INFORMATION COMPRESSION, INTELLIGENCE, COMPUTING, AND MATHEMATICS

Dr Gerry Wolff

CognitionResearch.org

OVERVIEW

- Information, redundancy, and compression of information.
- Information compression in brains and nervous systems.
- Information compression in computing and mathematics.

INFORMATION AND REDUNDANCY (1)

- Information: anything that contains recognisable variations may be seen as information—light waves, sound waves, pictures, language, music, etc.
- Redundancy = repetition of information.
- Any body of information, *I*, may be seen to comprise non-redundant and redundant information:

Non-redundant information	Redundant information

INFORMATION AND REDUNDANCY (2)

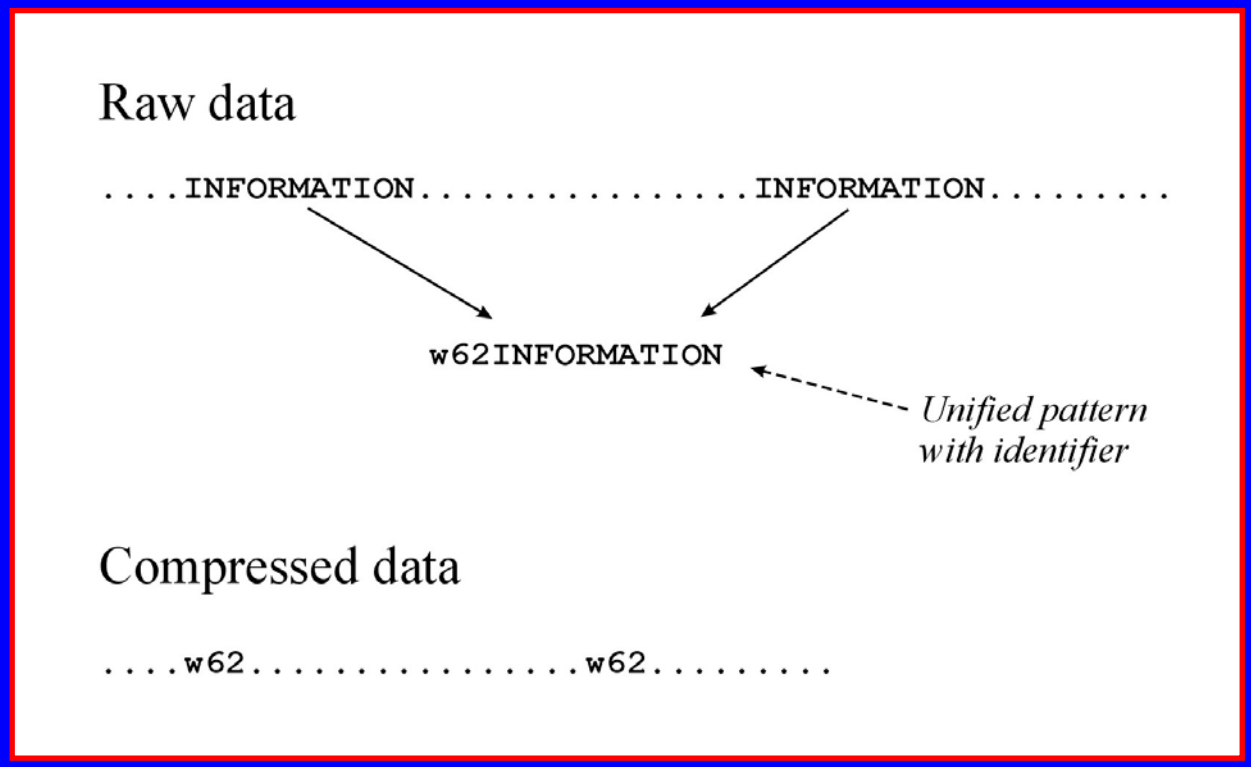
- Shannon's information theory: The communicative value of a symbol or other 'event' is related to its probability. There is redundancy in a body of information, *I*, if some symbol types are more probable than others.
- Algorithmic information theory: If a body of information, *I,* can be generated by a computer program that is shorter than *I* then the information is not random and contains redundancy.
- Redundancy as repetition of patterns:
  - Coherent patterns: I N F O R M A T I O N I N F O R M A T I O N
  - Discontinuous patterns:

