mop -> Face;
case = '_sub' then mop -> mopp;
case = 'fsub' then mop -> fmopp;
endswitchon endfor;
enddefine;

define compare_list( List1, List2 )-> List3;
;;;lists1&2 will have f-a which must be matched in pairs vars x y;
[] -> list3;
   [%
        for  x  from 1  by 2 to (length(List1) -1) do
```

```
if ((member(List1(x),List2)) and
(List2  matches [== ^(List1(x)) ^(List1(x+1)) == ]))
 then
 [%List1(x);List1(x+1)%]
endif endfor
%] ->List3;     ;;;a list containing matched face actions.

 enddefine;

define calc(k)-> m;
;;;removes duplicates from a list
k->n;
[% for t in k do
if member(t,n) then delete(t,n)->n;
delete(t,k)->k;
;;;'delete' removes all instances
t   ;;;puts t on stack
endif
endfor;
%]->o; ;;;gathers stack in list
[^^o ^^n]->m
enddefine;

define subtract(k,n)-> m; ;;;subtracts one list from another
vars t;
for t in k do
if member(t,k) then delete(t,n)->n;
endif
endfor;
n->m;
enddefine;

define whatsum(anymop) -> thesum;

 /*the variable "thesum" takes different value depending on the MOP*/
switchon anymop
case = "mop_happy_gen_Face" then happysum -> thesum;
case = "mop_sad_gen_Face" then sadsum -> thesum;
case = "mop_disgusted_gen_Face" then digustsum ->
      thesum;
case = "mop_afraid_gen_Face" then afraidsum -> thesum;
case = "mop_angry_gen_Face" then angrysum -> thesum;
case = "mop_surprised_gen_Face" then surprisesum ->
      thesum;
endswitchon
enddefine;
```

```
define greatest(list4) -> longest;
/*the proportion of input f-a which match each FaceMOP f-a is
the basis of comparison used in deciding which FaceMOP should
 be traversed. Should a tie occur, a procedure "findbestmop" applies
a heuristic to the input to find the best mop and should this also produce
a tie, a random choice is made from the finalists.*/

lvars longest sum succ part unit;
vars x y biggest last_biggest bestbet;
1000 -> biggest;
/* list4 ~ [[[matched f-a][matched f-a]..]node[FaceMOP]][ditto][ditto]] */
for x from 1 to length(list4) do
whatsum(list4(x)(3)) -> sum;
/*dividing no. of input matching f-a to the no. in the respective Frame gives
a comparable proportion, the greatest 'wins out'*/
if (length(list4(x)(1)) / sum ) > biggest
then
length(list4(x)(1)) / sum ->> biggest ->last_biggest;
list4(x)-> longest;
list4(x)(3)->> bestmop -> last_bestmop;
list4(x)(1)->rest;
true ->succ;
elseif
( length(list4(x)(1)) / sum ) =  biggest
then
[%
list4(x)(3), last_bestmop
%]-> bestbet;
false -> succ;
endif
endfor;

if succ = false ;;; if it ends in a tie
then
findbestmop(entr) -> bestmop;   ;;;on to the heuristic
if bestmop = undef ;;;if still undecided
then
oneof( bestbet )-> bestmop; /*random choice from finalists */
valof(bestmop) <- gf_a -> gen;  ;;; is chosen and its f-a grabbed
compare_list(gen,target) -> rest;    ;;;and compared with input
if facemop matches [ == [?y ^bestmop] == ]  ;;;Facemop's node found
then
[^rest ^y ^bestmop] -> longest;
endif
endif
endif;
```

```
decompose(valof(bestmop) <- gf_a) -> gen;
;;;FaceMOP's f-a made a l-o-l
for part in rest do
valof(bestmop) <- support_attribute(part,1);      ;;;Frame f-a support updated
  delete(part,gen) -> gen;
endfor;
for unit in gen do
valof(bestmop) <- support_attribute(unit,-1)
;;;Frame f-a non-support updated
endfor;

enddefine;

define printsub(nomen,b,lmop);
;;;announces the formation of a subMOP instance
 lvars nomen b lmop;
pr('\n'); pr('                INSTANTIATING  SUB_MOP : ');
pr('                                  ');
pr(nomen); pr('\n');
pr('\n  with generalization :\n');
pr('                         ');
pr(b);
pr('\n');
pr(' under the SPEC_MOP : ');              ;;; the FaceMOP in which it occurs
pr('                         ');
pr(lastmop);
pr('\n the content-frame of which contains the following generalizations:);
pr(\n');
pr('\n');
pr(valof(lastmop) <- common);   ;;;lists the face actions and interpretation
enddefine;

define face_acts ->result;
facts ->result;     ;;;makes a procedure function as a list used in matching
enddefine;

define entail(way,E,nodal,pppp,q,r);
/*for each input event an instance of the FLAVOUR 'evvent' is made. Its
variables have information about its sub- and FaceMOP*/
lvars way E nodal pppp q r rr  tag;
vars route d e g y no pec_support pec_match ub_support;
vars ub_match;
 (consword("EV_" >< hd(E) )) ->tag;

unless r = undef
```

```
then
valof(r) <- support -> pec_support;             ;;; support its FaceMOP f-a
have
valof(r) <- common -> pec_match; ;;; the FaceMOP f-a endun-
less;
unless q = undef then
valof(q) <- support -> ub_support;            ;;;support its subMOP f-a has
valof(q) <- common  -> ub_match;                ;;;its subMOP f-a endun-
less;
[% tag ,[%way, nodal%] %] -> route;
 if
routes matches [??d [^tag ?rr]??e ]
then
[^^d [^tag ^rr ^route] ^^e] -> routes;
else
unless ((info(cnode) = []) ;;;the node is empty
or (info(cnode) matches [[?y ?z] == ])
or (info(cnode ) matches [?y ?z] ))

then
hd(info(cnode))->icn;
else "nothing" -> icn;
endunless;
unless   ((Info=[])
  or (Info = [[?y ?z] == ]) or (Info matches [?y ?z] ))
then
  hd(Info)->inf; else "nothing" -> inf;           endunless;
if null( alist ) ;;;arrived at the node at the end of the given path
then
if         ;;;if node has a face feature(e.g. 'brows')
((gg matches [?b : feat]) and
(entry matches [ == ?v4]) and (Info matches      [^b]));;;likewise Info
then ;;; an fsubMOP instance is created
substring(5,(length(lastmop) - 13),lastmop) -> mopem;
consword('mop_' >< mopem >< b >< '_fsub') -> tag;
tag -> latermop;
mophead([^lastmop ^latermop ^latestmop]) -> mopp ->fmopp ->Face;
make_instance([fsub_mop name ^tag common [^b]
gf_a [^b]  Spec ^Face   path ^road  node ^nodename]
) -> valof(tag);
tag -> ivalof(valof(Face),"fSub"); ;;;inform FaceMOP
tag :: ivalof(valof(Face),"Face_fsub_list") ->
ivalof(valof(Face),"Face_fsub_list");
printsub(tag,b,Face);
[^tag] -> info(cnode);
if present([^gg == ^nodename]) ;;;inform db
```

```
then
remove(it);
add([[^tag] ^(links(cnode)) ^nodename]);
endif;
elseif substring(1,3,icn) = 'mop' and
(Info matches [?b : nevtec] or Info matches [?b : nomop])
;;;if there is a MOP in the node and Info is a face feature, say 'brows'
;;; if the face feature is the feature in the fsub_mop
;;;(which it has to be by design) do nothing
then
icn -> latermop;   ;;;set the fsub flag for the path
valof(icn) <- common -> comm;
unless  ((member(b,comm))  or ([^b] = comm))
then
  t_add([^b],Info,nodename);
endunless
elseif ((substring(1,3,icn) = 'mop') and
(Info matches [?b : feat ?c : olist]))
;;; there is a MOP in the node and the new arrival is a face action
then
valof(icn) <- common -> comm;
decompose(comm)-> comm;
if  ((member([^b ^c],comm)) or ([^b ^c] = comm)) ;;;it's a sub_mop
then
  icn -> latestmop; ;;;set flag for sub_mop on path
endif
elseif gg = [] and
(Info matches [?b : feat ?c ])
and entry matches [ == ?v2]
 /*the node is empty and Info is a face action eg 'brows raised' */
then
consword('EV_'><v2)->v2;
Info -> info(cnode);                    /*node content is now the face action */
if present([[] == ^nodename])
;;; the record is changed in db
then
  remove(it);
  add([^(info(cnode)) ^(links(cnode)) ^nodename]);
endif;
elseif gg = [] and Info matches [?h : feat]
;;;the node is empty and Info is a face feature eg 'brows'

then
Info -> info(cnode); ;;;node content is now the feature
if present([^gg == ^nodename]) ;;;change record in db.
then
```

```
  remove(it);
  add([^(info(cnode)) ^(links(cnode)) ^nodename]);
endif;
elseif gg = []
then
if Info matches [?m : evtec]
;;;node is empty and Info is an input face event ID eg 'ev0'
then
  hd( Info ) :: info(cnode) -> info(cnode);  /*Info added to node list*/
  if present([^gg == ^nodename])
;;; the record is changed in db
  then
remove(it);
add([^(info(cnode)) ^(links(cnode)) ^nodename]);
endif;
mophead([^lastmop ^latermop ^latestmop]) ->mopp ->fmopp ->Face;
if (mopp = undef and fmopp = undef and Face /= undef)
;;;no sub_mop or fsub_mop on its path, indexed directly off Face_mop
 Frame
then
  decompose(valof(Face) <- gf_a) -> Like; ;;;Frame f-a made a l-o-l
if Like matches [??z ^sub1 ==]
then
;;;increases support for the Frame f-a used as indexes
  (((valof(Face) <- support)(length(z)+1)) + 1) ->
  ((valof(Face) <- support)(length(z)+1));
endif
elseif (mopp /= undef and Face /= undef)
;;;if path has sub_mop
then  ;;;retrieve the Face_mop & sub_mop f-a
  valof(mopp) <- gf_a -> Like;
  valof(Face) <- gf_a -> ALike;
  for fa in Like do
  if ALike matches [??zz ^fa ==]
  then  ;;;update their f-a support
    (((valof(Face) <- support)(length(zz)+1)) - 1) ->
    ((valof(Face) <- support)(length(zz)+1));
    (((valof(mopp) <- support)(length(zz)+1)) + 1) ->
    ((valof(mopp) <- support)(length(zz)+1));
 endif
 endfor
endif;
mophead([^lastmop ^latermop ^latestmop]) ->
mopp -> fmopp -> Face;
;;;checks & assigns path MOPs to variables
      entail([^^road ],Info,nodename,mopp,fmopp,Face);
```

```
;;;procedure 'entail' will form an 'evvent' FLAVOUR of this input
;;;corresponding to its representation in this branch: part of
;;;a distributed storage of the input.
elseif ((Info matches [?n ?u]) and member(n,face)
and member(u,o))   ;;;if its a face action eg 'brows raised'
then   ;;;it becomes the node content
  Info  -> info(cnode);
;;; the record is changed in db
if present([^gg == ^nodename])
then
          remove(it);
  add([^(info(cnode)) ^(links(cnode)) ^nodename]);
endif;
endif;
elseif (substring(1,2,icn) = 'ev' ;;;if an input ID meets an input ID
and substring(1,2,inf) = 'ev' )
then
hd(Info) :: info(cnode) -> info(cnode); ;;;add input ID to the contents
if present([^gg == ^nodename])
;;; the record is changed in db
then
  remove(it);
  add([^(info(cnode)) ^(links(cnode)) ^nodename]);
endif;
mophead([^lastmop ^latermop ^latestmop]) -> mopp -> fmopp -> Face;
;;;checks & assigns path MOPs to variables
entail([^^road ],Info,nodename,mopp,fmopp,Face);
/*procedure 'entail' will form an 'evvent' FLAVOUR of this input
  corresponding to its representation in this branch: part of a distributed
  storage of the input.*/
elseif ((gg matches [?x : nevtec ?y : nevtec]) and
(Info matches [?w : nevtec ?z : nevtec])  and
(entry matches [ == ?v3]))
and [^x ^y] = [^w ^z]
;;;if Info = f-a eg 'brows raised' and node content is the same
then
substring(5,(length(lastmop) - 13),lastmop) -> mopem;
;;;a sub_mop is formed:
(consword('mop_' >< mopem >< w ><'_'>< z >< '_sub')) -> tag;
tag -> latestmop;
mophead([^lastmop ^latermop ^latestmop])
-> mopp -> fmopp -> Face ;
make_instance([sub_mop name ^tag  Spec ^Face common [^x ^y]
gf_a [^x ^y] support [20]
fSub ^fmopp ^^Info node ^nodename]) -> valof(tag);
valof(tag) <- subb;
```

```
true -> facet;
printsub(tag,[^x ^y],Face);
[^tag] -> info(cnode);
if present([^gg == ^nodename])
;;; the record is changed in db
then
  remove(it);
  add([^(info(cnode)) ^(links(cnode)) ^nodename]);
endif;
unless Face = undef
then ;;;inform Face_mop
  tag -> ivalof(valof(Face),"Sub");
endunless;
unless fmopp = undef
then ;;;inform fsub_mop
  tag -> ivalof(valof(fmopp),"Sub");
endunless;
tag -> latestmop
elseif
gg matches [?e : evtec ?f : evtec ==] and Info matches [?g : evtec]
;;;Info has an input ID and the node contains other input IDs
then
  delete(hd(Info),(info(cnode))) -> info(cnode);
  ;;;removes duplicates if any
[^^Info  ^^(info(cnode))] -> info(cnode); ;;; the record is changed in db
if present([^gg == ^nodename])
then
  remove(it);
  add([^(info(cnode)) ^(links(cnode)) ^nodename]);
endif;
mophead([^lastmop ^latermop ^latestmop]) ->
mopp -> fmopp -> Face;
entail([^^road ],Info,nodename,mopp,fmopp,Face);
/*procedure 'entail' will form an 'evvent' FLAVOUR of this input
 corresponding to its representation in this branch: part of a distributed
 storage of the input.*/
else
if present([^gg ?v ^nodename])
then
  remove(it);.
  add([^(info(cnode)) ^v ^nodename])
endif
endif
else
/*alist still has indexes, & unless node's db 'links' field gives the daughter
 node, they are created & added to db*/
```

```
unless links( cnode ) matches [== ^(hd(alist)) ?newnode ==]
then
gensym("m") -> newnode ;  ;;;give ID to new node
;;; the record is made in db
add([[] [] %newnode%] );
consnode( [], [] ) -> valof( newnode ); ;;;makes the record
hd(alist) :: (newnode :: (links( cnode ))) -> links( cnode );
    ;;;adds the links
if present([^gg == ^nodename])
then
  remove(it);
  add([^gg ^(links(cnode)) ^nodename]);
endif;
endunless ;
t_add( tl( alist ), Info, newnode ) ;;;recursion with tail of alist
endif;
enddefine ;

define show_tree( nodename ) ;
/*uses a lib.pack. to draw a tree from a list supplied by gentree*/
define gentree( nodename ) -> List ;
vars cnode L ;
valof( nodename ) -> cnode ;
[% info(cnode) >< ' ' >< nodename;
for L on links( cnode ) do
if length( L ) mod 2 = 0 and length( L ) > 0
then
[% hd( L ) ; gentree( hd(tl( L )) ) %]
endif
endfor %] -> List;
enddefine ;
show_tree( gentree( nodename ) )
 enddefine ;
```

## A.4 POP11 code transforming face geometry into face actions:

```
define pics_to_feat(veclist) -> factlist;
;;;the face geometry is in 'veclist'
vars pp name;
load garryrules.p
[
brows_contracted brows_low brows_raised mouth_bared
eyes_u_lid_tensed
brows_centre_raised brows_in_raised eyes_inlid_raised
eyes_u_lid_lowered
 eyes_l_lid_lowered
eyes_down eyes_closed eyes_u_lid_raised eyes_wide eyes_narrowed
eyes_l_lid_raised eyes_l_lid_tensed mouth_l_lip_tensed
mouth_u_lip_everted
 nose_screwed nose_flared mouth_down mouth_wide mouth_slightly_open
 mouth_open mouth_square mouth_u_lip_tensed mouth_l_lip_everted
 mouth_shut mouth_pulled mouth_compressed mouth_up
 mouth_l_lip_raised mouth_u_lip_raised mouth_bared mouth_l_lip_lowered
 mouth_l_lip_tensed cheeks_n_l_vert jaw_drop cheeks_raised
cheeks_dropped
] -> pp;
[%
  for name in pp do
    if valof(name)(veclist) ;;;if face action rule of that name is true
    then
      name ;;;put the name in a list
    endif
  endfor
%] -> factlist;
pr('These are the face actions expressed: \n');
factlist=>
enddefine;
```

## A.5 POP11 code transforming a face expression in geometric format into an interpretation:

```
;;;interprets a geometric face description in terms of emotion
;;;called by pic_to_feat(geometric 34 points description in normgarrylist)
load oracle6.p;
load /pics_to_feat.p;

vars ghost;

;;;takes underscored face-features and returns them in feature action format
define exscore(list1) -> list2;
vars numb first second entry newentry newentrylist ;
[] -> newentrylist;
for entry in list1 do
;;;for stri in entry do
entry >< '' ->entry;
unless (locchar('_',1,entry) = false )then
locchar('_',1,entry) -> numb;
consword(substring(1,(numb - 1),entry)) -> first;
consword(allbutfirst(numb,entry)) -> second;
second :: newentrylist ->newentrylist;
first :: newentrylist ->newentrylist;
else entry :: newentrylist -> newentrylist
;;;revs(newentrylist) -> list2;
;;;endfor;
endunless
endfor;
newentrylist -> list2;
enddefine;

define pic_to_interp(ghost);
vars underscore ghoul;
maketree1();
pics_to_feat(ghost) -> underscore; ;;;see overleaf
;;; returns a list of face-actions from an input list of xy face-vectors
;;;representing 34 points on a face which is run through garryrules.p