    I N F a b O R M A c d e T I O N x y I N p q F O R r s M A T I t u O N

INFORMATION AND REDUNDANCY (3)

- In ‘redundancy as repetition of patterns’, there are two key variables:
  - The sizes of patterns.
  - The frequencies of patterns.
- Given the close connection between frequency and probability, there are also close connections between probability, redundancy, and compression.
- More generally, information compression and probabilistic inference may be seen as two sides of the same coin.

COMPRESSION OF INFORMATION BY THE MATCHING AND UNIFICATION OF PATTERNS



- The idea may be generalised to discontinuous patterns like

I N F a b O R M A c d e T I O N x y I N p q F O R r s M A T I t u O N

TECHNIQUES FOR COMPRESSING INFORMATION

- Chunking-with-codes: each repeating ‘chunk’ of information is given a short ‘code’.
- Schema-plus-correction: a generalised pattern is ‘corrected’ with choices at specific points, eg choices in a restaurant menu.
- Run-length coding: eg ‘I N F O R M A T I O N’ × 100.

  Cut out repetition and mark transitions from one type of pattern to another.

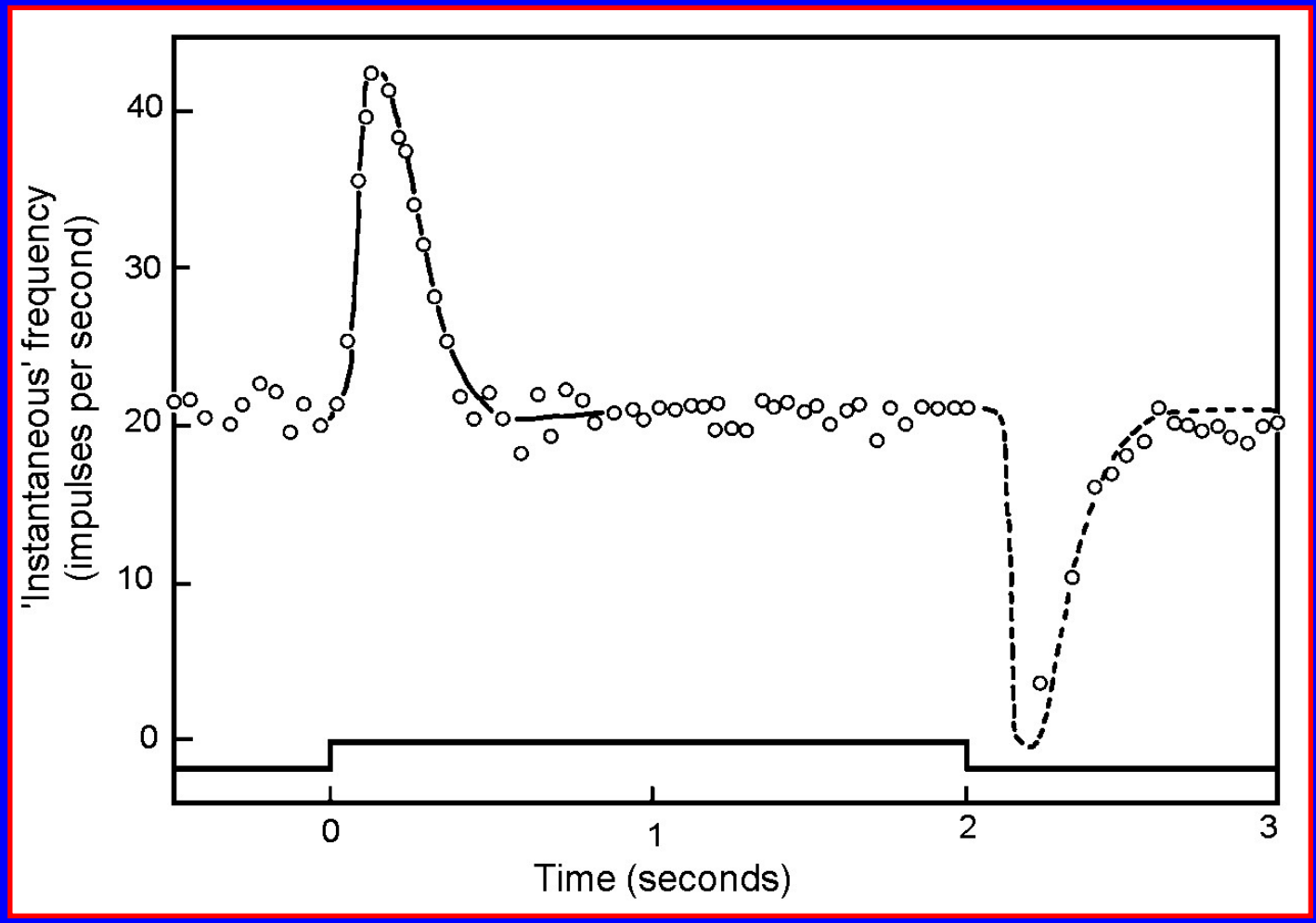
INFORMATION COMPRESSION AND NATURAL SELECTION

- Promoting economies in storage.
- Promoting efficiency and speed in the processing and transmission of information.
- Corresponding savings in energy (the brain is 2% of total body weight but it demands 20% of our resting metabolic rate).
- Perhaps more importantly, it is the key to predicting the future from the past.

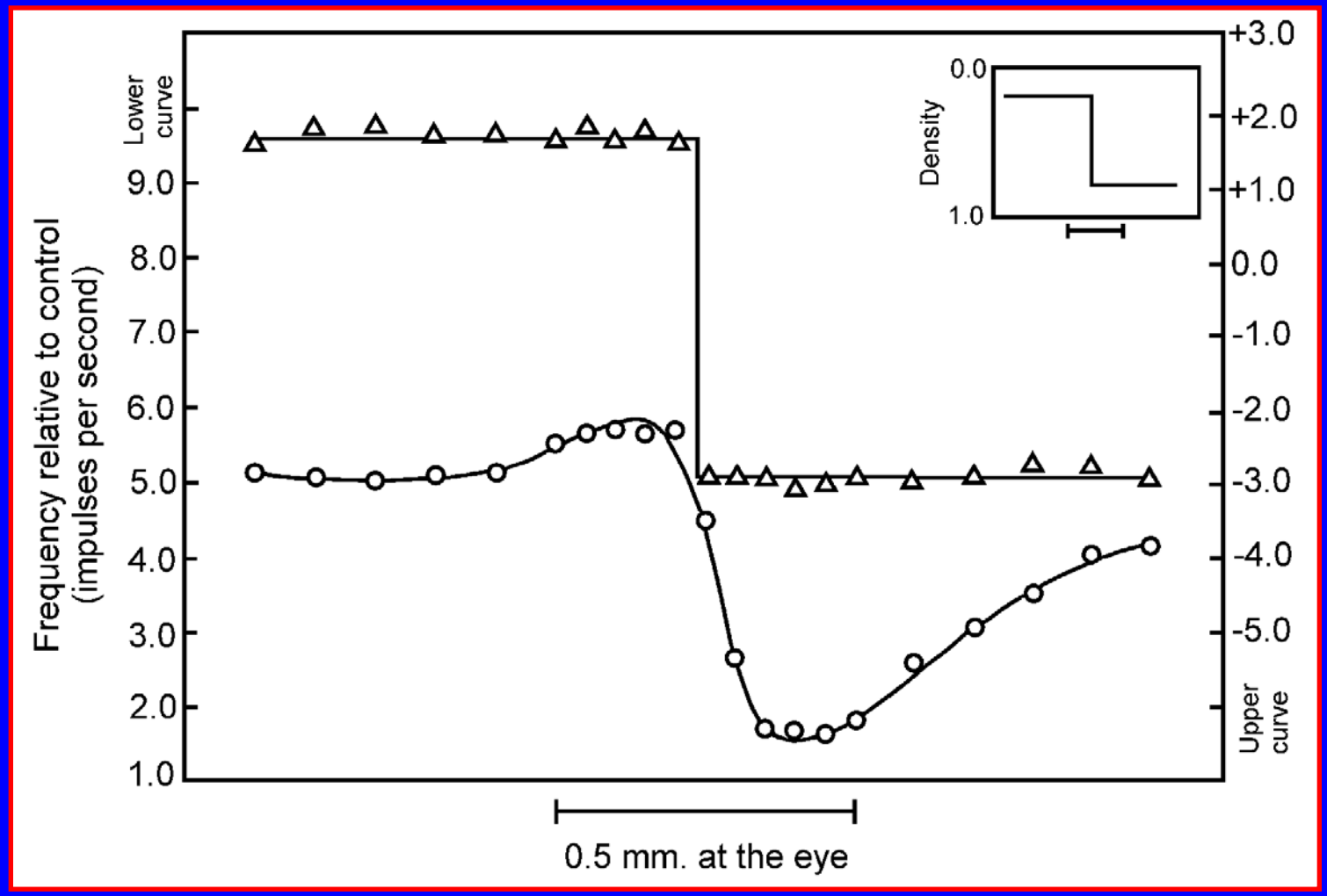
ADAPTATION AND INHIBITION IN THE NERVOUS SYSTEM

- Adaptation:
  - If someone turns on a fan, we notice the sound at first and then (normally) cease to notice it.
  - When the fan is turned off, we notice the quietness at first and then (normally) cease to notice it.
  - We do not normally notice our clothes, even though they are touching our skin all the time.
- Inhibition in the nervous system appears to be the mechanism for adaptation.
- Adaptation and inhibition are widespread in brains and nervous systems.
- Adaptation and inhibition as run-length coding: cut out repetition and mark transitions between one pattern and another.

ADAPTATION IN ONE OMMATIDIUM OF LIMULUS



EDGE DETECTION IN THE EYE OF LIMULUS



ADAPTATION, MICROSACCADES AND TREMOR IN THE MAMALIAN RETINA

- If we look very steadily at something, perhaps with artificial aids to steady one's eye, the image is likely to fade.
- But small movements of the eye ("microsaccades") or tremor in the eye will restore the image.
- As in the eye of Limulus, constant stimulation leads to adaptation, reversed by changes in stimulation.
- As before, adaptation may be seen as information compression.

INFORMATION COMPRESSION BETWEEN THE RETINA AND THE BRAIN

- The retina contains about 126 million photoreceptors.
- The optic nerve, connecting the retina to the brain, contains only about 1 million fibres.
- This suggests that there is likely to be a large reduction in redundant information between the retina and the brain.

BINOCULAR VISION

- Barlow (1969):

  “In an animal in which the visual fields of the two eyes overlap extensively, as in the cat, monkey, and man, one obvious type of redundancy in the messages reaching the brain is the very nearly exact reduplication of one eye’s message by the other eye.”

A RANDOM-DOT STEREOGRAM (JULESZ, 1971)