exscore(underscore) -> ghoul;
traverse(ghoul,"top1"); ;;;enters the face actions into Janus in RETRIEVE
mode
enddefine;
```

UNIVERSITY OF GREENWICH LIBRARY

# Appendix B:

# knowledge acquisition material.

## B.1: Face action questionnaire format

Use a separate answer book for each photo and please <u>record the number</u>.
The female faces should be judged in comparison with photo 4 (photo norm ).
The male faces should be judged in comparison with photo 6 ( " " ).
Photo
No. · (Here please record the number on the back of the photo)

---

Please judge whether each of the actions specified is present compared to photo norm. Put a tick against those you judge present
The example Nos. reference a printed example of the face action

**The EYEBROWS are:** (one or more of the below may be

example

6 a) pulled together
This makes the eyebrows slant towards the nose

5,14 b) lowered
The whole of each eyebrow is lowered onto the eye

11,2 c) raised
Both brows arch up causing creases accross whole forehead

12,13 d) the inner part of each eyebrow is raised and pulled together causing the centre only of the forehead to show arched creases - the eyebrows slant upwards perhaps

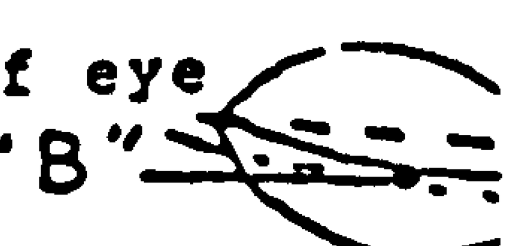
10 e) straight across often with low eye-brow laterally with the inner part of the eyebrow slightly raised, flattening the eyebrow - the skin below it is often triangular over the lid.

f) other: (write what):

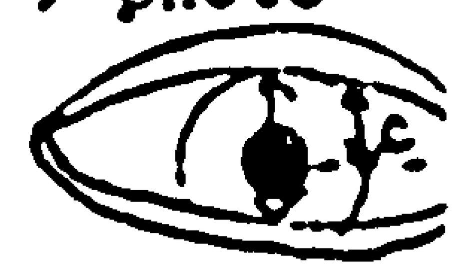
**The EYES are:** (one or more of the below may be true)

4 a) closed

4,7,8,10 b) upper lids down: "B" is below inner angle of eye
or "B" is below "B" of photo norm

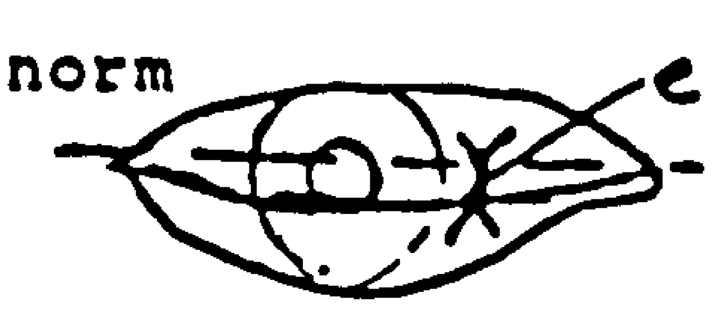



2 c) upper lids raised: most over the pupil: "c" > photo norm or:



5 d) inner part of upper lid raised most: "d" > photo norm

12 e) lower lid raised: "e" < photo norm



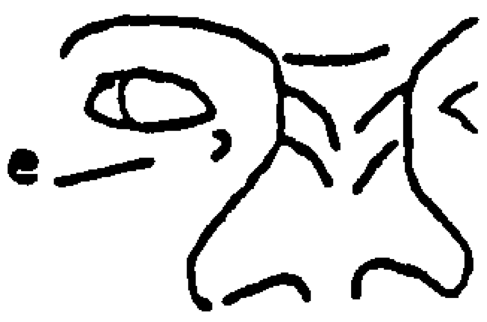

9 f) widely-open: "gap" > photo norm
or:
9,12,13 g) narrowly-open: "gap" < photo norm


h) other:

## B.1: Face action questionnaire format

### The NOSE is:
(one or more of the below may be true)

13,8 a) widened at the nostrils cf photo norm
and/or:



eyes 12 nose 1,4 b) screwed-up causing a furrow across the bridge or at the sides of the bridge - these should be looked for

and/or:

4,1 but not 5,7,8 b) nose-mouth grooves are curving down and more accentuated due to nostrils pulled up or mouth corners down or both.

c) other:
(write what):

### The MOUTH is:
(one or more of the below may be true)

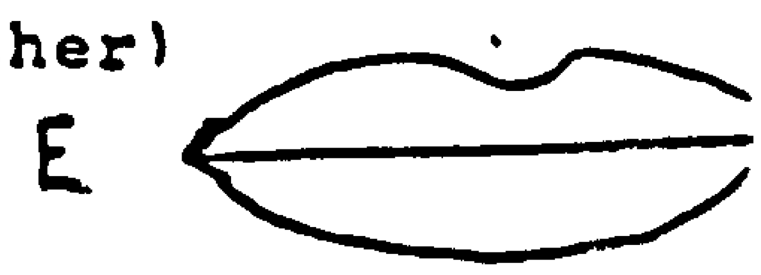
1,4 a) teeth show

8 b) corners of mouth are down
or:

3,7 h) corners of mouth are up

13 c) lips and mouth wide-open

7,4 d) lips slightly apart(teeth may be together)




3,14 e) lips shut

8,1 f) lips pulled back(ie E-F wider) cf photo norm

10 g) lips compressed i.e. M-U (see j below) < photo norm the lower lip compressed against it.

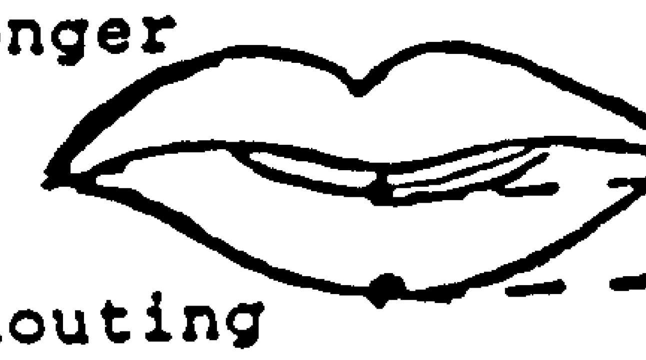
1 h) lower lip lowered (ie "L" is lowered ) cf photo norm

10,8 i) lower lip is raised (ie "L" is raised ) cf photo norm

1,13,8 j) upper lip is raised (ie "U" is raised ) cf photo norm

1,13 k) upper lip is tensed ( ie "M-U" is shorter ) cf photo norm

1,4 l) lower lip is everted (turned out and down) i.e. "L" is higher than in photo norm and L-N is longer



13 m) widely-open mouth with tension as in shouting -the teeth are separated and the upper lip is thinner than in photo norm

o) other:
(write what):

**The CHEEKS are:**
(one or more of the below may be true)

5,13,4,8,7 a) raised-up tending to push up skin under eyes in extreme or fold or bunch over the cheek-bones in lesser degree,

6 c) Dropped, more hollowed or toneless

d) other
(write what):

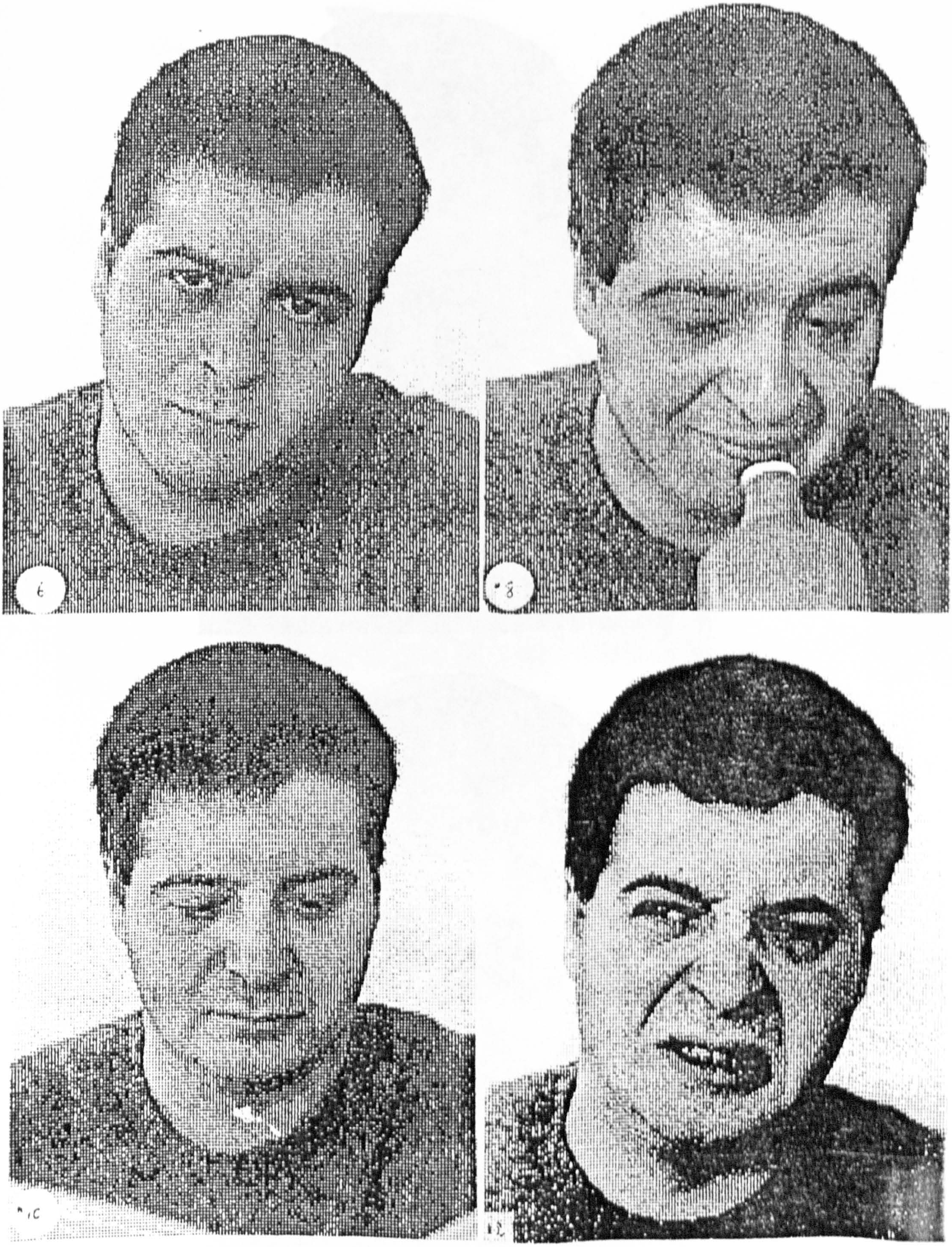
**The jaw is:**

13 a) dropped ( The mouth must be open and the chin lower down than the chin in photo norm.

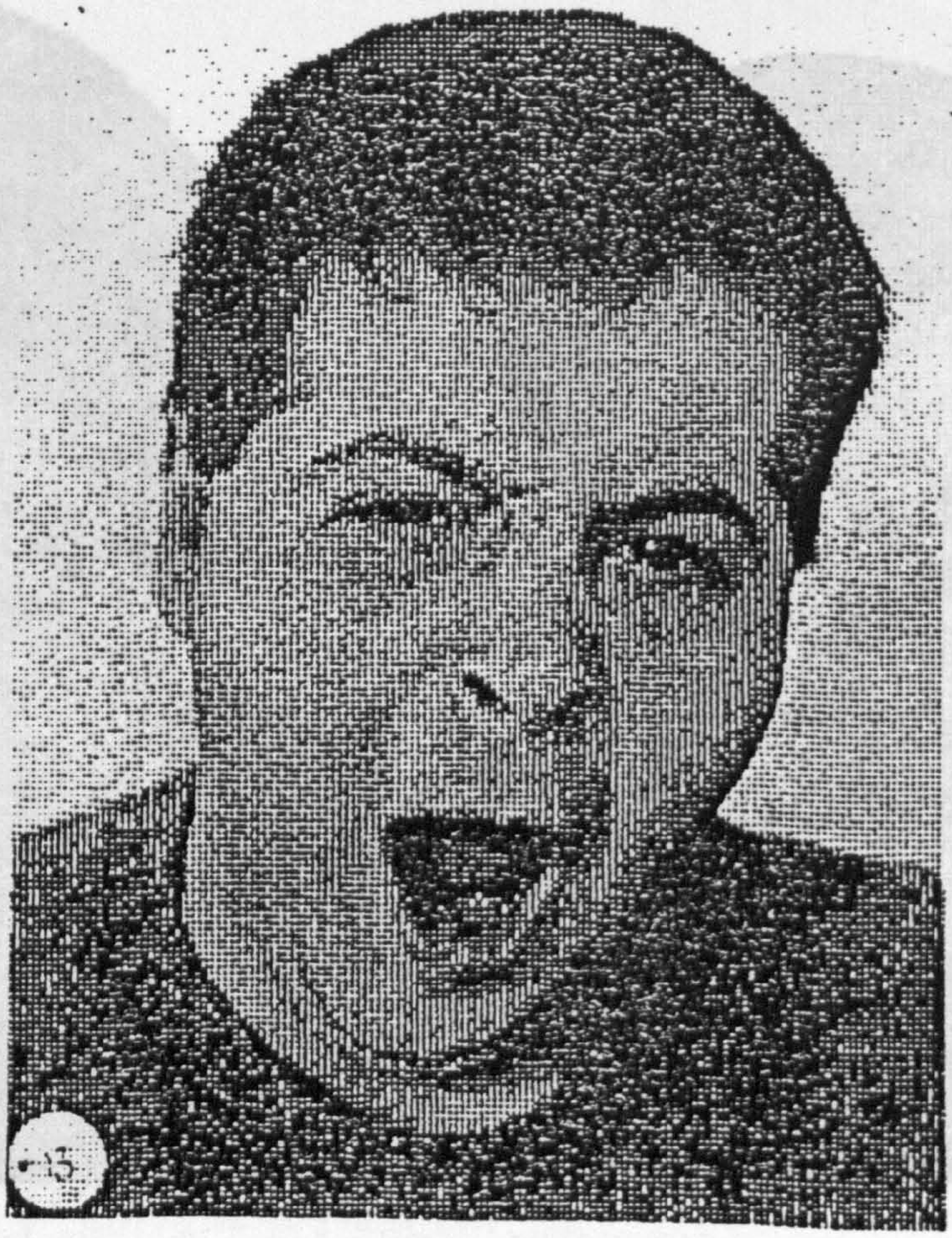
b) other:
(write what):

## B.2: T-series photographs

## B.2: T-series photographs



## B.2: T-series photographs



B.3: Garryphotos



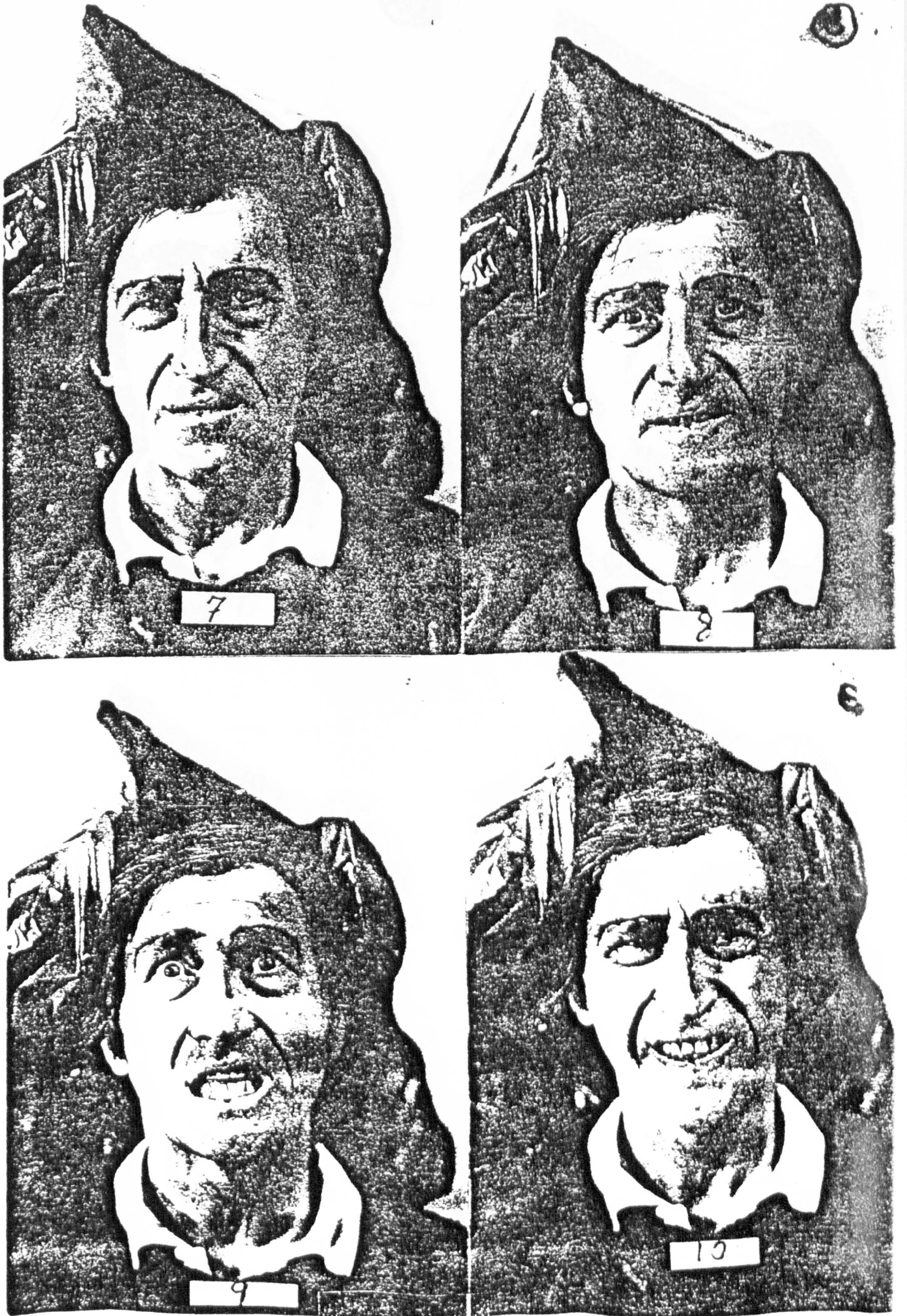
3


4


5


6

## B.3: Garryphotos




11


12


13


14

## B.3: Garryphotos




## B.4: Instruction sheet accompanying questionnaire:

Instructions accompanying the questionnaire to judges of 'PFA' facial expressions. Their task was to judge which face actions were present in photographs. The photographs from the transparencies of"Pictures of Facial Affect" were made with the consent of the Publishers, The Consulting Psychologists Press, Inc.

These are the things to know about filling in the answers:

1. There is a questionnaire to be answered for each face to be judged and there are five faces to be judged and two comparison faces: male and female. Please enter the number which is on the back of the photo on the questionnaire.

2.Do not answer questionnaires on the two comparison neutral photos 4 & 6. The questions should be answered after comparing the target face with the relevant neutral photo (eg photo 4 for a female and photo 6 for a male) and after looking at any examples of the face action supplied.

3.The photos to be judged are in the envelope labelled "Test Photographs". They are numbered 1, 2, 3, 7, 8.

4.The judgements required are about facial actions and I have tried in many cases to provide some help in the form of drawings and/or photocopy prints. These latter are in the envelope marked "Examples" and are classified into 3 areas of face. Some judgements require an examples for two areas eg top of nose is seen with the eye examples and the bottom with the cheeks & mouth area examples.

5 The questionnaire is in multiple-choice format. Please TICK the face actions you judge to be present(there may be more than one) and leave the others blank. The intensity of the action doesn't matter: if it shows at all, it's present!.

6. The object is to test a computer's judgement against human experts. The computer judges very finely by mathematical measurements whether, say, the mouth corners of the test photograph are higher than those of the neutral expression in the photo norm. Accordingly, some of the judgements required of the experts are very fine.

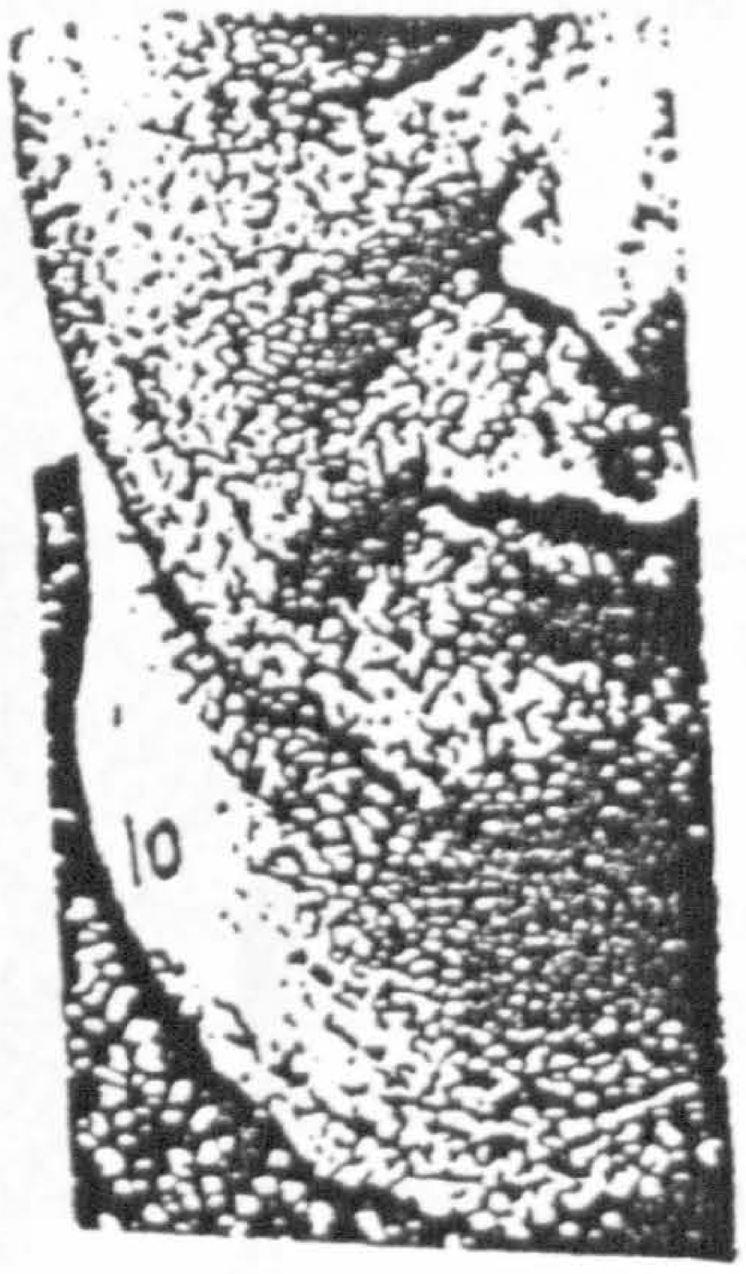
7. The most difficult brow action seems to be the "straight" variety which is also called "in_raised" since the important aspect is the raising of the inner ends so that their normal downcurve is eliminated, giving a straighter look. This does not mean, however, that the outer down curve is eliminated.

8. I do not think that it is at all easy and am truly grateful for your help. Thank you very much.

B.5a: Lower face action examples accompanying face action questionnaire:

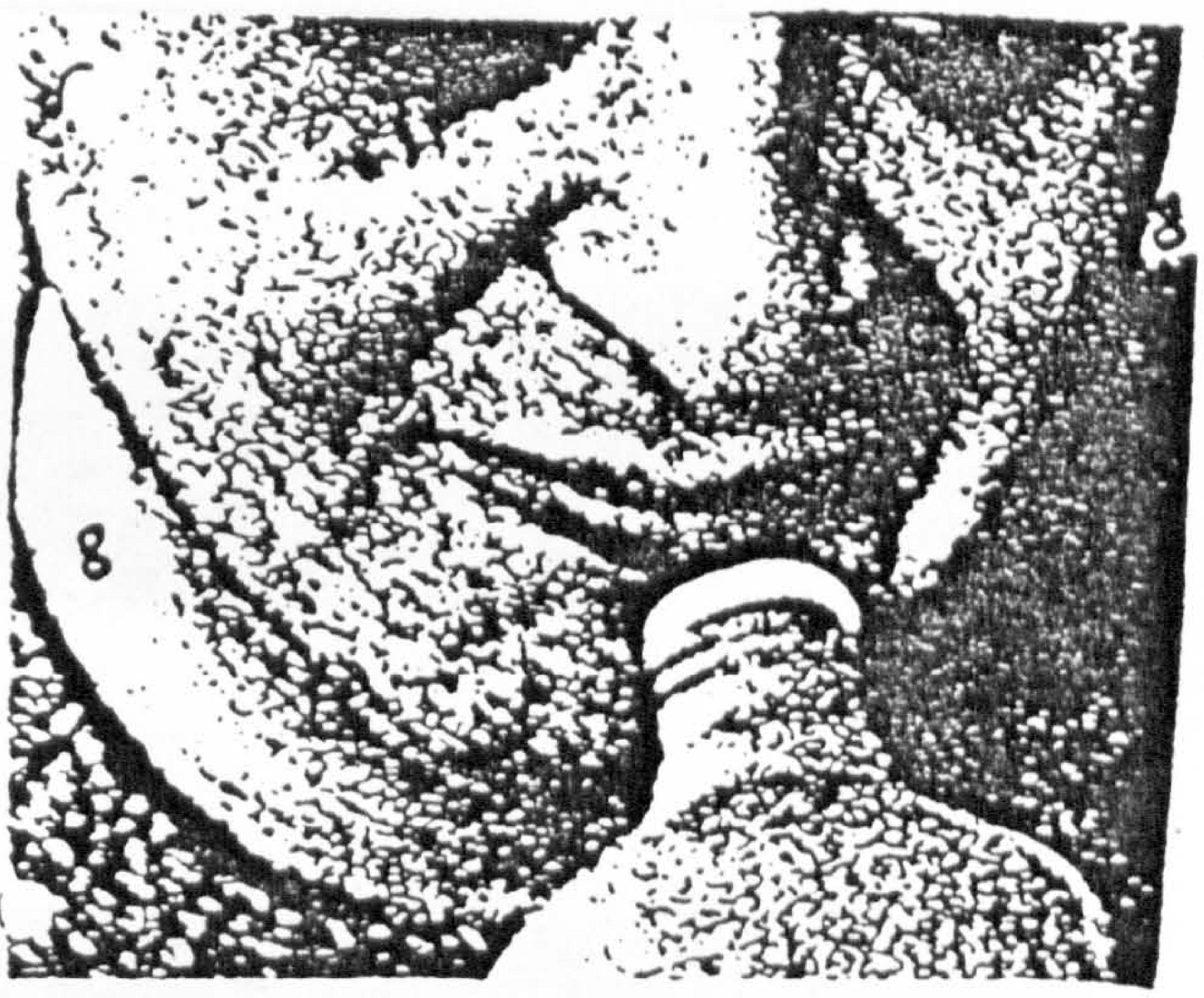


B.5a: Lower face action examples accompanying face action questionnaire:

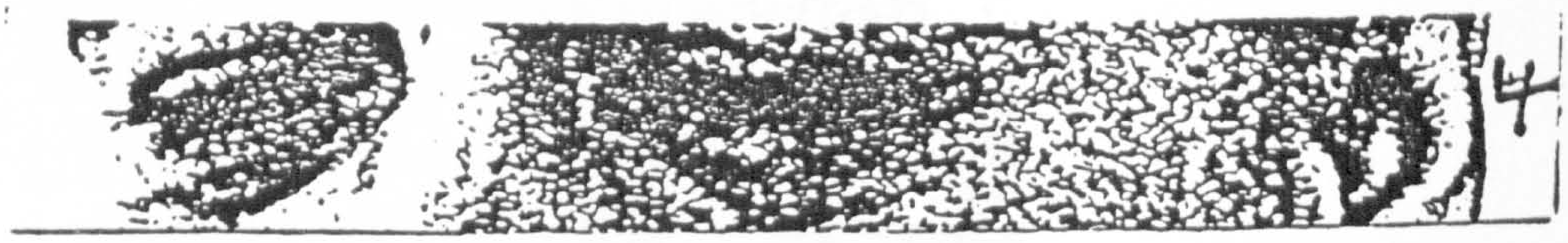


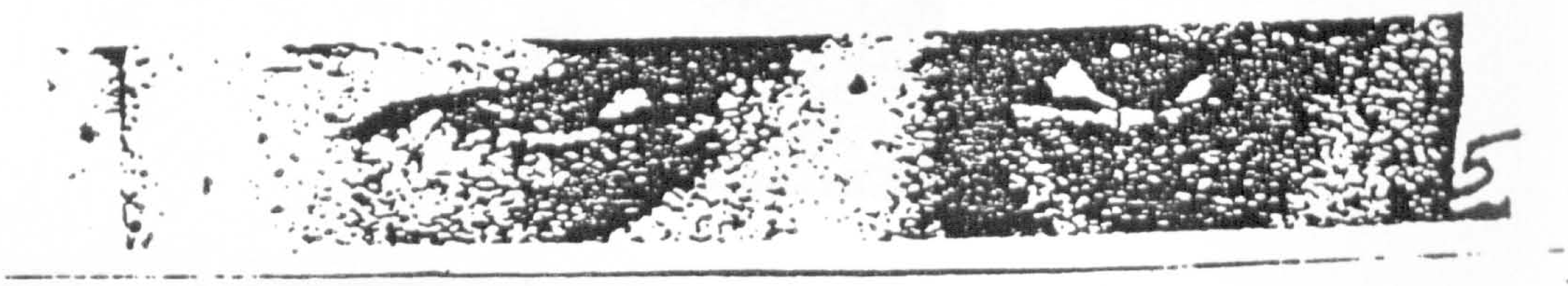


B.5a: Lower face action examples accompanying face action questionnaire:

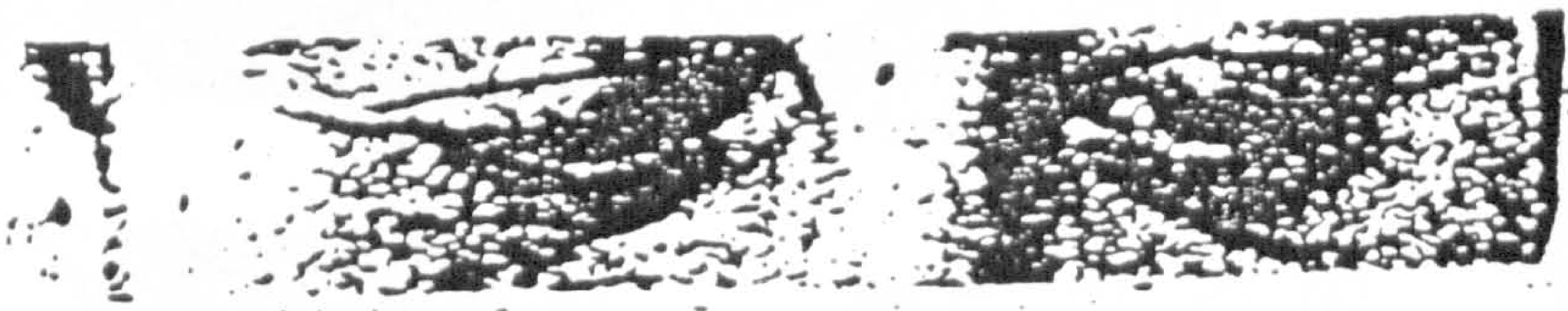


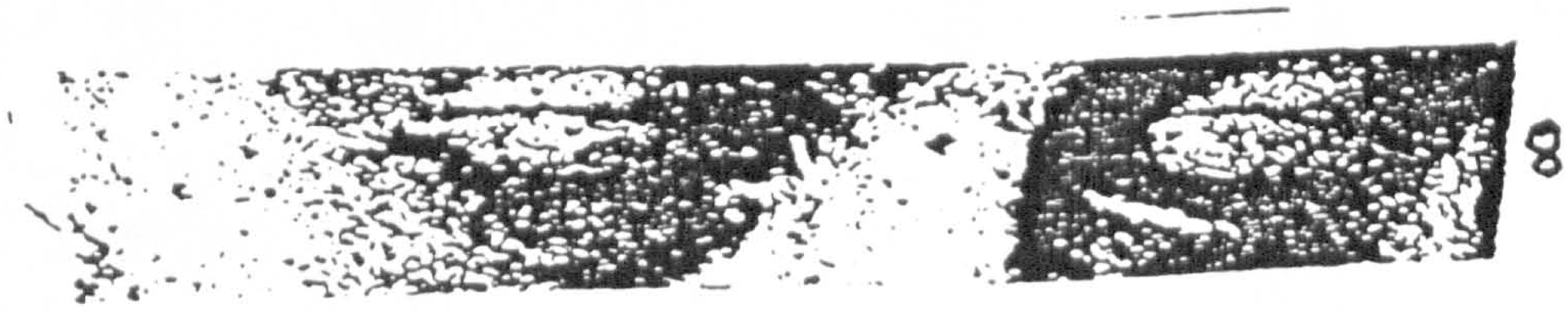

B.5b: Eye action examples accompanying the face-action questionnaire:



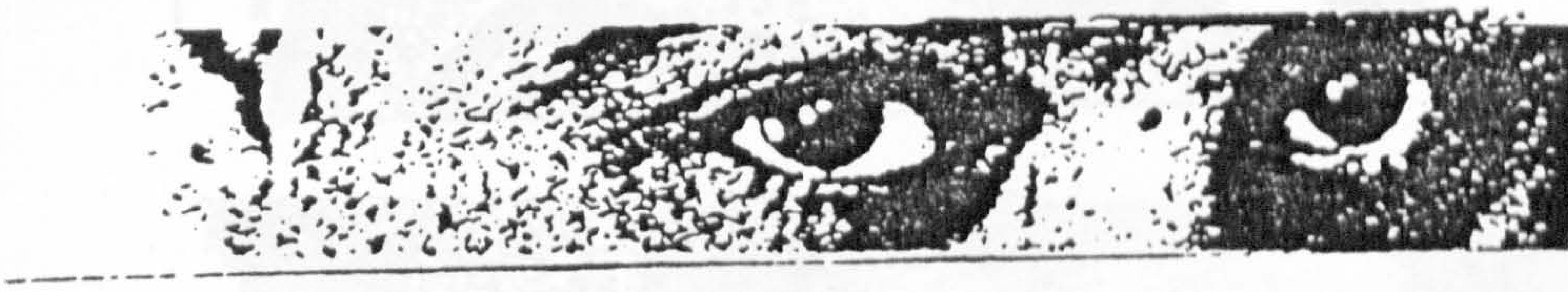





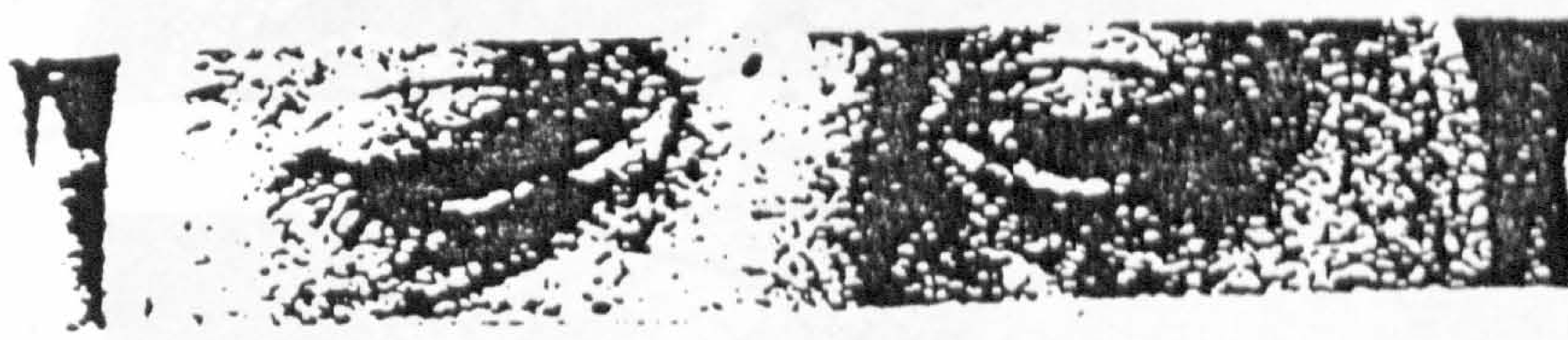


B.5b: Eye action examples accompanying the face-action questionnaire:

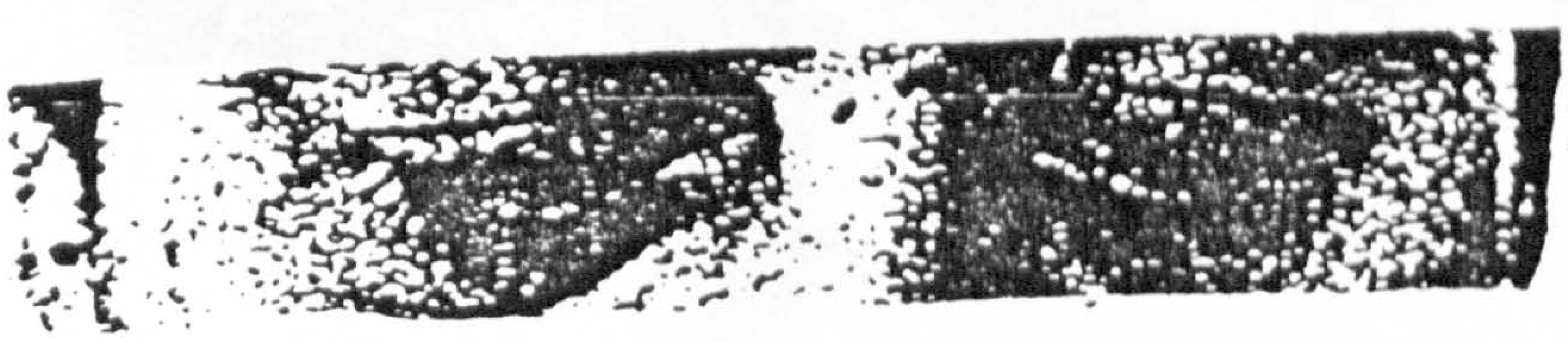


9

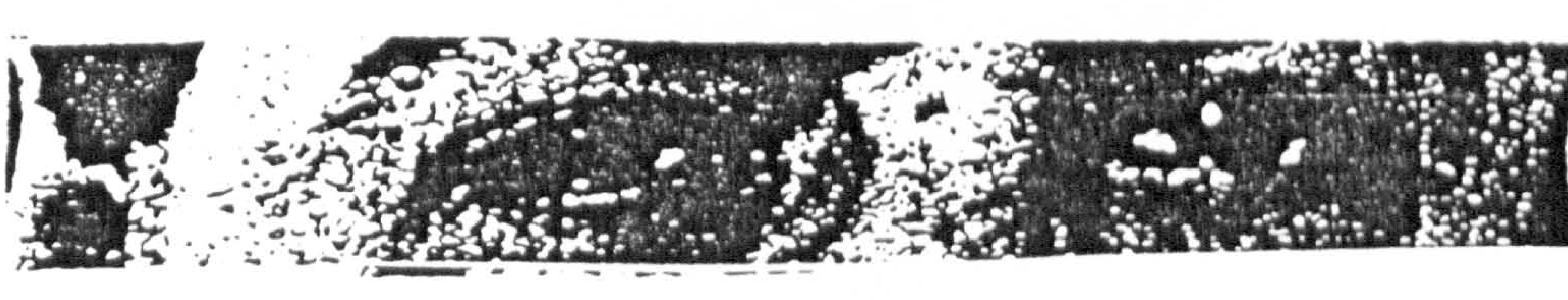


11

12

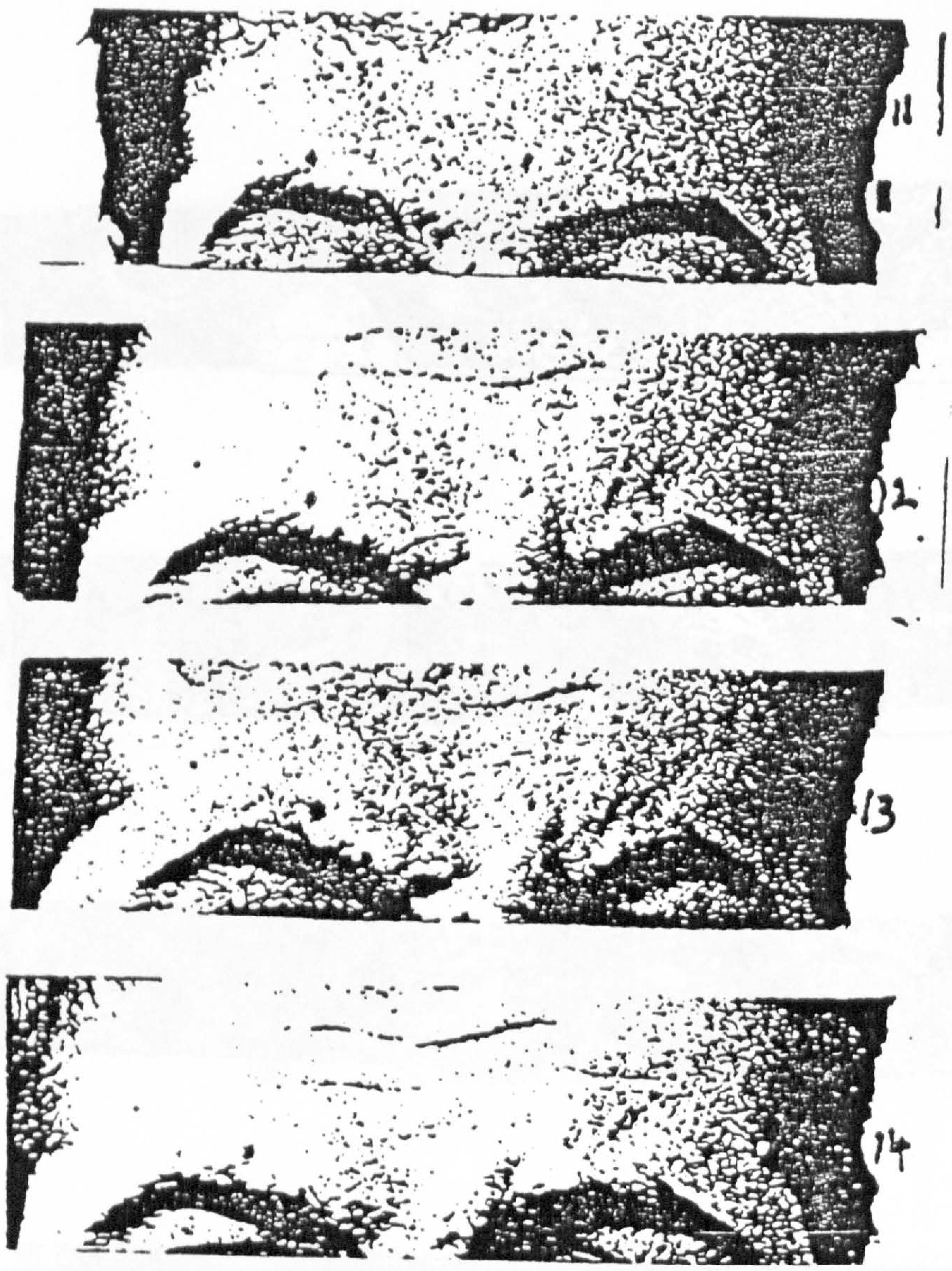


10

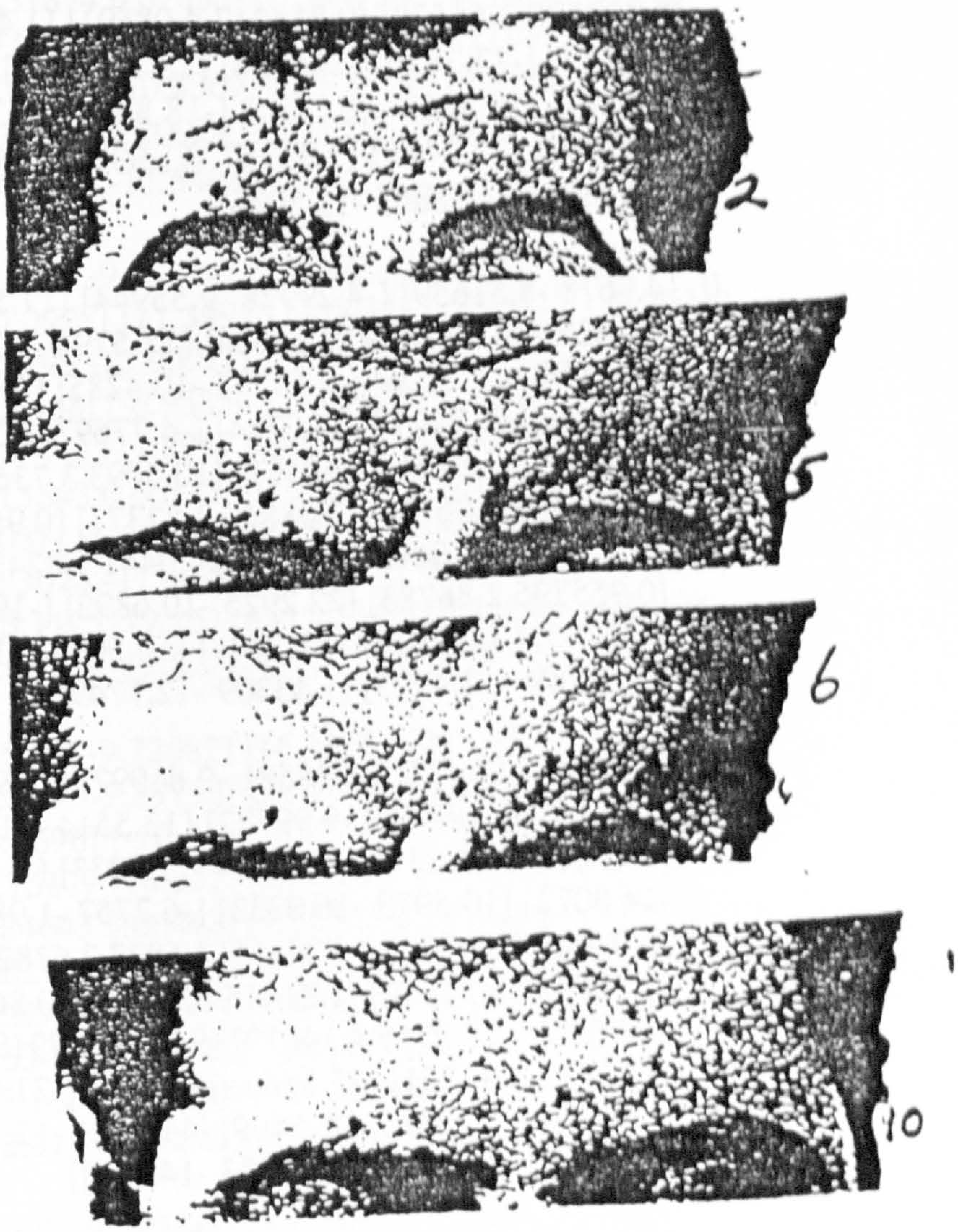


13

B5c: Brow action examples accompanying the face action questionnaire:



B5c: Brow action examples accompanying the face action questionnaire:



B6: The normalized representation of garryphotos - 34 points per face:

```
vars normgarrylist;
[

[[-15.057 -7.90729] [-5.34734 -9.23997] [13.9312 -10.8392] [4.3623
  -10.7504] [-8.16173 -9.6842] [9.99109 -10.928] [-9.56893 -10.0396]
  [7.88029 -11.3723] [-9.85036 -10.8392] [-9.70965 -9.06228] [-10.2725
  -14.2153] [7.03598 -15.7257] [-6.19166 -1.77692] [5.62878 -2.39884]
  [0.0 0.0] [-8.16173 2.57653] [10.4132 1.33269] [0.281439 1.59923]
  [0.703598 2.66538] [0.844317 4.08691] [1.4072 3.02076] [1.97007
  10.5727] [-23.7816 0.266538] [18.7157 0.0] [1.12576 2.93191] [1.12576
  2.93191] [19.7007 -10.7504] [-18.8564 -8.52921] [0.985037 3.10961]
  [-1.68863 -21.9449] [7.73957 -11.9942] [7.59885 -10.6615] [-4.64374
  -13.3269] [1.68863 -13.5046]]

[[-14.9678 -8.51659] [-4.29928 -9.55944] [17.5156 -10.2547] [5.57314
  -10.4285] [-7.32469 -9.64634] [12.5794 -10.863] [-9.55395 -9.82015]
  [10.5093 -11.2106] [-9.55395 -10.4285] [-9.55395 -8.95111] [-11.3055
  -12.9487] [8.75779 -14.0784] [-4.77697 -2.34641] [7.48392 -2.2595]
  [0.0 0.0] [-7.64316 4.60591] [10.5093 3.73687] [1.27386 1.21666]
  [0.955395 2.86783] [2.38849 3.82377] [0.955395 2.86783] [4.14004
  10.2547] [-27.0695 6.6047] [24.0441 1.82498] [0.955395 2.86783]
  [0.955395 2.86783] [22.2925 -10.6892] [-19.1079 -7.90826] [0.955395
  2.86783] [-0.159232 -21.3784] [10.5093 -11.732] [10.5093 -9.82015]
  [-4.14004 -12.4273] [1.43309 -12.7749]]

[[-16.6786 -8.95977] [-5.38581 -9.61997] [18.9372 -10.3745] [4.51713
  -10.3745] [-8.68679 -9.99722] [13.5514 -11.129] [-11.1191 -10.0915]
  [10.9454 -11.2233] [-10.9454 -11.2233] [-11.2928 -9.24271] [-12.1615
  -14.8072] [10.5979 -16.0333] [-6.7757 -1.98058] [8.16558 -2.26352]
  [0.0 0.0] [-7.47064 4.9043] [12.6827 3.67822] [1.04241 3.4896]
  [0.868679 4.33842] [1.21615 6.50762] [0.868679 5.47018] [-2.77977
  13.7698] [-27.2765 4.14979] [9.7292 1.79195] [0.521207 4.43273]
  [1.38989 4.99861] [25.8866 -10.1858] [-21.8907 -8.77114] [0.868679
  4.43273] [-1.38989 -22.5409] [11.1191 -12.7323] [10.9454 -9.90291]
  [-4.51713 -14.2413] [2.08483 -14.147]]

[[-19.0919 -9.04544] [-6.89429 -9.54244] [15.5563 -11.6299] [3.0052
  -10.3376] [-10.0763 -10.2382] [12.1976 -11.5304] [-12.5511 -10.3376]
  [9.54594 -11.8287] [-13.0815 -11.0334] [-12.9047 -9.54244] [-13.435
  -15.6059] [9.36916 -17.0969] [-7.24784 -1.49101] [6.89429 -1.88861]
  [0.0 0.0] [-9.19239 4.17482] [11.4905 3.18081] [0.53033 4.77122]
  [0.353553 4.57242] [1.06066 6.56043] [1.06066 5.36762] [2.2981
  13.3197] [-29.6985 4.37362] [25.9862 1.0934] [0.176777 4.27422]
  [0.707107 4.97002] [22.981 -11.0334] [-22.981 -8.94604] [0.53033
  4.37362] [-2.47487 -22.7627] [9.36916 -12.3257] [9.19239 -10.8346]
  [-5.65685 -14.8107] [2.12132 -14.6119]]
```

*B6: The normalized representation of garryphotos - 34 points per face:*

[[-17.8894 -8.70826] [-6.30387 -9.28881] [16.1856 -10.8369] [3.4075 -10.6434] [-9.54099 -10.1596] [11.4151 -11.224] [-12.0966 -10.2564] [9.37062 -11.7078] [-11.7559 -10.8369] [-11.9262 -9.28881] [-10.5632 -14.6105] [10.0521 -15.8684] [-7.49649 -1.35462] [6.98537 -2.61248] [0.0 0.0] [-9.20024 3.38655] [11.2447 2.41896] [0.511125 2.99951] [0.170375 3.4833] [0.851874 6.67633] [0.34075 4.74116] [1.87412 12.7721] [-22.1487 1.93517] [21.4672 -0.387034] [0.170375 4.06385] [0.170375 4.06385] [20.9561 -10.8369] [-23.3414 -10.8369] [0.34075 4.54765] [-2.0445 -22.7382] [9.20024 -12.1916] [9.71136 -10.1596] [-5.28162 -13.6429] [1.87412 -13.5462]]

[[-17.5117 -7.77144] [-5.61696 -9.58814] [17.3465 -10.7993] [4.29532 -10.3956] [-9.58187 -10.0928] [11.3991 -11.0011] [-11.5643 -10.2946] [9.25146 -11.203] [-11.7295 -10.9002] [-11.5643 -9.0835] [-11.7295 -13.6252] [6.11257 -15.6438] [-6.93859 -1.00928] [6.60818 -1.8167] [0.0 0.0] [-8.09503 -5.95474] [11.8947 4.7436] [0.330409 3.93618] [0.826023 5.34917] [1.15643 6.66123] [0.660819 5.65195] [2.97368 14.7355] [-13.712 6.5603] [16.5205 6.05567] [0.330409 4.94546] [0.330409 4.94546] [22.3026 -11.4048] [-22.3026 -8.47793] [0.495614 5.24824] [-0.991228 -23.1125] [9.25146 -12.3132] [9.25146 -10.0928] [-3.30409 -13.6252] [1.65205 -13.8271]]

[[-15.6929 -9.01385] [-5.88482 -9.72082] [16.1455 -10.0743] [4.07411 -10.251] [-9.20447 -9.89756] [10.1098 -10.5162] [-10.5625 -9.98593] [8.14822 -10.8696] [-10.4116 -10.5162] [-10.5625 -9.63245] [-9.05357 -13.4324] [8.75179 -14.051] [-6.79018 -2.82788] [4.97947 -3.18136] [0.0 0.0] [-7.84643 3.00462] [9.35536 2.38602] [-0.603572 2.12091] [-0.301786 2.82788] [-0.301786 5.12552] [-0.603572 3.71159] [1.35804 10.3394] [-32.8947 -3.26973] [24.7464 0.0] [-0.452679 2.91625] [-0.754465 3.53484] [20.9741 -10.1627] [-19.9179 -8.83711] [-0.603572 3.00462] [-3.16875 -21.2091] [7.99732 -11.2231] [7.39375 -9.98593] [-3.62143 -12.8138] [1.35804 -12.9906]]

[[-19.6684 -8.36455] [-7.0877 -9.57388] [18.0736 -10.1786] [2.83508 -10.2793] [-10.9859 -10.0778] [11.8719 -11.0855] [-13.4666 -9.97699] [9.3912 -11.1863] [-13.2894 -10.9848] [-13.2894 -8.86844] [-13.1122 -13.4034] [10.1 -15.8221] [-8.68243 -1.10855] [7.0877 -2.82178] [0.0 0.0] [-9.21401 5.03889] [11.1631 3.72878] [1.06315 3.12411] [1.59473 4.03111] [-0.177192 6.95366] [0.354385 5.34122] [1.59473 15.1167] [-37.0332 4.93811] [26.4017 1.61244] [0.531577 4.23266] [0.531577 4.23266] [23.921 -10.7832] [-24.9841 -8.06222] [0.531577 4.23266] [-2.65789 -22.675] [8.85962 -12.1941] [9.3912 -10.4809] [-3.89823 -13.5042] [0.531577 -13.8065]]
[[-16.3228 -8.16191] [-6.06274 -9.29839] [15.0791 -10.5382] [3.73092

B6: The normalized representation of garryphotos - 34 points per face:

-10.6415] [-8.39457 -10.0216] [10.5709 -11.2614] [-10.5709 -9.91828] [9.01639 -11.3647] [-10.8819 -11.0547] [-10.7264 -8.98844] [-10.5709 -14.4642] [7.92821 -16.6338] [-6.68456 -1.03315] [7.30639 -2.37625] [0.0 0.0] [-7.61729 6.40555] [9.94912 4.95914] [0.62182 2.99615] [0.93273 4.33925] [1.86546 10.3315] [1.3991 8.47186] [2.33183 17.047] [-42.4392 20.1465] [14.6128 -2.89283] [1.55455 5.3724] [1.55455 5.99229] [21.7637 -10.3315] [-20.3646 -7.95528] [1.3991 5.78566] [-1.55455 -23.4526] [9.17184 -12.6045] [9.17184 -10.0216] [-3.73092 -14.0509] [1.3991 -14.2575]]

[[-16.4584 -7.82008] [-5.9433 -9.5681] [14.1725 -10.1201] [3.9622 -10.3961] [-9.14354 -9.9361] [10.5151 -10.4881] [-10.2103 -9.8441] [7.9244 -10.9481] [-10.5151 -10.3041] [-10.2103 -9.4761] [-8.38158 -13.1561] [8.07679 -14.4441] [-7.01005 -2.85203] [5.63852 -3.49604] [0.0 0.0] [-8.22918 3.03603] [10.0579 2.02402] [0.457177 1.84002] [1.06675 2.76003] [1.06675 6.44007] [0.457177 4.87605] [2.59067 12.2361] [-32.4596 2.57603] [27.1258 0.460005] [0.761961 3.86404] [0.761961 3.86404] [21.0301 -10.3041] [-21.1825 -8.09608] [0.304785 4.04804] [-2.89545 -21.5282] [8.38158 -11.1321] [8.07679 -10.0281] [-3.9622 -13.4321] [0.761961 -13.3401]]

[[-16.3147 -9.08523] [-5.547 -9.69091] [15.8253 -10.3975] [4.40497 -10.2966] [-9.29939 -9.89281] [11.4203 -10.7004] [-11.2572 -10.0947] [9.29939 -11.2051] [-11.4203 -10.8013] [-11.2572 -8.5805] [-12.0729 -15.8487] [9.29939 -17.0601] [-7.17847 -1.11042] [7.34162 -2.01894] [0.0 0.0] [-9.78883 4.54262] [11.5834 3.33125] [0.489441 5.35019] [0.815736 4.84546] [0.815736 7.77292] [0.326294 6.25872] [2.12091 14.5364] [-22.8406 -1.7161] [17.4567 -0.504735] [0.652589 4.94641] [0.489441 5.35019] [22.3512 -10.2966] [-21.6986 -9.18618] [0.652589 5.1483] [-1.30518 -22.814] [9.13624 -12.0127] [9.13624 -9.79186] [-4.24183 -14.6373] [2.44721 -14.9402]]

[[-17.3212 -8.36271] [-6.45609 -9.58422] [15.1167 -9.96008] [3.46424 -10.3359] [-9.13301 -9.67819] [11.0226 -10.7118] [-11.0226 -10.054] [8.50315 -10.9937] [-11.1801 -10.8057] [-11.1801 -9.20837] [-11.0226 -14.5643] [7.71582 -15.8797] [-7.40089 -2.25511] [6.61356 -3.00682] [0.0 0.0] [-9.44794 2.72493] [12.2823 1.50341] [0.314931 2.25511] [0.157466 2.81889] [0.314931 5.91967] [0.314931 4.60419] [1.57466 12.4031] [-21.4153 -0.657741] [17.7936 -2.16115] [-0.629863 3.66456] [0.157466 4.51022] [19.8407 -10.148] [-21.1004 -8.6446] [0.314931 4.13437] [-2.36199 -22.1753] [8.50315 -11.6514] [8.18822 -10.148] [-4.25157 -13.9065] [2.20452 -14.4703]]

**B6: The normalized representation of garryphotos - 34 points per face:**

[[-14.7597 -9.38295] [-6.3019 -9.67617] [14.0964 -11.0445] [3.64847 -10.2626] [-9.45285 -9.87165] [11.4429 -11.1423] [-10.7796 -10.0671] [9.61869 -11.4355] [-11.6088 -10.849] [-11.6088 -9.18747] [-11.6088 -14.2699] [9.61869 -15.736] [-6.63358 -2.15026] [6.96526 -2.05252] [0.0 0.0] [-7.1311 5.47339] [11.1112 4.30052] [0.663358 3.32313] [0.829198 5.27791] [1.16088 6.15756] [0.663358 5.27791] [2.48759 12.5106] [-34.6605 10.2626] [24.7101 3.81182] [0.829198 4.98469] [0.829198 4.98469] [0.829198 4.98469] [22.0567 -11.1423] [-22.0567 -9.57843] [0.829198 4.98469] [-1.82423 -23.1642] [9.28701 -12.1196] [9.45285 -10.1649] [-5.14103 -14.0744] [2.32175 -14.2699]]

[[-15.0173 -9.16975] [-5.15744 -9.4624] [14.8656 -10.633] [4.70237 -10.5355] [-8.03954 -9.95015] [10.6183 -11.3159] [-10.1632 -10.0477] [8.64629 -11.6085] [-10.4666 -11.7061] [-10.6183 -8.97465] [-10.3149 -15.3154] [9.70812 -16.486] [-6.52264 -1.46326] [6.06758 -2.04856] [0.0 0.0] [-7.43278 4.97508] [9.70812 3.70692] [0.455068 3.41427] [0.455068 4.58488] [0.303379 7.31629] [0.606758 5.56038] [2.57872 13.2669] [-16.8375 1.75591] [17.596 2.92652] [0.455068 4.58488] [0.455068 4.68242] [21.5399 -10.7306] [-20.3264 -8.97465] [0.455068 4.77998] [-2.12365 -22.5342] [8.34292 -13.0718] [8.49461 -10.4379] [-3.94392 -15.0228] [2.12365 -14.7301]]

[[-14.0201 -9.22774] [-5.33373 -9.53533] [14.1725 -10.4581] [4.57177 -10.4581] [-8.22918 -9.84292] [11.4294 -11.3809] [-10.2103 -9.84292] [8.83875 -11.1758] [-9.9055 -11.3809] [-9.9055 -8.92015] [-10.9722 -17.4302] [10.3627 -18.1479] [-5.63852 -0.820243] [7.31483 -2.15314] [0.0 0.0] [-6.09569 6.04929] [10.2103 4.20375] [1.37153 3.79362] [1.67632 4.61387] [1.67632 8.71508] [1.06675 7.27966] [2.13349 15.8922] [-17.2203 0.307591] [17.6775 0.10253] [1.37153 5.33158] [1.67632 5.84423] [21.0301 -10.8682] [-19.9634 -9.12521] [1.82871 5.63917] [-0.761961 -22.7617] [8.68636 -12.7138] [9.44832 -9.94545] [-3.50502 -15.7897] [3.04785 -15.9947]]

[[-17.0929 -8.4465] [-4.85797 -9.61154] [17.6326 -10.7766] [5.0379 -10.3882] [-9.53601 -10.097] [13.1345 -10.9707] [-11.875 -10.3882] [10.6156 -11.6503] [-11.5152 -10.8737] [-12.055 -9.32028] [-12.5947 -13.5921] [10.2557 -15.2425] [-7.73677 -2.03881] [7.73677 -2.52424] [0.0 0.0] [-8.27654 4.36888] [12.055 3.30093] [0.0 4.46597] [-0.179925 5.33974] [0.539774 6.99021] [0.0 5.631] [2.69887 13.3008] [-29.3277 3.4951] [24.1099 1.16503] [0.179925 4.07762] [0.179925 4.07762]

**B6: The normalized representation of garryphotos - 34 points per face:**

[25.5493 -10.8737] [-24.1099 -8.15524] [0.719699 4.2718] [-0.899624 -22.1357] [10.2557 -12.1358] [10.2557 -10.4853] [-3.41857 -14.8542] [1.4394 -14.7571]]

[[-17.3212 -8.54369] [-6.77102 -9.53714] [15.4316 -10.3319] [3.14931 -10.3319] [-9.44794 -9.83517] [9.92034 -10.928] [-11.8099 -9.93452] [8.03075 -11.0273] [-11.6525 -10.3319] [-11.8099 -9.23911] [-11.3375 -13.6103] [6.45609 -15.3985] [-7.87329 -1.78821] [5.66876 -3.37774] [0.0 0.0] [-9.44794 4.56988] [9.29047 3.0797] [0.314931 2.68232] [0.472397 3.37774] [0.787328 7.35155] [0.314931 5.86137] [1.88959 14.3057] [-22.0452 -0.19869] [17.1638 -0.794762] [1.10226 4.86792] [1.10226 4.86792] [21.7303 -10.0339] [-22.2027 -8.04696] [0.472397 5.2653] [-2.99185 -22.8494] [8.03075 -11.9214] [7.71582 -10.4312] [-4.56651 -13.7096] [0.787328 -13.809]]

[[-11.8289 -9.71601] [-4.78471 -9.71601] [12.2276 -10.6593] [5.18344 -10.282] [-6.91125 -10.1877] [9.43652 -10.848] [-9.03779 -10.0933] [7.17707 -10.9423] [-9.56943 -10.282] [-9.70234 -9.62168] [-9.1707 -13.4892] [6.64544 -14.6212] [-6.24671 -2.64124] [6.37962 -3.58455] [0.0 0.0] [-6.77834 3.77321] [7.5758 2.54691] [0.0 2.35825] [0.0 3.01857] [0.664543 6.22579] [0.398726 4.71651] [1.19618 12.1686] [-13.5567 4.0562] [17.8098 2.73557] [0.0 4.15053] [0.0 4.15053] [19.5376 -9.999] [-17.411 -8.58405] [0.531635 4.24486] [-1.06327 -22.0733] [7.44289 -11.4139] [7.44289 -10.0933] [-3.18981 -12.8289] [1.32909 -12.7346]]

] -> normgarrylist;

# Appendix C:

# Validation of the rule base

## C.1: Pair-wise chi-square test: Janus included:

Quantitative validation of the rule base: Statistical analysis of multiple-choice selections: pair-wise comparison of experts9-12 including Janus(J): Aggreement/Disagreement in Face Actions in each of six Feature Areas (brows, eyes, nose, mouth,cheeks and jaw) summed over seventeen face photographs: agreement includes actions absent as well as present: chi-square test.

Key: A = agree, D-A = disagree, E = expected frequencies

BROWS

| EXP | 9-10 | 9-11 | 9-12 | 10-11 | 10-12 | 11-12 | J-9 | J-10 | J-11 | J-12 | total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 63 | 56 | 64 | 60 | 58 | 61 | 65 | 62 | 60 | 65 | 614 |
| E | 61.4 | 61.4 | 61.4 | 61.4 | 61.4 | 61.4 | 61.4 | 61.4 | 61.4 | 61.4 | |
| D-A | 22 | 29 | 21 | 25 | 27 | 24 | 20 | 23 | 25 | 20 | 236 |
| E | 23.6 | 23.6 | 23.6 | 23.6 | 23.6 | 23.6 | 23.6 | 23.6 | 23.6 | 23.6 | |
| | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 850 |

$\chi^2 = 4.72$; d.f. = 9; $p \cong 0.85$

EYES:

| EXP | 9-10 | 9-11 | 9-12 | 10-11 | 10-12 | 11-12 | J-9 | J-10 | J-11 | J-12 | total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 94 | 94 | 101 | 98 | 98 | 90 | 86 | 85 | 88 | 88 | 922 |
| E | 92.2 | 92.2 | 92.2 | 92.2 | 92.2 | 92.2 | 92.2 | 92.2 | 92.2 | 92.2 | |
| D-A | 25 | 25 | 18 | 21 | 21 | 29 | 33 | 34 | 31 | 31 | 268 |
| E | 26.8 | 26.8 | 26.8 | 26.8 | 26.8 | 26.8 | 26.8 | 26.8 | 26.8 | 26.8 | |
| | 119 | 119 | 119 | 119 | 119 | 119 | 119 | 119 | 119 | 119 | 119 |

$\chi^2 = 15.177$; d.f. = 9; $p \cong 0.08$

NOSE:

| EXP | 9-10 | 9-11 | 9 12 | 10-11 | 10-12 | 11-12 | J-9 | J-10 | J-10 | J-12 | total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 25 | 22 | 22 | 26 | 29 | 24 | 26 | 23 | 20 | 23 | 240 |
| E | 24.0 | 24.0 | 24.0 | 24.0 | 24.0 | 24.0 | 24.0 | 24.0 | 24.0 | 24.0 | |
| D-A | 9 | 12 | 12 | 8 | 5 | 10 | 8 | 11 | 14 | 11 | 100 |
| E | 10.0 | 10.0 | 10.0 | 10.0 | 10.0 | 10.0 | 10.0 | 10.0 | 10.0 | 10.0 | |
| | 34 | 34 | 34 | 34 | 34 | 34 | 34 | 34 | 34 | 34 | 340 |

$\chi^2 = 8.5$; d.f. = 9; $p \cong 0.48$

## C.1: Pair-wise chi-square test: Janus included:

MOUTH:

| EXP | 9-10 | 9-11 | 9-12 | 10-11 | 10-12 | 11-12 | J-9 | J-10 | J-11 | J-12 | total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 164 | 162 | 155 | 156 | 166 | 162 | 162 | 156 | 155 | 159 | 1597 |
| E | 159.7 | 159.7 | 159.7 | 159.7 | 159.7 | 159.7 | 159.7 | 159.7 | 159.7 | 159.7 | |
| D-A | 57 | 59 | 66 | 65 | 55 | 59 | 59 | 65 | 66 | 62 | 613 |
| E | 61.3 | 61.3 | 61.3 | 61.3 | 61.3 | 61.3 | 61.3 | 61.3 | 61.3 | 61.3 | |
| | 221 | 221 | 221 | 221 | 221 | 221 | 221 | 221 | 221 | 221 | 2210 |

$\chi^2 = 3.31$; d.f.= 9; $p \cong 0.95$

CHEEKS:

| EXP | 9-10 | 9-11 | 9-12 | 10-11 | 10-12 | 11-12 | J-9 | J-10 | J-11 | J-12 | total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 40 | 36 | 33 | 39 | 32 | 34 | 35 | 37 | 33 | 33 | 352 |
| E | 35.2 | 35.2 | 35.2 | 35.2 | 35.2 | 35.2 | 35.2 | 35.2 | 35.2 | 35.2 | |
| D-A | 11 | 15 | 18 | 12 | 19 | 17 | 16 | 14 | 18 | 18 | 158 |
| E | 15.8 | 15.8 | 15.8 | 15.8 | 15.8 | 15.8 | 15.8 | 15.8 | 15.8 | 15.8 | |
| | 51 | 51 | 51 | 51 | 51 | 51 | 51 | 51 | 51 | 51 | 510 |

$\chi^2 = 6.20$; d.f. = 9; $p \cong 0.72$

JAW:

| EXP | 9-10 | 9-11 | 9-12 | 10-11 | 10-12 | 11-12 | J-9 | J-10 | J-11 | J-12 | total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 11 | 12 | 8 | 12 | 8 | 9 | 11 | 8 | 10 | 13 | 102 |
| E | 10.2 | 10.2 | 10.2 | 10.2 | 10.2 | 10.2 | 10.2 | 10.2 | 10.2 | 10.2 | |
| D-A | 6 | 5 | 9 | 5 | 9 | 8 | 6 | 9 | 7 | 4 | 68 |
| E | 6.8 | 6.8 | 6.8 | 6.8 | 6.8 | 6.8 | 6.8 | 6.8 | 6.8 | 6.8 | |
| | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 170 | 170 |

$\chi^2 = 7.74$; d.f. = 9; $p \cong 0.55$

## C.2: Pair-wise chi-square test: Janus excluded:

Quantitative validation of the rule base: Statistical analysis of multiple-choice selections: pair-wise comparison of experts9-12 excluding Janus(J): Aggreement/Disagreement in Face Actions in each of six Feature Areas (brows, eyes, nose, mouth, cheeks and jaw) summed over seventeen face photographs: agreement includes actions absent as well as present: chi-square test.

Key: A = agree, D-A = disagree, E = expected frequencies

BROWS

| EXP | 9-10 | 9-11 | 9-12 | 10-11 | 10-12 | 11-12 | total |
|---|---|---|---|---|---|---|---|
| A | 63 | 56 | 64 | 60 | 58 | 61 | 362 |
| E | 61.4 | 61.4 | 61.4 | 61.4 | 61.4 | 61.4 | . |
| D-A | 22 | 29 | 21 | 25 | 27 | 24 | 148 |
| E | 23.6 | 23.6 | 23.6 | 23.6 | 23.6 | 23.6 | |
| | 85 | 85 | 85 | 85 | 85 | 85 | 510 |

$\chi^2 = 2.59$; d.f. = 5; $p \cong 0.76$

EYES:

| EXP | 9-10 | 9-11 | 9-12 | 10-11 | 10-12 | 11-12 | total |
|---|---|---|---|---|---|---|---|
| A | 94 | 94 | 101 | 98 | 98 | 90 | 575 |
| E | 95.8 | 95.8 | 95.8 | 95.8 | 95.8 | 95.8 | |
| D-A | 25 | 25 | 18 | 21 | 21 | 29 | 139 |
| E | 23.2 | 23.2 | 23.2 | 23.2 | 23.2 | 23.2 | |
| | 119 | 119 | 119 | 119 | 119 | 119 | 714 |

$\chi^2 = 4.12$; d.f. = 5; $p \cong 0.52$

NOSE:

| EXP | 9-10 | 9-11 | 9 12 | 10-11 | 10-12 | 11-12 | total |
|---|---|---|---|---|---|---|---|
| A | 25 | 22 | 22 | 26 | 29 | 24 | 148 |
| E | 24.0 | 24.0 | 24.0 | 24.0 | 24.0 | 24.0 | |
| D-A | 9 | 12 | 12 | 8 | 5 | 10 | 56 |
| E | 10.0 | 10.0 | 10.0 | 10.0 | 10.0 | 10.0 | |
| | 34 | 34 | 34 | 34 | 34 | 34 | 204 |

$\chi^2 = 5.38$; d.f. = 5; $p \cong 0.37$

## C.2: Pair-wise chi-square test: Janus excluded:

MOUTH:

| EXP | 9-10 | 9-11 | 9-12 | 10-11 | 10-12 | 11-12 | total |
|---|---|---|---|---|---|---|---|
| A | 164 | 162 | 155 | 156 | 166 | 162 | 965 |
| E | 160.8 | 160.8 | 160.8 | 160.8 | 160.8 | 160.8 | |
| D-A | 57 | 59 | 66 | 65 | 55 | 59 | 361 |
| E | 60.2 | 60.2 | 60.2 | 60.2 | 60.2 | 60.2 | |
| | 221 | 221 | 221 | 221 | 221 | 221 | 1326 |

$\chi^2$ = 2.210; d.f.= 5; p ≅ 0.82

CHEEKS:

| EXP | 9-10 | 9-11 | 9-12 | 10-11 | 10-12 | 11-12 | total |
|---|---|---|---|---|---|---|---|
| A | 40 | 36 | 33 | 39 | 32 | 34 | 214 |
| E | 35.6 | 35.6 | 35.6 | 35.6 | 35.6 | 35.6 | |
| D-A | 11 | 15 | 18 | 12 | 19 | 17 | 92 |
| E | 15.3 | 15.3 | 15.3 | 15.3 | 15.3 | 15.3 | |
| | 51 | 51 | 51 | 51 | 51 | 51 | 306 |

$\chi^2$ = 4.99; d.f. = 5; p ≅ 0.41

JAW:

| EXP | 9-10 | 9-11 | 9-12 | 10-11 | 10-12 | 11-12 | total |
|---|---|---|---|---|---|---|---|
| A | 11 | 12 | 8 | 12 | 8 | 9 | 60 |
| E | 10 | 10 | 10 | 10 | 10 | 10 | |
| D-A | 6 | 5 | 9 | 5 | 9 | 8 | 42 |
| E | 7 | 7 | 7 | 7 | 7 | 7 | |
| | 17 | 17 | 17 | 17 | 17 | 17 | 102 |

$\chi^2$ = 4.36; d.f. = 5; p ≅ 0.50

## C.3: Williams $\hat{I}n$ - a worked example:

The following table presents the judgement of Janus(0) and five human experts (1-5) on five photographs (Nos. 84, 90, 91, 38, 37 in PFA, Ekman, 1975), as to whether the face action 'mouth u_lip_raised' is present (T = present, F = absent):

| photo No. | experts | | | | | |
|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 | 5 |
| 1 | T | F | F | T | T | F |
| 2 | T | F | F | T | F | F |
| 3 | F | T | F | T | T | T |
| 4 | T | T | F | T | T | T |
| 5 | T | F | F | F | T | T |

Following Williams (1976 ), we calculate $\hat{P}$ (a,b) which represents the proportion of observed agreements between raters a and b.

$\hat{P}(0,1) = \hat{P}(0,2) = 1/5$, $\hat{P}(0,3) = \hat{P}(0,4) = 3/5$, $\hat{P}(0,5) = 2/5$.
$\hat{P}(1,2) = \hat{P}(1,3) = \hat{P}(1,4) = 3/5$, $\hat{P}(1,5) = 4/5$.
$\hat{P}(2,3) = \hat{P}(2,4) = 1/5$, $\hat{P}(2,5) = 2/5$.
$\hat{P}(3,4) = 3/5$, $\hat{P}(3,5) = 2/5$, $\hat{P}(4,5) = 4/5$.

The overall group agreement among judges 1 - 5 is given by,

$$\hat{P}_5 = 2\,[\,5(5-1)\,]^{-1} \sum_{i=1}^{4} \sum_{j=i+1}^{5} \hat{P}(i,j) = 0.52.$$

The overall group agreement of Janus with experts 1 - 5 is given by,

$$\hat{P} = 1/5 \sum_{i=1}^{n} \hat{P}(0,i) = 0.4$$

The statistic $\hat{I}_5$ is given by,

$$\hat{I}_5 = Po/P_5 = 0.4/0.52 = 0.77.$$

# Appendix D:

# Validation of the dynamic memory

## D.1 Qualitative validation:

### D.1.1. Unprimed memory, basic emotions:

This validation addresses the question of whether typical face-action configurations access the correct FaceMOP emotions - using the greatest fraction of face action (FAC) matches between input and Face_mop frame FAC as a Face_mop selection criterion.

Descriptions of typical face actions and associated validated common emotions were adapted from "Unmasking The Face" by Paul Ekman & Wallace V. Friesen(1984) by the writer and entered into Janus in random order and in systematic reduction in number. The system response (an emotion) in each case was compared with the 'UTF' emotion in order to evaluate the system's ability to cope with incomplete and random order input.

In order to defend against a charge of possible "designer bias", similar descriptions were extracted from 'UTF' by a small group of persons individually and used as input. The results of these evaluations were overwhelmingly correct so that statistical analysis was not carried out.

## RESULTS:

### 'Surprised' configuration =

brows raised(br)
eyes wide(ew)
eyes upper-lid-raised(ulr)
eyes lower-lid-lowered(lll)
mouth-open(mo)
mouth-slightly-open(so)
mouth-wide-open(w)
jaw-drop(jd)

| Groupings | *System response* |
|---|---|
| 9-groupings: | |
| [ew mo w ulr ew w so so ew] | surprised |
| 8-groupings: | |
| [jd ew mo lll lll lll ew mo ] | surprised |
| 7-groupings: | |

| | | |
|---|---|---|
| [lll ew mo ew mo ulr jd | ] | surprised |
| [jd mo w br jd mo ulr | ] | surprised |
| [jd ulr mo mo ew jd mo | ] | surprised |

6-groupings:

| | | |
|---|---|---|
| [w ulr ew w so so | ] | surprised |
| [lll ulr ulr mo br br | ] | surprised |

5-groupings:

| | | |
|---|---|---|
| [so jd ew ulr mo | ] | surprised |
| [lll lll lll ew mo | ] | surprised |
| [ew mo ew jd jd | ] | surprised |

4-groupings:

| | | |
|---|---|---|
| [w ulr ew w | ] | surprised |
| [so so ew so | ] | surprised |
| [lll ulr ulr mo | ] | surprised |

3-groupings:

| | | |
|---|---|---|
| [jd mo ulr | ] | surprised |
| [ew ew ew | ] | surprised |
| [mo jd ulr | ] | surprised |
| [mo mo ew | ] | surprised |

2-groupings:

| | | |
|---|---|---|
| [ulr mo | ] | afraid |
| [mw | ] | angry |
| [jd ew | ] | surprised |
| [mo lll | ] | surprised |
| [lll lll | ] | surprised |
| [ew mo | ] | surprised |
| [ew jd | ] | surprised |
| [jd mo | ] | surprised |
| [mw br | ] | surprised |

1-groupings:

| | | |
|---|---|---|
| [ulr | ] | afraid |