THE STRUCTURE OF THE LEFT AND RIGHT IMAGES IN THE RANDOM-DOT STEREOGRAM

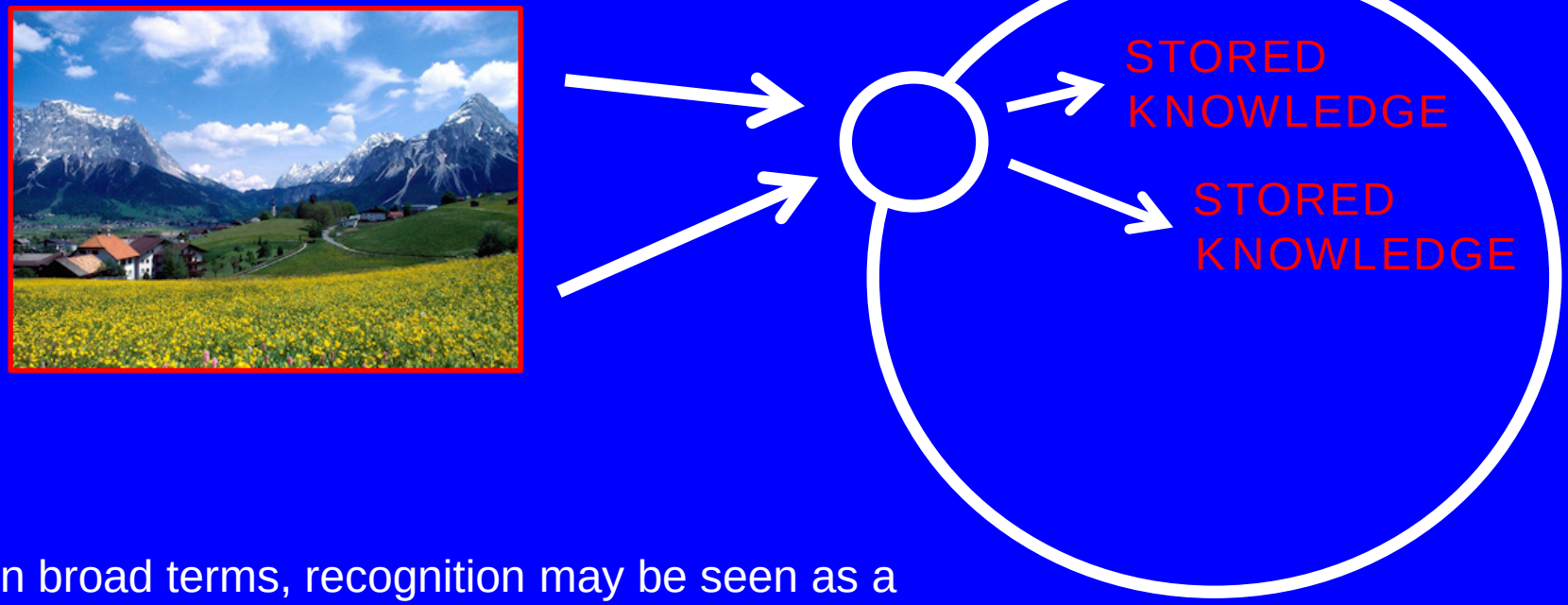
1	0	1	0	1	0	0	1	0	1
1	0	0	1	0	1	0	1	0	0
0	0	1	1	0	1	1	0	1	0
0	1	0	Y	*A*	*A*	*B*	*B*	0	0
1	1	1	X	*B*	*A*	*B*	*A*	0	1
0	0	1	X	*A*	*A*	*B*	*A*	1	0
1	1	1	Y	*B*	*B*	*A*	*B*	0	1
1	0	0	1	1	0	1	1	0	1
1	1	0	0	1	1	0	1	1	1
0	1	0	0	0	1	1	1	1	0

1	0	1	0	1	0	0	1	0	1
1	0	0	1	0	1	0	1	0	0
0	0	1	1	0	1	1	0	1	0
0	1	0	*A*	*A*	*B*	*B*	X	0	0
1	1	1	*B*	*A*	*B*	*A*	Y	0	1
0	0	1	*A*	*A*	*B*	*A*	Y	1	0
1	1	1	*B*	*B*	*A*	*B*	X	0	1
1	0	0	1	1	0	1	1	0	1
1	1	0	0	1	1	0	1	1	1
0	1	0	0	0	1	1	1	1	0

MERGING MULTIPLE VIEWS

If we close our eyes for a moment and open them again, we merge the ‘before’ and ‘after’ views.