| [br | ] | surprised |
| [ew | ] | surprised |
| [lll | ] | surprised |
| [mo | ] | surprised |
| [jd | ] | surprised |
| [so | ] | surprised |
| [w | ] | surprised |

## 'Angry' configuration =

brows lowered(bl)
brows contracted(bc)
eyes narrowed(en)
eyes inlid_raised
eyes lower-lid-raised(llr)
eyes lower-lid-tensed(llt)
*eyes upper-lid-lowered(ull)*
eyes upper-lid-tensed(ult)
nose flared(nf)
mouth compressed(mc)
mouth open(mo)
mouth bared(mb)
mouth wide(w)
mouth l_lip_tensed(llipt)
mouth u_lip_tensed(ulipt)
mouth square(ms)
cheeks n_l_vert(nlv)

| Groupings | System response |
|---|---|
| 9-groupings: | |
| [ull mb llt llr mc llr mc ult mbn] | nngry |
| [mb mc ull bc mo w ull bc mc] | angry |
| [en ult mb llr ult ult mc mo llt ] | angry |
| [mb w en mo en mc mo llt en] | angry |
| [en bl llt en en en bc w mc ] | angry |
| 8-groupings: | |
| [ull bl bc ult bc llt mb ull ] | angry |
| [ult mb bc bl w ull llr bl ] | angry |
| [llt bc mb mo mc bl bc ult ] | angry |
| [bl mc w llt mb w mo llt ] | angry |
| [ull bl bc ult bl bc ult w ] | angry |
| 7-groupings: | |
| [ult mc mo llt mb w en ] | angry |
| [mo ew mc mo llt en bc ] | angry |
| [en bl llt en en en bc ] | angry |
| [w mc mo bc llt en mo ] | angry |
| [bc bl w en bl llr mo ] | angry |
| 6-groupings: | |
| [ult bl bc ult w ull ] | angry |
| [mb llt llr mc llr mc ] | angry |

| | | |
|---|---|---|
| [ult mb mb mc ull bc | ] | angry |
| [mo w ull bc mc en | ] | angry |
| [mo en ult mb llr ult | ] | angry |

5-groupings:

| | | |
|---|---|---|
| [bc bl w ull llr | ] | angry |
| [bl llt bc mb mo | ] | angry |
| [mc bl ult bc bl | ] | angry |
| [mc w llt mb w | ] | angry |
| [mo llt ull bl bc | ] | angry |

4-groupings:

| | | |
|---|---|---|
| [llt en mo bc | ] | angry |
| [bl w en bl | ] | angry |
| [llr mo ull bl | ] | angry |
| [bc ult bc llt | ] | angry |
| [mb ull ult mb | ] | angry |

3-groupings:

| | | |
|---|---|---|
| [mc en mo | ] | angry |
| [en ult mb | ] | angry |
| [llr ult ult | ] | angry |
| [mc mo llt | ] | angry |
| [mb w en | ] | angry |

2-groupings:

| | | |
|---|---|---|
| [en llt | ] | angry |
| [mb ult | ] | angry |
| [ull bc | ] | angry |
| [llt llr | ] | afraid |
| [llr llr | ] | disgusted |
| [bc llt | ] | angry |
| [bc llt | ] | angry |
| [en ull | ] | angry |
| [ull llt | ] | angry |
| [en bc | ] | angry |

1-groupings:

| | | |
|---|---|---|
| [bl | ] | angry |
| [bc | ] | angry |
| [en | ] | angry |
| [llr | ] | disgusted |
| [llt | ] | angry |
| [ull | ] | angry |
| [ult | ] | angry |
| [mc | ] | angry |

| | | |
|---|---|---|
| [mb | ] | angry |
| [mo | ] | angry |
| [w | ] | angry |

## 'Happy' configuration =

eyes lower-lid-raised(llr)
mouth up(mu)
mouth-open(mo)
mouth slightly_open(so)
mouth wide(w)
mouth upper-lip-raised(ulipr)
mouth-bared(mb)
mouth pulled(mp)
cheeks raised(cr)

| Groupings | | System response |
|---|---|---|
| 7-groupings: | | |
| [mu cr cr mb llr mp cr | ] | happy |
| [llr mb mu llr cr llb mb | ] | happy |
| [mu mo mu cr llr mb mp | ] | happy |
| 6-groupings: | | |
| [mu mo mb mp llr cr | ] | happy |
| 5-groupings: | | |
| [mu mp mo mp llr | ] | happy |
| [mp mp mb mu llr | ] | happy |
| [mp mo mb mb mp | ] | happy |
| [ mp mb mp mp cr | ] | happy |
| [mo llr mb mu mu | ] | happy |
| [llr mp mb mb mu | ] | happy |
| 4-groupings: | | |
| [cr mu mb mp | ] | happy |
| [mu mb llr mp | ] | happy |
| 3-groupings: | | |
| [mu mu mp | ] | happy |
| [mp mb llr | ] | happy |
| [mb mo mp | ] | happy |
| 2-groupings: | | |

| | | |
|---|---|---|
| [cr mu | ] | happy |
| [cr mu | ] | happy |
| [mu llr | ] | happy |
| [mu mb | ] | happy |
| [mp mo | ] | afraid |
| [mp llr | ] | afraid |
| [mb llr | ] | angry |
| [mo mu | ] | happy |
| [mu mo | ] | happy |
| [mb mo | ] | angry |

1-groupings:

| | | |
|---|---|---|
| [mp | ] | afraid |
| [llr | ] | disgusted |
| [mb | ] | angry |
| [mo | ] | angry |
| [mu | ] | happy |
| [cr | ] | surprised |
| [mb | ] | happy |

**<u>'Sad' configuration =</u>**

brows contracted(bc)
brows centre-raised(bcr)
brows in_raised
eyes lower-lid-raised(llr)
eyes inlid-raised(eir)
eyes down(ed)
mouth down(md)

| Groupings | | System response |
|---|---|---|
| 4-groupings: | | |
| [bcr ed llr llr | ] | sad |
| [bcr md bc ed | ] | sad |
| [md ed bcr ed | ] | sad |
| [bc ed bcr bcr | ] | sad |
| [md md llr ed | ] | sad |
| 3-groupings: | | |
| [ed bcr llr | ] | sad |
| [llr bc bc | ] | angry |
| [ed md bcr | ] | sad |
| [eir md bc | ] | sad |

| | | |
|---|---|---|
| [bc md eir | ] | sad |

2-groupings:

| | | |
|---|---|---|
| [ed md | ] | sad |
| [md md | ] | sad |
| [llr bc | ] | angry |
| [eir bc | ] | sad |
| [bcr ed | ] | sad |
| [eir md | ] | sad |
| [ed llr | ] | sad |
| [eir ed | ] | sad |
| [bcr md | ] | sad |
| [ed mo | ] | angry |
| [llr eir | ] | sad |
| [llr md | ] | sad |

1-groupings:

| | | |
|---|---|---|
| [ed | ] | sad |
| [md | ] | sad |
| [eir | ] | sad |
| [llr | ] | disgusted |
| [bc | ] | angry |
| [bcr | ] | sad |

**'Disgusted' configuration =**

brows lowered(bl)
eyes lower-lid-raised(llidr)
eyes narrowed(en)
nose screwed(ns)
nose flared(nf)
mouth open
mouth compressed
mouth l-lip-tensed
mouth u-lip-tensed
mouth lower-lip-everted(llipe)
mouth upper-lip-everted(ulipe)
mouth upper-lip-raised(ulipr)
mouth lower-lip-raised(llipr)
mouth lower-lip-lowered(llipl)
cheeks raised(cr)
cheeks n-l-vert(nlv)

**Groupings** **System response**

12-groupings:

| | | |
|---|---|---|
| [llipl ulipe nf ns ulipe en llip nf ulipe nf ulipe nf | ] | disgusted |
| [nlv cr ulipe llipe nf llipr llipl ns ulipr nf en llipr | ] | disgusted |
| [nlv bl ulipe bl llidr cr ulipr llipr llipr bl en cr | ] | disgusted |
| [bl llipl bl ulipe ulipe nlv ns en ulipr llipl cr llipl | ] | disgusted |
| [llipe llipl ns bl ulipe nlv llipr nf llipe llipe nf llidr | ] | disgusted |

11-groupings:

| | | |
|---|---|---|
| [ulipr nlv ns llidr bl nvl ulipe en ulipe en llipr | ] | disgusted |
| [nvl en nvl cr llidr cr ulipr llipe llidr nf llipl | ] | disgusted |
| [en en llidr ulidr llipe cr llipl nf llipe llidr ulipe | ] | disgusted |
| [ns nlv llipl nlv cr llipl ulipe ulipr llipl bl nlv | ] | disgusted |
| [llipe cr nf nlv llidr ns ns cr nf bl cr | ] | disgusted |

10-groupings:

| | | |
|---|---|---|
| [ulipe llipe llipe nlv llipl llipl bl llipl llidr ulipr | ] | disgusted |
| [ulipr bl nf nf bl llidr nlv llipe llipr cr | ] | disgusted |
| [en llipl nf llidr ns ulipe nlv nf ulipr llipl | ] | disgusted |
| [bl nlv llipe llidr ns cr ulipe nlv llipe llipl | ] | disgusted |

9-groupings:

| | | |
|---|---|---|
| [ns nf ns en llipr nlv llipr cr llipe | ] | disgusted |
| [llidr en ns en llidr nf nlv bl cr | ] | disgusted |
| [llidr nf llidr llipe cr cr nlv ulr ulr | ] | disgusted |
| [bl llipr ulipr llipe ulipr en nf ulipe bl | ] | disgusted |

8-groupings:

| | | |
|---|---|---|
| [ns llipe ulipe llipr en lipr en ulipr | ] | disgusted |
| [nvl llipr nf bl cr ulipe ns llipe | ] | disgusted |
| [ns nlv nlv en ulipe llr nlv cr | ] | disgusted |

| | | |
|---|---|---|
| [cr ulipe llipe ulipr nlv llipe en llipe | ] | disgusted |
| [nlv ns llipr ulipe nf nf ns cr | ] | disgusted |
| [nvl llipr nf bl cr ulipe ns llipl | ] | disgusted |

7-groupings:

| | | |
|---|---|---|
| [llidr llipe ulipr llidr cr bl bl | ] | disgusted |
| [llidr llipr ulipr en llipr ns llidr | ] | disgusted |
| [cr ulipe nf cr llipr ulipr nf | ] | disgusted |
| [ulipe llipl llipe ulipr ns nlv llipr | ] | disgusted |
| [ulipe cr nf ulipr nf nf ns | ] | disgusted |

6-groupings:

| | | |
|---|---|---|
| [llidr ulipr nlv llipe llipr ulipe | ] | disgusted |
| [ulipe llipe ns llipe llipe nf | ] | disgusted |
| [cr ns llipl llipl llipl ulidr | ] | disgusted |
| [bl ns llipr ulipe bl en | ] | disgusted |
| [en llipl ulipr ulipe ns llipl | ] | disgusted |

5-groupings:

| | | |
|---|---|---|
| [nf bl cr ns cr | ] | disgusted |
| [llipl en nlv llipl llipr | ] | disgusted |
| [llipl llidr nlv llidr llipe | ] | disgusted |
| [llipl llidr nf bl ulipe | ] | disgusted |
| [en en nlv bl lipl | ] | disgusted |

4-groupings:

| | | |
|---|---|---|
| [llipl nf nf nf | ] | disgusted |
| [ulipe en ulipe nlv | ] | disgusted |
| [llidr nf nlv llidr | ] | disgusted |
| [llipe cr llipl llipe | ] | disgusted |
| [llipe ns ns llidr | ] | disgusted |

3-groupings:

| | | |
|---|---|---|
| [cr llipe bl | ] | disgusted |
| [en ulipr ns | ] | disgusted |
| [nf cr llipr | ] | disgusted |
| [ns llidr ulipr | ] | disgusted |
| [cr llipe ns | ] | disgusted |

2-groupings:

| | | |
|---|---|---|
| [en bl | ] | angry |
| [llipe llidr | ] | disgusted |
| [ulipr llipr | ] | disgusted |
| [en nlv | ] | disgusted |

| | | |
|---|---|---|
| [llidr ns | ] | disgusted |
| [llipl llipe | ] | disgusted |
| [bl llipl | ] | disgusted |
| [llipe llipr | ] | disgusted |
| [en ns | ] | disgusted |
| [ulipe cr | ] | disgusted |
| [llipe llipe | ] | disgusted |
| [en bl | ] | angry |

1-groupings:

| | | |
|---|---|---|
| [ulipr | ] | disgusted |
| [llipe | ] | disgusted |
| [cr | ] | disgusted |
| [nlv | ] | disgusted |
| [ns | ] | disgusted |
| [en | ] | angry |
| [b | ] | angry |
| [llipl | ] | disgusted |
| [llipr | ] | disgusted |
| [llidr | ] | disgusted |
| [nf | ] | disgusted |
| [ulipe | ] | disgusted |

**'Afraid' configuration =**

brows raised(br)
brows in-raised(bir)
brows contracted(bc)
eyes lower-lid-raised(llidr)
eyes lower-lid-tensed(llidt)
eyes upper-lid-raised(ulidr)
eyes inlid-raised(ilidr)
mouth wide
mouth bared
mouth lower-lip-tensed(llipt)
mouth upper-lip-tensed(ulipt)
mouth pulled(mp)
mouth open(mo)

| Groupings | | System response |
|---|---|---|
| 15-groupings: | | |
| [br bc bc br br mp br llidt mp llipt ulidr ulidr br br llidt | ] | afraid |

9-groupings:

[llipt ulidr ulidr br br llidt bc
llipt bc ] afraid

[mp llidr llidt bc br mp br llidr
llipt ] afraid

8-groupings:

[llidr llidt ulipt br llipt llidt mo
llidt ] afraid

[mp llidt br llipt mp mo mp mo ] afraid

[ulidr mp br llidt ulidr mp mo mo] afraid

7-groupings:

[br llidr mo mp llidr mo ulidr ] afraid

[ mo mo ulidr ulipt llidr bc ] afraid

[llipt ulidr mo ulidr bc ulipt llipt ] afraid

[br br llipt mo bc bc ulidr ] afraid

[llidt ulidr llidt ulidr bc llidt llidt ] afraid

[ bc ulidr bc ulipt llidt mp
ulipt ] afraid

6-groupings:

[mp llipt ulipt llidr ulidr br ]. afraid

[ulipt br mo llidt mo bc ] afraid

5-groupings:

[mo br mp ulipt llipt ] afraid

4-groupings:

[llipt llipt bc br ] afraid

[bc llipt br ulipt ] afraid

[bc ulidr bc br ] afraid

[llidt llidt llipt mp ] afraid

[bc bc llidr llidr ] afraid

3-groupings:

[bc llipt llidt ] afraid

[llidt ulidr br ] afraid

[mo mo ulidr ] surprised

[ulidr br llipt ] afraid

2-groupings:

[ulidr llidt ] afraid

[ulipt ulidr ] afraid

| | | |
|---|---|---|
| [llidr br | ] | afraid |
| [ulidr llipt | ] | afraid |
| [ulidr llipt | ] | afraid |
| [bc llidr | ] | afraid |
| [mo mo | ] | angry |
| [mo llidr | ] | afraid |
| [llidt llidt | ] | afraid |

1-groupings:

| | | |
|---|---|---|
| [mp | ] | afraid |
| [ulipt | ] | afraid |
| [ulidr | ] | surprised |
| [mp | ] | afraid |
| [br | ] | afraid |
| [bc | ] | angry |
| [llidt | ] | afraid |
| [llidr | ] | disgusted |
| [mo | ] | angry |

## D.1.2: System learning and retrieval:

**Validation of system learning and retrieval: .Once the trail down a branch of the tree has been blazed by an input configuration of face actions in "LEARN" mode, the same input in RETRIEVE mode can retrace the path and access the emotion labels of the leaf. This allows access to the emotion associated with the separate events which traversed that path.**

'+' (in these descriptions) means 'with the addition of'; '-' means 'without the inclusion of'; 'int.' means interpretation; 'FAC' means face action components.

**a) database unprimed at outset.**

| system mode | face actions input | system output | user suggested interpretation |
|---|---|---|---|
| RETRIEVE | brows raised<br>brows wrinkled<br>eyes upper_lid_raised<br>eyes lower_lid_neutral<br>mouth upper_lip_neutral<br>mouth_upper_lip_lowered<br>cheeks raised | Surprised | Shocked |
| LEARN | FAC as above<br>+ int. = shocked | | |
| RETRIEVE | FAC as above | Shocked<br>under Surprised | |

Comment: Satisfactory retrieval of learned interpretation, demonstrating that the same new additions to the path have been retraced.

**b) database primed as above.**

| | input | output | |
|---|---|---|---|
| RETRIEVE | brows low<br>*eyes* u_lid_lowered<br>eyes *l_lid_neutral*<br>nose neutral<br>mouth u_lip_drawn<br>mouth l_lip_lowered<br>cheeks neutral | *Disgusted* | 1. Vicious<br>*2. Angry* |
| LEARN | FAC as above<br>+ int. = vicious | | |
| RETRIEVE | FAC as above | Vicious<br>under Disgusted | |

Comment: Although the input did not self-select the 'Angry' branch (because characteristic actions in eyes and mouth were not recorded by the lay-expert, the system showed correct retrieval of the learned interpretation, demonstrating that the same additions to the path have been retraced.

**c) database primed as above.**

| | input | | output | user |
|---|---|---|---|---|
| RETRIEVE | brows | low | | |
| | eyes | u_lid_lowered | | |
| | eyes | l_lid_neutral | | |
| | nose | neutral | Happy | 1. Slightly-amused |
| | mouth | pulled | | 2. Interested |
| | cheeks | raised | | |
| LEARN | FAC as above<br>+ int. = slightly-amused | | | |
| RETRIEVE | FAC as above | | Slightly-amused under Happy | |

Comment: The system showed correct retrieval of the learned interpretation, demonstrating that the same additions to the path have been retraced.

**d) database primed as above.**

| | input | | output | user |
|---|---|---|---|---|
| RETRIEVE | brows | neutral | | |
| | eyes | u_lid_lowered | | |
| | eyes | l_lid_neutral | | |
| | nose | neutral | Disgusted | 1. Disgusted |
| | mouth | u_lip_everted | | 2. Indifferent |
| | mouth | l_lip_lowered | | |
| LEARN | FAC as above<br>+ int. = Indifferent | | | |
| RETRIEVE | FAC as above | | 1. Indifferent or<br>2 vicious<br>under Disgusted | |

Comment: Although the System interpretation: "Disgusted" agreed with the lay-expert's 1st choice, it is a default interpretation. The system was therefore taught the 2nd choice for this face description. The system showed correct retrieval of the learned interpretation, demonstrating that the same additions to the path have been retraced.

e)database primed as above.

| | input | output | user |
|---|---|---|---|
| RETRIEVE | brows low<br>brows straight<br>eyes u_lid_lowered<br>eyes l_lid_raised<br>eyes narrowed<br>nose screwed<br>mouth u_lip_drawn<br>mouth l_lip_everted<br>cheeks raised | Disgusted | 1.Contemptuous<br>2.Disgusted |
| LEARN | FAC as above<br>+ int. = contemptuous | | |
| RETRIEVE | FAC as above<br>-int. = contemptuous | Contemptuous<br>under Disgusted | |

Comment: The system showed correct retrieval of the learned interpretation, demonstrating that the same additions to the path have been retraced.

f)database primed as above.

| | input | output | user |
|---|---|---|---|
| RETRIEVE | brows raised<br>eyes wide<br>eyes u_lid_raised<br>eyes l_lid_lowered<br>mouth bared<br>cheeks raised<br>jaw drop | Surprised | 1.happy-surprised<br>2.gleeful-recognition |
| LEARN | FAC as above<br>+ int. = happy-surprised | | |
| RETRIEVE | FAC as above<br>-int. = happy-surprised | happy-surprised<br>under Surprised | |

Comment: The system showed correct retrieval of the learned interpretation, demonstrating that the same additions to the path have been retraced.

g)database primed as above.

| | input | output | user |
|---|---|---|---|

| | | | | |
|---|---|---|---|---|
| RETRIEVE | brows | raised | Surprised | Horror |
| | eyes | staring | | |
| | eyes | wide | | |
| | eyes | u_lid_raised | | |
| | eyes | l_lid_lowered | | |
| | eyes | protuberant | | |
| | mouth | bared | | |
| | mouth | u_lip_raised | | |
| | mouth | l_lip_lowered | | |
| LEARN | FAC as above | | | |
| | + int. = horror | | | |
| RETRIEVE | FAC as above | | Horror | |
| | - int. = horror | | under Surprised | |

Comment: Although the input did not self-select the Fear branch (because characteristic tense actions in eyes and mouth were not recorded by the lay-expert, the expression is a blend of surprise and fear actions and the system decided for the Surprised branch. The system showed correct retrieval of the learned interpretation, demonstrating that the same additions to the path have been retraced.

h) database primed as above.

| | input | | output | user |
|---|---|---|---|---|
| RETRIEVE | brows | low | | |
| | brows | contracted | | |
| | eyes | narrowed | Angry | 1.amused-crossed |
| | eyes | l_lid_raised | | 2.vindictive |
| | eyes | u_lid_lowered | | |
| | mouth | compressed | | |
| | mouth | up | | |
| | mouth | pulled | | |
| | nose | tightened | | |
| | cheeks | firmed | | |
| LEARN | FAC as above | | | |
| | + int. = amused-crossed | | | |
| RETRIEVE | FAC as above | | amused-crossed | |
| | | | under Angry | |
| | - int. = amused-crossed | | | |

Comment: The system showed correct retrieval of the learned interpretation, demonstrating that the same additions to the path have been retraced.

i) database primed as above. system user

| | | | | |
|---|---|---|---|---|
| RETRIEVE | brows | low | | |
| | brows | contracted | | |
| | eyes | narrowed | Angry | 1.amused-crossed |
| | | | | 2.vindictive |
| | eyes | l_lid_raised | | |
| | eyes | u_lid_lowered | | |
| | mouth | compressed | | |
| | mouth | up | | |
| | mouth | pulled | | |
| | nose | tightened | | |
| | cheeks | firmed | | |
| LEARN | as above | | | |
| | + int. = vindictive | | | |
| RETRIEVE | as above | | amused-crossed | |
| | - int. = vindictive | | vindictive | |
| | | | under Angry | |

Comment: The system showed correct retrieval of the learned interpretation, demonstrating that the same additions to the path have been retraced.

| j)database primed as above. | | | system | user |
|---|---|---|---|---|
| RETRIEVE | | | | |
| | eyes | smiling | | 1.amiable |
| | eyes | l_lid_raised | | 2.happy-relaxed |
| | eyes | u_lid_lowered | Happy | |
| | mouth | shut | | |
| | mouth | up | | |
| | cheeks | relaxed | | |
| LEARN | as above | | | |
| | + int. = amiable | | | |
| RETRIEVE | as above | | amiable | amiable |
| | - int. = amiable | | under Happy | |
| LEARN | as above | | | |
| | + int. = happy-relaxed | | | |
| RETRIEVE | as above | | amiable | |
| | - int. = amiable | | happy-relaxed | |
| | - int. = happy-relaxed | | under happy | |

Comment: The system showed correct retrieval of the learned interpretation, demonstrating that the same additions to the path have been retraced.

| k) database primed as above. | | system | user |
|---|---|---|---|
| RETRIEVE | brows low<br>brows contracted<br>nose flared<br>eyes l_lid_tensed<br>eyes u_lid_lowered<br>mouth compressed<br>mouth down<br>mouth shut<br>nose n_l_groove<br>jaw tensed | Angry | 1.very-angry<br>2.furious |
| LEARN | as above<br>+ int. = furious | | |
| RETRIEVE | as above<br>- int. = amused-crossed | furious<br>under Angry | furious |

Comment: The system showed correct retrieval of the learned interpretation, demonstrating that the same additions to the path have been retraced.

## CONCLUSION:

The above examples illustrate the system's ability to learn new face actions from idiosyncratic perceptions of specific expressions and interpretations and to order these in Memory in the existing tree branches. Furthermore provided that some "given" face actions are included in an input description, the examples show that when no precedent for the rest of the input face actions exist (and when, therefore, there is nothing in the System's experience available as an interpretation), the System interprets from prototypic matching. Inspection of the database confirms that the appropriate submops are established when a retraced path brings similar components into juxta-position. These abstractive concepts provide learning and possible prediction on a different level: it is an automatic organizing of the knowledge extracted from experience so that it is immediately available to assist in processing new input in both RETRIEVE and LEARN mode (it is "remembered" when similar input occurs).

The database changes produced by the above learning are shown below in a comparison: before versus after.

**The initial database:**

[[mop_angry_gen_spec][]m13]
[[mop_afraid_gen_spec][]m12]

```
[[mop_disgusted_gen_spec][]m11]
[[mop_sad_gen_spec][]m10]
[[mop_happy_gen_spec][]m8]
[[mop_surprised_gen_spec][]m7]
[[][raised m8 n_l_vert m11 raised m11] m6]
[[] [drop m7 tense m11]m5]
[[] [drop m7 smile m8 open m7 open m12 open m13 compressed m13
  u_lip_raised m11 l_lip_lowered m11 l_lip_lowered m11 pulled m12
  down m10 down m11 shut m8 slightly_open m7 wide m7 wide m13
  l_lip_everted m11 u_lip_everted m11 l_lip_raised m11
  u_lip_tensed m12 l_lip_tensed m12 square m13
  open m8 open m12 open m13 bared m8 up m8 bared m7 bared m13]m4]
[[] [screwed m11 flared m11 flared m13 ]m3]
[[] [  narrowed m11 wide m7 wide m12 l_lid_tensed m13 l_lid_tensed   m12
  u_lid_tensed m13 l_lid_raised m13 u_lid_lowered m13
  l_lid_raised m10  u_lid_lowered m8  narrowed m13
  l_lid_lowered m7 u_lid_raised m12 l_lid_raised m12 wide m7
  u_lid_raised m7 l_lid_raised m11 in_lid_raised m10 down m10 ]m2]
[[] [raised m7  raised m10 low m11 low m13
  raised m12 contracted m12 centre_raised m10 contracted m10
  contracted m13 straight m12]m1]
[[] [brows m1 eyes m2 nose m3 mouth m4 jaw m5 cheeks m6 ]top1]]
 -> database;
```

**Explanation:** each sublist is a record with two fields: [node content information] and [associated links in the tree]. The record ID is the node to which these refer, thus:
[[info][links]top1]. "Top1" is the top node in the tree.

**The database after the inputs in the validation above is:**