INFORMATION COMPRESSION IN RECOGNITION



In broad terms, recognition may be seen as a process of matching incoming information with stored knowledge, merging or ‘unifying’ patterns that are the same, and thus compressing information.

OBJECTS AND CLASSES IN PERCEPTION AND COGNITION

- Objects: we collapse the ‘cinema frames’ of a moving object into a single object and single background.
- Classes: Attributes which are shared by all members of a class need be recorded only once and not repeated for every member.

NATURAL LANGUAGES

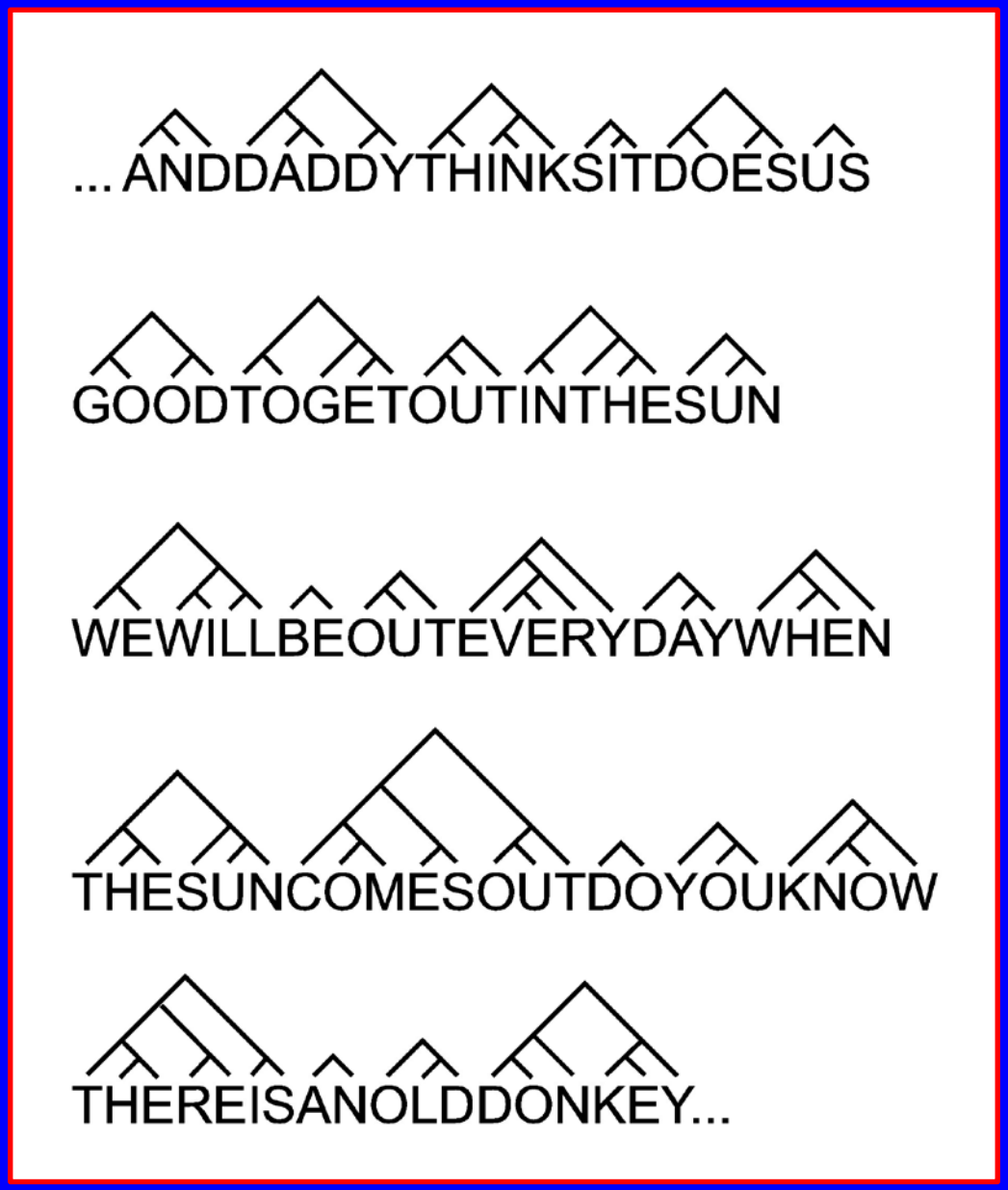
- Every noun, verb, adjective or adverb, may be seen as a ‘code’ for a relatively complex ‘chunk’ of information (the word’s meaning).
- Imagine saying “a horizontal platform with four, sometimes three, vertical supports, normally about three feet high, normally used for …” every time we wanted to refer to a “table”—like the slow language of the Ents in Tolkien’s *The Lord of the Rings*.