```
[[[ev6 ev120] [] m102]
[[mop_happycheeks_relaxed_sub] [event m102] m101]
[[mop_happycheeks_fsub] [relaxed m101] m100]
[[ev6 ev3 ev2 ev1 ev0] [] m49]
[[ev6 ev120] [] m99]
[[mop_happyeyes_smiling_sub] [event m99] m98]
[[ev5] [] m116]
[[] [event m116] m115]
[[] [eyes m115] m114]
[[mop_happy_gen_spec]
 [smiling m114 cheeks m100 nose m52 eyes m47 brows m44]
 m8]
[[ev123] [] m113]
[[jaw tensed] [event m113] m112]
[[jaw] [tensed m112] m111]
```

[[mop_angry_gen_spec]
[jaw m111 cheeks m41 mouth m36 nose m33 eyes m30 brows m27]
m13]
[[ev123] [] m110]
[[nose n_l_groove] [event m110] m109]
[[mop_angrynose_fsub]
[n_l_groove m109 flared m103 tightened m34] m33]
[[ev123] [] m108]
[[mouth down] [event m108] m107]
[[mop_angrymouth_fsub] [down m107 shut m105 pulled m39 up m37]
m36]
[[ev123] [] m106]
[[mouth shut] [event m106] m105]
[[ev123] [] m104]
[[nose flared] [event m104] m103]
[[ev120 ev3 ev2 ev1 ev0] [] m54]
[[mop_happyeyes_fsub]
[smiling m98 u_lid_raised m96 narrowed m94 l_lid_neutral m50
u_lid_lowered m48] m47]
[[ev120] [] m97]
[[eyes u_lid_raised] [event m97] m96]
[[ev120] [] m95]
[[eyes narrowed] [event m95] m94]
[[ev118] [] m93]
[[eyes protuberant] [event m93] m92]
[[mop_surprisedeyes_fsub] [protuberant m92 l_lid_slack m19] m18]
[[ev118] [] m91]
[[nose n_l_groove] [event m91] m90]
[[mop_surprisednose_fsub] [n_l_groove m90 relaxed m22] m21]
[[ev118] [] m89]
[[mouth l_lip_lowered] [event m89] m88]
[[mop_surprisedmouth_fsub]
[l_lip_lowered m88 down m86 u_lip_slack m84 up m25] m24]
[[ev118] [] m87]
[[mouth down] [event m87] m86]
[[ev118] [] m85]
[[mouth u_lip_slack] [event m85] m84]
[[ev118] [] m83]
[[cheeks dropped] [event m83] m82]
[[mop_surprisedcheeks_fsub] [dropped m82 raised m80] m79]
[[ev116 ev115] [] m81]
[[cheeks raised] [event m81] m80]
[[mop_surprised_gen_spec]
[cheeks m79 mouth m24 nose m21 eyes m18 brows m15] m7]
[[ev113] [] m78]
[[mouth pulled] [event m78] m77]

[[mop_disgustedmouth_fsub]
[pulled m77 compressed m75 shut m67] m66]
[[ev113] [] m76]
[[mouth compressed] [event m76] m75]
[[ev113] [] m74]
*[[eyes u_lid_lowered]* [event m74] m73]
*[[mop_disgustedeyes_fsub] [u_lid_lowered m73* down m59] m58]
[[ev113] [] m72]
[[brows straight] [event m72] m71]
[[mop_disgustedbrows_fsub]
[straight m71 lowered m69 neutral m64 centre_raised m56] m55]
[[ev113] [] m70]
[[brows lowered] [event m70] m69]
[[ev4] [] m68]
[[mouth shut] [event m68] m67]
[[mop_disgusted_gen_spec]
[mouth m66 cheeks m61 eyes m58 brows m55] m11]
[[ev4] [] m65]
[[brows neutral] [event m65] m64]
[[ev4] [] m63]
[[] [event m63] m62]
[[] [n_l_groove m62] m61]
[[ev4] [] m60]
[[eyes down] [event m60] m59]
[[ev4] [] m57]
[[brows centre_raised] [event m57] m56]
[[ev110 ev107] [] m43]
[[ev110 ev107] [] m40]
[[mop_angrymouth_pulled_sub] [event m40] m39]
[[ev110 ev107] [] m38]
[[mop_angrymouth_up_sub] [event m38] m37]
[[ev110 ev107] [] m35]
[[mop_angrynose_tightened_sub] [event m35] m34]
[[ev110 ev107] [] m32]
[[mop_angryeyes_wide_sub] [event m32] m31]
[[mop_angryeyes_fsub] [wide m31] m30]
[[ev110 ev107] [] m29]
[[mop_angrybrows_lowered_sub] [event m29] m28]
[[mop_angrybrows_fsub] [lowered m28] m27]
[[ev3 ev2 ev1 ev0] [] m51]
[[ev3 ev2 ev1 ev0] [] m46]
[[mop_happynose_neutral_sub] [event m54] m53]
[[mop_happynose_fsub] [neutral m53] m52]
[[mop_happyeyes_l_lid_neutral_sub] [event m51] m50]
[[mop_happyeyes_u_lid_lowered_sub] [event m49] m48]
[[mop_happybrows_low_sub] [event m46] m45]