SCIENCE AS INFORMATION COMPRESSION

- John Barrow:

  "Science is, at root, just the search for compression in the world. ... the world is surprisingly compressible and the success of mathematics in describing its workings is a manifestation of that compressibility."

- The SP theory: mathematics is largely a set of techniques for compressing information (more later).



A parsing of text with no spaces or punctuation — developed by program MK10 without any prior knowledge of words.

The key is compression of information via the matching and unification of patterns.

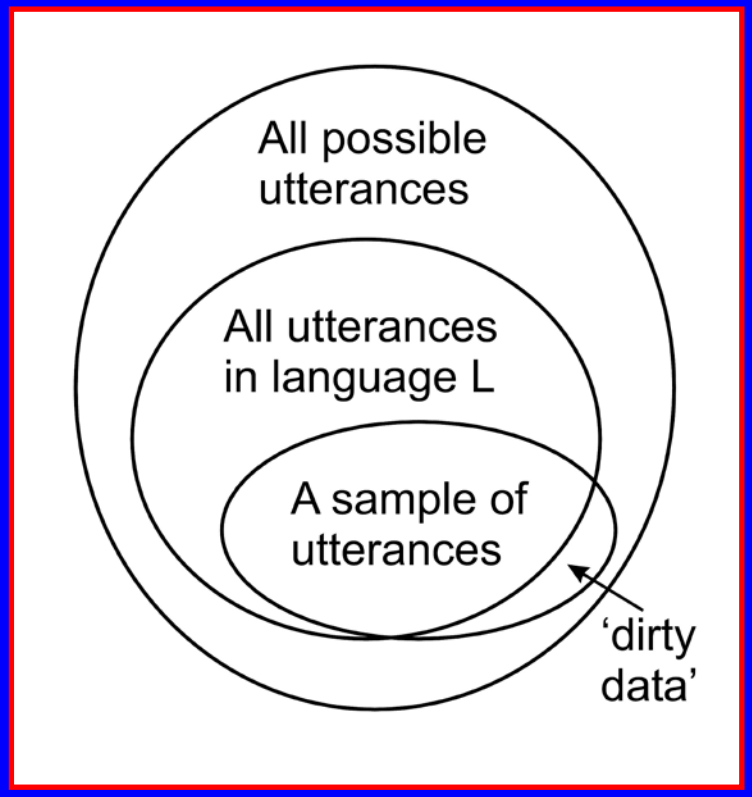
GRAMMATICAL INFERENCE: PROBLEMS OF GENERATION AND 'DIRTY DATA'

Problems:

- How to generalise without over-generalising?
- How to learn despite errors in what children hear ('dirty data')?

Gold (1967): learning needs correction by a 'teacher' or other aids.

No, this is only with a narrow definition of learning. Children can learn without these things.



Information compression provides a solution:

Minimise (G + E), where

- G is the size of the grammar, and
- E is the size of the sample when it is encoded in terms of the grammar.