```
[[mop_happybrows_fsub] [low m45] m44]
[[] [event m43] m42]
[[] [firm m42] m41]
[[ev105 ev104 ev102 ev100] [] m26]
[[mop_surprisedmouth_up_sub] [event m26] m25]
[[ev105 ev104 ev102 ev100] [] m23]
[[mop_surprisednose_relaxed_sub] [event m23] m22]
[[ev105 ev104 ev102 ev100] [] m20]
 [[mop_surprisedeyes_l_lid_slack_sub] [event m20] m19]
[[ev105 ev104 ev102 ev100] [] m17]
[[mop_surprisedbrows_wide_apart_sub] [event m17] m16]
[[mop_surprisedbrows_fsub] [wide_apart m16] m15]
[[mop_afraid_gen_spec] [] m12]
[[mop_sad_gen_spec] [] m10]
[[] [raised m8 n_l_vert m11 raised m11] m6]
[[] [drop m7 tense m11] m5]
[[] [drop m7 smile m8 open m7 open m12 open m13 compressed m13
        u_lip_raised m11 l_lip_lowered m11 l_lip_lowered m11
        pulled m12 down m10 down m11 shut m8 slightly_open m7
        wide m7 wide m13 l_lip_everted m11 u_lip_everted m11
        l_lip_raised m11 u_lip_tensed m12 l_lip_tensed m12
        square m13 open m8 open m12 open m13 bared m8 up m8
        bared m7 bared m13] m4]
[[] [screwed m11 flared m11 flared m13] m3]
[[] [narrowed m11 wide m7 wide m12 l_lid_tensed m13 l_lid_tensed
          m12 u_lid_tensed m13 l_lid_raised m13 u_lid_lowered
          m13 l_lid_raised m10 u_lid_lowered m8 narrowed m13
          l_lid_lowered m7 u_lid_raised m12 l_lid_raised m12
          wide m7 u_lid_raised m7 l_lid_raised m11
          in_lid_raised m10 down m10] m2]
[[] [raised m7 raised m10 low m11 low m13 raised m12 contracted
        m12 centre_raised m10 contracted m10 contracted m13
        straight m12] m1]
[[] [brows m1 eyes m2 nose m3 mouth m4 jaw m5 cheeks m6] top1]]
```

**Explanation:** One branch below m7 (mop_surprised_gen_spec) is via the link: "brows" to m15 (mop_surprisedbrows_fsub) via "wide_apart" to m16 and then via "event" to the leaf node: m17 which stores referents to four separate input events: ev100, ev102, ev104, and ev105. From these, the four interpretations (of the same or different lay-experts), tagged to an identical or very similar configuration of perceived face-actions, can be retrieved to offer a user as an interpretation for a similar input.

## D.2 Quantitative validation of Janus:

### D.2.1 Unprimed memory:

For the purpose of this evaluation, FOUR lay-experts and Janus judged 17 garryphotos. The lay-experts were College academics in the Computer and Information Technology Department who were aware of the purposes of the system. Two of each gender were approached on the basis of availability only. Judging each photograph in turn, with reference to a forced - choice constraint of the following:

"I think he is feeling:
Sad Happy Angry Afraid Surprised Disgusted (Ring 1st choice)."

Their responses are listed below alongside those of Janus, and form the information on which the following statistical tests were carried out:

1. The KAPPA statistic: to assess the agreement between the lay-experts with and without Janus, hence any significant difference.

2. Williams' In statistic: to test Janus's agreement with the set of lay-experts acting as standard, taking into account the agreement between each standard.

3. Friedman's ANOVA: testing that each lay-expert's and Janus's ratings (attributed by other "meta-judges" on their agreement with the lay-experts' and Janus's assignations) could come from the same population. In this procedure, the meta-judges were unaware of the fact that one of the sources of the information was a computer.

The source judgments on which these statistics were performed are shown in Table D1:

Table D.1: Judgements on 17 garryphotos by Janus and 4 experts:

| photo | Janus | Expert A | Expert B | Expert C | Expert D |
|---|---|---|---|---|---|
| 1 | Happy | Sad | Happy | Disgusted | Sad |
| 3 | Afraid | Afraid | Surprised | Surprised | Surprised |
| 4 | Afraid | Sad | Surprised | Happy | Happy |
| 5 | Happy | Happy | Happy | Happy | Happy |
| 6 | Angry | Surprised | Surprised | Angry | Sad |
| 7 | Disgusted | Disgusted | Disgusted | Sad | Disgusted |
| 8 | Afraid | Angry | Angry | Angry | Disgusted |
| 9 | Afraid | Afraid | Afraid | Angry | Afraid |
| 10 | Disgusted | Sad | Disgusted | Happy | Disgusted |
| 11 | Surprised | Happy | Surprised | Happy | Happy |
| 12 | Happy | Happy | Happy | Happy | Happy |
| 13 | Sad | Sad | Sad | Sad | Sad |
| 14 | Afraid | Afraid | Surprised | Angry | Surprised |
| 15 | Surprised | Surprised | Surprised | Angry | Surprised |
| 16 | Sad | Sad | Sad | Sad | Sad |
| 17 | Angry | Angry | Angry | Angry | Angry |
| 18 | Disgusted | Disgusted | Disgusted | Surprised | Disgusted |

## Table D.2: The Kappa Statistic (Cohen 1960 & 1968)

For this purpose the above information is represented in a frequency table drawn under the six common emotions upon which it is based:

| photo | Happy | Afraid | Disgusted | Surprised | Sad | Angry |
|---|---|---|---|---|---|---|
| 1 | 2 | 0 | 1 | 0 | 2 | 0 |
| 3 | 0 | 2 | 0 | 3 | 0 | 0 |
| 4 | 2 | 1 | 0 | 1 | 1 | 0 |
| 5 | 5 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 2 | 1 | 2 |
| 7 | 0 | 0 | 4 | 0 | 1 | 0 |
| 8 | 0 | 1 | 1 | 0 | 0 | 3 |
| 9 | 0 | 4 | 0 | 0 | 0 | 1 |
| 10 | 1 | 0 | 3 | 0 | 1 | 0 |
| 11 | 3 | 0 | 0 | 2 | 0 | 0 |
| 12 | 5 | 0 | 0 | 0 | 0 | 0 |
| 13 | 0 | 0 | 0 | 0 | 5 | 0 |
| 14 | 0 | 2 | 0 | 2 | 0 | 1 |
| 15 | 0 | 0 | 0 | 4 | 0 | 1 |
| 16 | 0 | 0 | 0 | 0 | 5 | 0 |
| 17 | 0 | 0 | 0 | 0 | 0 | 5 |
| 18 | 0 | 0 | 4 | 1 | 0 | 0 |

### D.2.1.1. The Kappa statistic:

Denoting by $n_{ij}$, the number of judges assigning the $i^{th}$ photograph to the $j^{th}$ category (i.e. $n_{13} = 1$), we define the column totals $C_j$ by

$$C_j = \sum_{i=1}^{N} n_{ij}, \qquad \text{(N = no. of photographs = 17).} \qquad (1)$$

The proportion of objects assigned to the $j^{th}$ category is given by:

$$P_j = C_j / Nk, \qquad \text{(k = no. of judges = 5).} \qquad (2)$$

The Kappa coefficient is given by (Cohen 1968),

$$K = \frac{P(A) - P(E)}{1 - P(E)}, \qquad (3)$$

$$P(E) = \sum_{j=1}^{m} P_j^{\,2}, \qquad \text{(m = no. of categories = 6).} \qquad (4)$$

$$P(A) = \left( \frac{1}{Nk(k-1)} \cdot \sum_{i=1}^{N} \sum_{j=1}^{m} n_{ij}^{\,2} \right) - \frac{1}{k-1} \qquad (5)$$

For the $n_{ij}$ given in the table,

$$P(E) = .1720 \,, \qquad (6)$$

Denoting by $f(l)$ the number of $n_{ij}$ equal to $l$, we obtain,

$$f(1) = 66,\ f(2) = 14 \ldots f(5) = 5. \qquad (7)$$

The double sum in (5) is given by,

$$\sum_{l=0}^{5} l^2 f(l) = 275. \qquad (8)$$

P(A) is then given by (using 5) ,

$$P(A) = \frac{275}{17 \times 5 \times 4} - \frac{1}{4} = 0.559. \qquad (9)$$

The coefficient is thus,

$$K = \frac{.559 - .172}{1 - .172} = 0.467. \quad (10)$$

The variance of K can be calculated using,

$$Var(k) = \frac{2}{Nk(k-1)} \cdot \frac{P(E) - (2k-3)[P(E)]^2 + 2(k-2) \ \Sigma P_i^3}{[1 - P(E)]^2} \quad (11)$$

$$= 0.001266 \ ,$$

from which we derive ,

$$Z = \frac{K}{\sqrt{var(K)}} = 13.13.$$

## D.2.1.2 Williams' $\hat{I}_n$ Statistic:

For the calculation of this statistic, the attributed emotions are coded by integers as follows: *Happy* = 1; Sad = 2; Angry = 3; Disgusted = 4; Surprised = 5 and Afraid = 6. The following table (Table D.3) of the experts' assignations are so coded in the leftmost column for each photo and it is upon these codings that the statistic is worked. Agreements between experts then is facilitated and counts of such over the photograph range is a matter of counting matches in these integers since they always are ordered horizontally thus:Janus, expert1 - expert4.

### Table D.3 Coded interpretations from Janus and four lay-experts for 17 garryphotos:

| codes | Janus | Expert 1 | *Expert 2* | Expert 3 | Expert 4 | photo |
|---|---|---|---|---|---|---|
| 12142 | Happy | Sad | *Happy* | Disgusted | Sad | 1 |
| 66555 | Afraid | Afraid | Surprised | Surprised | Surprised | 3 |
| 62511 | Afraid | Sad | Surprised | Happy | Happy | 4 |
| 11111 | Happy | Happy | Happy | Happy | Happy | 5 |
| 35532 | Angry | Surprised | Surprised | Angry | Sad | 6 |
| 44424 | Disgusted | Disgusted | Disgusted | Sad | Disgusted | 7 |
| 63334 | Afraid | Angry | Angry | Angry | Disgusted | 8 |
| 66636 | Afraid | Afraid | Afraid | Angry | Afraid | 9 |
| 42414 | Disgusted | Sad | Disgusted | Happy | Disgusted | 10 |
| 51511 | Surprised | Happy | Surprised | Happy | Happy | 11 |
| 11111 | Happy | Happy | Happy | Happy | Happy | 12 |
| 22222 | Sad | Sad | Sad | Sad | Sad | 13 |
| 66535 | Afraid | Afraid | Surprised | Angry | Surprised | 14 |
| 55535 | Surprised | Surprised | Surprised | Angry | Surprised | 15 |
| 22222 | Sad | Sad | Sad | Sad | Sad | 16 |
| 33333 | Angry | Angry | Angry | Angry | Angry | 17 |
| 44454 | Disgusted | Disgusted | Disgusted | Surprised | Disgusted | 18 |

$\hat{P}(0,1)$ = 11/17 = agreement between Janus and expert1

$\hat{P}(0,2)$ = 12/17 = agreement between Janus and expert 2

$\hat{P}(0,3)$ = 6/17 = agreement between Janus and expert 3

$\hat{P}(0,4)$ = 10/17 = agreement between Janus and expert 4

$\hat{p}(1,2)$ = 11/17 = " " expert1 and expert 2

$\hat{p}(1,3)$ = 7/17 = " " expert1 and expert 3

$\hat{p}(1,4)$ = 11/17 = " " expert1 and expert 4

$\hat{p}(2,3)$ = 7/17 = " " expert 2 and expert 3

$\hat{p}(2,4)$ = 12/17 = " " expert 2 and expert 4

$\hat{p}(3,4)$ = 8/17 = " " expert 3 and expert 4

The agreement among experts1-4 ($\hat{P}_n$)

$= 2/(4\times3) \times ( \hat{p}(1,2)+\hat{p}(1,3)+\hat{p}(1,4)+\hat{p}(2,3)+\hat{p}(2,4)+\hat{p}(3,4) )$

= 0.549

The agreement between Janus and experts1-4 ($\hat{P}0$)

$= 1/(4) \times ( \hat{p}(0,1)+\hat{p}(0,2)+\hat{p}(0,3)+\hat{p}(0,4) )$

= 0.5735

$\hat{I}_4$ is the ratio of $\hat{P}0 / \hat{P}n$

= 1.04

**Estimation of the confidence intevals for Williams $\hat{I}_n$**

Following Williams (1976), we *calculate* an upper bound for the population $\hat{I}_n$ at the 5% confidence level. We make the assumption that $\hat{P}_0$ and $\hat{P}_4$ are asymptotically normally distributed, and calculate their variances (V0, Vn) and covariance (Cv) as detailed in the paper. The quantity, Z, given by,

$$Z = \frac{\hat{P}_0 - I_4 \hat{P}_n}{[V_0 - 2(\hat{I}_4)Cv + \hat{I}_n \quad V_n]^{1/2}}$$

is distributed as a normal deviate.

For the 5% significance level, Z = 1.645 (1-sided). Using the calculated P0, Pn, V0, Cv and Vn, we can solve the above equation to get an upper bound for the population $\hat{I}_4$. We summarise the results in Table D.4.

**Table D.4 Variances and covariances for Janus and four lay-experts' interpretations of 17 garryphotos:**

| Test Expert | $\hat{P}_0$ | $\hat{P}_n$ | $V_0$ | $C_v$ | $V_n$ |
|---|---|---|---|---|---|
| Janus | .574 | .549 | .0077 | .0056 | .0061 |
| Expert 1 | .588 | .539 | .0073 | .0056 | .0063 |
| Expert 2 | .618 | .520 | .0061 | .0056 | .0067 |
| Expert 3 | .412 | .657 | .0104 | .0030 | .0084 |
| Expert 4 | .603 | .529 | .00695 | .0054 | .0066 |
| 'Disgusted' Janus | .1471 | .5490 | .0039 | -.0013 | .0061 |
| Random Janus | .1617 | .5490 | .00603 | .0024 | .0061 |

### D.2.1.3 Friedman's Anova:

**Use of meta-judges in a 'blind' rating of emotions attributed by Janus and four lay-experts to the same set of photographs. Friedman's ANOVA was used to test that such ratings could come from the same population.**

In this procedure, the meta-judges were unaware of the fact that one of the sources of the information was a computer. They were asked to rate each member of a quintet of interpretations for each photograph in turn using the following semantic codes:
A = "What I would have said"
B = "Not what I would have said but I can see it's there"
C = "I can not see it at all"

Three "meta-judges" - actually different lay-experts - were selected on the basis of availability without intended bias. They were aged 31, 34,and 35 years against the test-judges' 39, 35, 26 and 24 years. They were executives with university degrees or chartered professional status with no special training in reading faces. They rated the interpretations for a photograph with the same photograph and a neutral face photograph before them. The interpretations of the four lay-experts and Janus were the same in the above studies. These ratings are tabulated in Table D.5.

Their ratings were processed in the following way: the number of meta-judges who agreed that each assignation merited either of codes A or B was calculated for each lay-expert and for Janus. The results are tabulated in Table D.6, together with the ordering(in brackets) across each photograph required for the Friedman Anova - the test applied to decide whether these columns of ordinal-based ratings could come from the same population - if they can, the rationale that Janus's judgements are on a par with those of human lay - experts, is supported.

Table D.5 Three meta-judges' ratings of interpretations of 17 garryphotos.
The ratings (r) shown are the better of first and second(if any) interpretations of each lay-expert compared to the interpretation by the meta-judge of the same photograph:

'r' = three meta-judges ratings of the same photograph

| photo no. | Janus | r | expert1 | r | expert2 | r | expert3 | r | expert4 | r |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | happy | bab | sad | cbc | happy | bab | disgusted | ccc | sad | cbc |
| 3 | afraid | baa | afraid | baa | afraid | baa | afraid | baa | surprised | bba |
| 4 | afraid | ccc | sad | bcc | surprised | bcb | happy | bbb | happy | bba |
| 5 | happy | aaa | happy | aaa | happy | aaa | happy | aaa | happy | aaa |
| 6 | angry | cca | disgusted | bcb | sad | bac | afraid | bbb | sad | cab |
| 7 | disgusted | aaa | disgusted | aaa | disgusted | aaa | sad | aaa | disgusted | aaa |
| 8 | afraid | aab | afraid | aab | angry | acb | angry | acb | afraid | baa |
| 9 | afraid | aab | afraid | aab | afraid | aab | angry | bcc | afraid | baa |
| 10 | disgusted | bba | pained | aaa | disgusted | bba | disgusted | bba | disgusted | abb |
| 11 | surprised | bbb | happy | aab | surprised | bbb | happy | aab | happy | baa |
| 12 | happy | aaa | happy | aaa | happy | aaa | happy | aaa | happy | aaa |
| 13 | sad | aaa | sad | aaa | sad | aaa | sad | aaa | sad | aaa |
| 14 | afraid | baa | afraid | baa | afraid | baa | disgusted | ccb | surprised | bba |
| 15 | surprised | bab | afraid | aab | afraid | aab | angry | ccb | afraid | baa |
| 16 | sad | baa | sad | baa | sad | baa | sad | baa | sad | aab |
| 17 | angry | aaa | angry | aaa | angry | aaa | angry | aaa | angry | aaa |
| 18 | disgusted | aaa | disgusted | aaa | disagusted | aaa | disgusted | aaa | disgusted | aaa |

**Table D.6 Meta-judges' consensus of 'A' & 'B' ratings on interpretations of 17 garryphotos ranked for the Friedman Anova:**

| photo | Janus | Expert 1 | Expert 2 | Expert 3 | Expert 4 | $\Sigma(t^3-t)$ |
|---|---|---|---|---|---|---|
| 1 | 3 (4.5) | 1 (2.5) | 3 (4.5) | 0 (1.0) | 1 (2.5) | 12 |
| 3 | 3 (3.0) | 3 (3.0) | 3 (3.0) | 3 (3.0) | 3 (3.0) | 120 |
| 4 | 0 (1.0) | 1 (2.0) | 2 (3.0) | 3 (4.5) | 3 (4.5) | 6 |
| 5 | 3 (3.0) | 3 (3.0) | 3 (3.0) | 3 (3.0) | 3 (3.0) | 120 |
| 6 | 1 (1.0) | 2 (3.0) | 2 (3.0) | 3 (5.0) | 2 (3.0) | 24 |
| 7 | 3 (3.0) | 3 (3.5) | 3 (3.5) | 3 (1.0) | 3 (3.5) | 120 |
| 8 | 3 (4.0) | 3 (4.0) | 2 (1.5) | 2 (1.5) | 3 (4.0) | 30 |
| 9 | 3 (3.5) | 3 (3.5) | 3 (3.5) | 1 (1.0) | 3 (3.5) | 60 |
| 10 | 3 (3.0) | 3 (3.0) | 3 (3.0) | 3 (3.0) | 3 (4.0) | 120 |
| 11 | 3 (3.0) | 3 (3.0) | 3 (3.0) | 3 (3.0) | 3 (3.0) | 120 |
| 12 | 3 (3.0) | 3 (3.0) | 3 (3.0) | 3 (3.0) | 3 (3.0) | 120 |
| 13 | 3 (3.0) | 3 (3.0) | 3 (3.0) | 3 (3.0) | 3 (3.0) | 120 |
| 14 | 3 (3.5) | 3 (3.5) | 3 (3.5) | 1 (1.0) | 3 (3.5) | 60 |
| 15 | 3 (3.5) | 3 (3.5) | 3 (3.5) | 1 (1.0) | 3 (3.5) | 60 |
| 16 | 3 (3.0) | 3 (3.0) | 3 (3.0) | 3 (3.0) | 3 (3.0) | 120 |
| 17 | 3 (3.0) | 3 (3.0) | 3 (3.0) | 3 (3.)) | 3 (3.0) | 120 |
| 18 | 3 (3.0) | 3 (3.0) | 3 (3.0) | 3 (3.0) | 3 (3.0) | 120 |
| R | (51.0) | (52.0) | (52.5) | (45.0) | (54.5) | 1452= T |
| R x R | 2601 | 2704 | 2756.25 | 2025 | 3136 | |

$\Sigma(R^2)$ = 13036.5 = H

$$\text{test statistic, } \chi^2_r = \frac{\frac{12(H)}{NK(K+1)} - 3N(K+1)}{1 - \frac{T}{N(K^3 - K)}}$$

$$\chi^2_r = \frac{\frac{12 \times 13036.5}{17 \times 5 \times 6} - 3 \times 17 \times 6}{1 - \frac{1452}{17 \times 120}}$$

$$\chi^2_r = 2.5694,\ d.f. = 4,\ p \cong 0.5$$

The n.s. result supports the hypothesis that the scores could be drawn from the same population. The inference is therefore supported that the meta-judges as a group do not discriminate overall in their agreement with the lay - experts and Janus.

The above studies support the claim that under the conditions stated and with the one set of face photographs of the one actor portraying the six common emotions: happy, sad, disgusted, angry, afraid and surprised, Janus makes attributions which agree with those of human lay - experts.

## D.2.2 Primed memory:

Janus's interpretations used in the above tests were those of the "Unprimed" system, i.e. the system which contained only the "given" face actions(FAC) typical each of the common emotions: happy, sad, angry, disgusted, afraid and surprised. Janus's knowledge was augmented by teaching it new expressions labelled with emotion. 50 persons judged expressions on a series of face photographs of a different actor than was used in the development of the system and the knowledge was taught to Janus, resulting in a "Primed" version of Janus. In order to evaluate this primed version, the FAC for these taught expressions are now re-entered in RETRIEVE mode in search of Janus-produced interpretations for each of the 17 photographs used in the unprimed evaluations above. Janus responds by outputting the "given", general emotion for each and the leaf node interpretations resulting from the priming, depending on whether the priming FAC were fully subsumed by the Face_mop Frame FAC or not. For each of the 17 photographs then, there are a variable number of "primed" interpretations resulting from this procedure and independent judges were requested to rate the latter as "good", "fair" or "poor" in comparison with their own judgements of these photographs. In fact these raters were not informed that the primed or unprimed interpretations had anything to do with a computer and the photograph was presented to them with the request that they should tell the emotion portrayed by the expression. This was recorded but played no part in the procedure that followed, They were told that "Other people" shown this photograph had said it was "such and such"(the retrieved labels) and asked to rate each of these as a good, fair or poor interpretation. Apart from avoiding possible computer bias, this apparent subterfuge is not so far from the truth, since only unprimed interpretations are built into Janus a priori. Primed interpretations of course originate from users in the knowledge acquisition, and Janus's memory is automatically organized by their associated FAC. Each judge rated up to three sets of primed interpretations. An example completed rating sheet is reproduced in Table D.7. (g) = good, (f) = fair, (p) = poor ratings:

Table D.7 User-rated interpretations of primed memory (ratings in brackets) - a specimen sheet

| photo no. | User-judged attributes | Janus-judged attributes + User's rating | |
|---|---|---|---|
| 4 | humoured/happy | surprised(f); receptive to argument(f) interested(p). | |
| 13 | glum | sad(g); having distaste(p) depressed(g) | unhappy(p) |
| 14 | shocked | afraid(g); puzzled(f); incomprehending(p); fearful(p) | |

reproduced in Table D.7. (g) = good, (f) = fair, (p) = poor ratings: The 'user-judged attributes' column contains the rater's judgement of the current photograph and is ignored. The system interpretations for the same photograph are on the right and contain the single Face_MOP "given" emotion and a varying number of learned interpretations which the rater has marked with a (g), (f), or (p).
It is easy to compare the rating for the Face_mop interpretation with that of the learned ones using these ordered ratings. Some raters rate 3 faces, some less, and this accounts for there being a total of 149 FaceMOP interpretations from 55 raters. Of all the 149 photographs rated in this validation method, the analysis reveals that 105(70.5%) received a 'good', 35(23.5%), a 'fair' and 9(6.0%), a 'poor' rating, when Face_mop and learned interpretations were considered together, giving a combined 'good and fair' rating to almost 94% overall. The percentage of these rating categories *per photograph* follows:

| photo no | good | fair | poor | total | total% all |
|---|---|---|---|---|---|
| 1 | 9 | 0 | 1 | 10 | 6.6 |
| 3 | 7 | 1 | 0 | 8 | 5.2 |
| 4 | 1 | 3 | 0 | 4 | 2.6 |
| 5 | 7 | 0 | 0 | 7 | 4.7 |
| 6 | 6 | 4 | 4 | 14 | 9.3 |
| 7 | 10 | 2 | 0 | 12 | 8.05 |
| 8 | 4 | 4 | 1 | 9 | 6.04 |
| 9 | 8 | 1 | 0 | 9 | 6.04 |
| 10 | 3 | 3 | 0 | 6 | 4.03 |
| 11 | 5 | 1 | 0 | 6 | 4.03 |
| 12 | 3 | 3 | 0 | 6 | 4.03 |
| 13 | 3 | 0 | 0 | 3 | 2 |
| 14 | 9 | 3 | 0 | 12 | 8.05 |
| 15 | 6 | 1 | 0 | 7 | 4.6 |
| 16 | 4 | 3 | 2 | 9 | 6.04 |
| 17 | 7 | 2 | 1 | 10 | 6.6 |
| 18 | 13 | 4 | 0 | 17 | 11.3 |

The complete ratings follow:

rater c.k.: photo 8: afraid(p), puzzled(f),incomprehending(g),fearful(p).
photo 5: happy(g),cheerful(f),anticipating pleasure(p)
photo 3: afraid(p),puzzled(f),fearful(f),incomprehending(g).

rater y.a.: photo16: sad(p),depressed(p),miserable(p).
photo 12: happy(g),anticipating pleasure(g).
photo 8: afraid(f),puzzled(p),incomprehending(p),fearful(p).

rater m.i.: photo 7: disgustedg,disliking(f),displeased(f).
photo 14: afraid(g), puzzled(p),incomprehending(p),fearful(p).

rater c.p. :photo 14: afraid(f),puzzled(p),incomprehending(f),fearful(g).
photo 6: angry(g),disliking(g),glowering(f),hostile(g),disagreeable(f).
photo 16: sad(f),depressed(g),miserable(f).

rater b.b. :photo 16: sad(f),depressed(g),miserable(g).
photo 6: angry(f),disliking(p),glowering(f),hostile(g),disagreeable(p).
photo 10: disgusted(f),disliking(f),displeased(f).

rater j.j. : photo 18: disgusted(f),disliking(p),displeased(f)
photo 11: surprised(g),receptive to argument(f),interested(f).

rater c.i. photo 1: happy(f),cheerful(p),anticipating pleasure(g).

photo 18: disgusted(f),disliking(g),displeased(p).
photo16: sad(f),depressed(f),miserable(f).
raterd.b.: photo12: happy(f),anticipating pleasure(f)..
photo 9: afraid(p),incomprehending(f),puzzled(f),fearful(p).
photo 1: cheerful(g),happy(f),anticipating pleasure(f).
rater u.k.: photo 7: disgusted(7),disliking(p),displeased(p).
photo 10: disgusted(g),displeased(g),disliking(g).
photo 9: afraid(g),puzzled(p),incomprehending(f),fearful(g).
photo 6: angry(f),disliking(f).
photo18: dislike(g),disgusted(g),displeased(f).
rater j.b.: photo 7: disgusted(g),disliking(g),displeased(g).
photo 7: disgusted(g),disliking(g),displeased(g).
photo 3: afraid(g),puzzled(f),incomprehending(p),fearful(g).
rater v.s. photo 18: disgusted(g),disliking(g),displeased(g).
photo 14: afraid(p),puzzled(g),incomprehending(g),fearful(p).
photo 17: angry(p),disliking(f).
rater s.s.: photo 17: angry(p),disliking(p).
photo 14: afraid(g),puzzled(p),incomprehending(p),fearful(g).
photo 3: afraid(p),puzzled(g),incomprehending(p),fearful(g).
rater s.b.: photo 9: afraid(g),incomprehending(p),puzzled(g),fearful(p).
photo 17: angry(g),disliking(g).
rater s.r. photo 17: angry(f),disliking(f).
photo 9: afraid(g),puzzled(g),incomprehending(g),fearful(f).
photo 15: surprised(g),receptive to argument(f),interested(f0.
rater m.i photo 16: sad(g),depressed(f),miserable(g).
photo 4; surprised(g),receptive to argument(f),interested(f).
rater d.p. photo 6: angry(f),disliking(p).
photo 9: afraid(g),puzzled(p),incomprehending(f),fearful(g).
photo 7: disgusted(g),disliking(g),displeased(f).
rater j.s. photo 5: happy(f),cheerful(f),anticipating pleasure(f).
photo 18: disgusted(f),disliking(p),displeased(p).
rater m.w. photo 14: afraid(f),puzzled(f),incomprehending(f),fearful(p).
photo 16: sad(g),depressed(g),miserable(f).
photo 17: angry(g),disliking(p).
rater r.b. photo 4: surprised(p),receptive to argument(f),interested(p).
photo 6: angry(g),disliking(g),glowering(g),hostile(g).
photo 3: afraid(g),fearful(g),puzzled(p),incomprehending(f).
rater r.b. photo 7: disgusted(f),disliking(f),displeased(f).
photo 15: surprised(f),receptive to argument(f),interested(p).
photo 14: afraid(f),puzzled(g),incomprehending(f),fearful(f).
rater u.k. photo 11: surprise(f),receptive to argument(f),interested(p).
rater p.m. photo 6: angry(p),disliking(p).
photo 7: disgusted(g),disliking(f),displeased(f).
photo 1: cheerful(f),anticipating pleasure(f),happy(f).
rater g.a. photo 8: afraid(g),puzzled(p),incomprehending(p),fearful(g).
photo 18: disgusted(f),displeased(f).
rater m.g. photo 6: angry(f),disliking(f).
photo 1: cheerful(g),anticipating pleasure(g),happy(g).
rater j.s. photo 12: happy(g),anticipating pleasure(g).
photo 8: afraid(p),puzzled(f),incomprehending(g),fearful(p).
photo 9: afraid(g),incomprehending(f),puzzled(p).
rater i.l. photo 12: happy(f),anticipating pleasure(f).
photo 15: surprised(g),receptive to argument(p),interested(p).
photo 7: disgusted(g),disliking(g),displeased(g).

rater p.b. photo 17: angry(g),disliking(g).
photo 12: happy(f),anticipating pleasure(f).
photo 15: surprised(g),receptive to argument(p),interested(p).
rater r.g. photo 1: cheerful(f),anticipating pleasure(g),happy(g).
photo 15: surprised(g),receptive to argument(f),interested(f).
photo 10: disgusted(p),disliking(f),displeased(p).
rater k.s. photo 11: surprised(g),interested(g),receptive to argument(f).
photo 8: afraid(p),puzzled(g),incomprehending(g),fearful(p).
photo14: afraid(g),puzzled(f),incomprehending(f),fearful(g).
rater g.b. photo 11: surprised(g),interested(g),receptive to argument(p).
photo 7: disgusted(g),displeased(g),disliking(g).
photo 1: cheerful(f),happy(g),anticipating pleasure(g).
rater t.k. photo 6: angry(p),disliking(g).
photo 18: disgusted(f),disliking(p),displeased(p).
rater t.z. photo 5: happy(g),anticipating pleasure(g),cheerful(f).
photo 11: surprised(g),receptive to argument(p),interested(f)
rater n.k. photo 3: afraid(p),puzzled(p),incomprehending(p),fearful(f).
photo 6: angry(f),disliking(g).
rater f.m. photo 11: surprised(f),interested(g),receptive to argument(p).
photo 18: disgusted(f),displeased(g),disliking(g).
rater m.r. photo 6: angry(p),disliking(p).
photo 16: sad(f),depressed(p),miserable(p).
photo 3: afraid(f),incomprehending(g),puzzled(g),fearful(p).
rater j.m. photo 15: surprised(g),receptive to argument(f),interested(f).
photo 16: sad(f),depressed(p),miserable(p).
photo 4: surprised(p),fearful(p),puzzled(f),incomprehending(f).
rater a.a. photo 8: afraid(f),puzzled(f),incomprehending(p),fearful(f).
photo 5: happy(g),anticipating pleasure(p),cheerful(g).
photo 6: angry(p),disliking(p).
rater n.s. photo 18: disgusted(g),disliking(f),displeased(f).
photo 12: happy(g),anticipating pleasure(p).
photo 3: afraid(g),puzzled(p),incomprehending(p),fearful(g).
rater s.s. photo 1: cheerful(p),anticipating pleasure(p),happy(p).
photo 14: afraid(f),incomprehending(p),puzzled(p),fearful(p).
photo 10: disgusted(g),disliking(g),displeased(f).
rater s.k. photo 14: afraid(g),puzzled(p),incomprehending(p),fearful(g).
photo 6: angry(p),disliking(f),glowering(p),hostile(f),disagreeable(g).
photo 9: afraid(g),puzzled(p),incomprehending(p),fearful(g).
rater l.m. photo 13: sad(g),depressed(g),unhappy(f),having distaste(p).
photo 18: disgusted(g),displeased(p),disliking(p).
rater j.p. photo 15: surprised(g),receptive to argument(f),interested(f).
photo 5: happy(g),cheerful(f),anticipating pleasure(g).
photo 17: angry(g),disliking(f),displeased(f)
rater j. photo 9: afraid(g),puzzled(p),incomprehending(f),fearful(g).
photo 18: displeased(g),disgusted(g),disliking(f).
photo 8: afraid(p),puzzled(p),incomprehending(p),fearful(p).
rater m. photo 14: afraid(g),puzzled(f),incomprehending(g),fearful(g).
photo 10: disgusted(g),disliking(f),displeased(f).
photo 8: afraid(f),puzzled(p),incomprehending(p),fearful(f).
rater j.o. photo 18: disgusted(g),disliking(f),displeased(g).
photo 1: happy(f),cheerful(g),anticipating(f) pleasure(g).
photo 5: happy(g),cheerful(g), anticipating(f),pleasure(g).
rater j.h. photo 17: angry(g),disliking(f),displeased(p).
photo 10: disgusted(f),disliking(p),displeased(f).

| | |
|---|---|
| | photo 7: disgusted(f),disliking(p),displeased(g). |
| rater s.t. | photo 5: happy(g),cheerful(g),anticipating(f),pleasure(f). |
| | photo 18: disgusted(g),disliking(f),displeased(g). |
| rater c. | photo 4: surprised(f),receptive to argument(f),interested(p). |
| | photo 13: sad(g),depressed(g),unhappy(p),having distaste(p). |
| | photo 14: afraid(g),puzzled(f),incomprehending(p),fearful(f). |
| rater d. | photo 18: disgusted(f),disliking(g),displeased(p). |
| | photo 7: disgusted(g),disliking(g),displeased(p0. |
| | photo17: angry(g),disliking(p),displeased(g). |
| rater jem | photo6: angry(p),disliking(p),glowering(p),hostile(p),disagreeable(p). |
| | photo 14: afraid(f),puzzled(p),incomprehending(p),fearful(f). |
| | photo 5: happy(g),cheerful(g),anticipating(p),pleasure(g). |
| rater s.t. | photo 7: disgusted(g),disliking(f),displeased(g). |
| | photo 18: disgusted(g),disliking(f),displeased(p). |
| | photo 8: afraid(p),puzzled(p),incomprehending(p),fearful(f). |
| rater v.i. | photo 6: angry(p),disliking(g),glowering(p),hostile(p),disagreeable(f). |
| | photo 9: afraid(g),puzzled(f),incomprehending(p),fearful(g). |
| | photo 3: afraid(p),puzzled(f),fearful(p),incomprehending(g). |
| rater c.a. | photo 1: happy(f),cheerful(p),anticipating(g),pleasure(f). |
| | photo 18: disgusted(g),disliking(g),displeased(g). |
| | photo 13: sad(g),depressed(f),unhappy(g),having distaste(f). |
| rater l.i. | photo 17: angry(f),disliking(g),displeased(g). |
| | photo 18: disgusted(g),disliking(f),displeased(g). |
| | photo 1: happy(f),cheerful(g),anticipating(f),pleasure(f). |

UNIVERSITY C
NWICH LIBRARY