PERCEPTUAL CONSTANCIES

- Size constancy: We judge the size of an object to be constant despite wide variations in the size of its image on the retina.
- Brightness constancy: We judge the brightness of an object to be constant despite wide variations in the intensity of its illumination.
- Colour constancy: We judge the colour of an object to be constant despite wide variations in the colours of its illumination.
- Without these constancies, memories for objects and events would be much more complicated than is our ordinary experience.

MATCHING AND UNIFICATION OF PATTERNS IN COMPUTING

- The 'Post Canonical System', an equivalent of the Turing machine, is essentially a system for the matching and unification of patterns (MUP).
- Query-by-example, and other forms of information retrieval, are largely MUP.
- MUP is a prominent feature of Prolog and other versions of logic programming.
- Dereferencing of identifiers requires MUP.
- Access to and retrieval of information from computer memory requires MUP.
- Etc.

CHUNKING-WITH-CODES IN COMPUTING

- A named ‘function’, ‘procedure’ or ‘sub-routine’ may be referenced from two or more parts of a program.
- Named objects in object-oriented systems.
- Named records in databases.
- Named files.
- Named folders or directories.
- Etc.

SCHEMA-PLUS-CORRECTION IN COMPUTING

- A program or named procedure:
  - The body of the program or procedure = schema.
  - Parameters are empty slots or variables within the schema.
  - Values for those variables provide corrections to the schema.
  - Conditional statements apply those corrections within the schema.
- A class (in an object-oriented system) = schema. Objects derived from a class contain specific values or corrections to the schema.

RUN-LENGTH CODING IN COMPUTING

- Repeat … until …, while …, for …, eg

```
s = 0;
for (i = 1; i <= 100; i++) s += i;
```

- Recursion, eg

```
int factorial(int x)
{
   if (x == 1) return(1) ;
   return(x * factorial(x - 1)) ;
}
```

CONCEPTS OF MATHEMATICS

- Mathematical Platonism: Mathematical concepts are “numinous and transcendent entities, existing independently of both the phenomena they order and the human mind that perceives them.” (Hersh, 1997).
- SP view: Mathematical concepts are forms of information in brains or computers, like other concepts.

INFORMATION COMPRESSION IN MATHEMATICS

- John Barrow:

  "For some mysterious reason mathematics has proved itself a reliable guide to the world in which we live and of which we are a part. Mathematics works: as a result we have been tempted to equate understanding of the world with its mathematical encapsulization. ... Why is the world found to be so unerringly mathematical?"

- Suggested answer: because mathematics is largely a set of techniques for the compression of information. Likewise for logic.

A 'FUNCTION' IS A COMPRESSED TABLE

Newton's second law of motion:

$$s = gt^2 / 2$$

where

- **s** is distance
- **g** is acceleration
- **t** is time

This is a compressed representation of a very large table, part of which is shown on the right.

The function uses techniques for compression of information (next).

Distance (m)	Time (sec)
0.0	0
4.9	1
19.6	2
44.1	3
78.4	4
122.5	5
176.4	6
240.1	7
313.8	8
Etc	Etc

CHUNKING-WITH-CODES IN MATHEMATICS

- Names are widespread in mathematics: We use names for functions, sets, members of sets, numbers, variables, etc. In most cases, the name represents a relatively large chunk of information.
- Base 1 numbers: Each unit is a generalised name of an object, typically a real-world object such as a goat. Each object is itself a conceptual ‘chunk’.
- Numbers with a base greater than 1: Each digit is the name of a chunk. Eg, ‘253’ (base 10) = a chunk of a 200 units + a chunk of 50 units + a chunk of 3 units.

SCHEMA-PLUS-CORRECTION IN MATHEMATICS

- Any kind of structure containing variables may be seen as a schema, with the value of each variable as a correction to the schema.
- Just like a program or procedure in computing, a mathematical function may be seen as a schema, and the arguments provide corrections to the schema.

RUN-LENGTH CODING IN MATHEMATICS

- Multiplication (eg, 3 × 4) is a shorthand for repeated addition.
- Division (eg, 12 / 3) is a shorthand for repeated subtraction.
- The power notation (eg, 10^9) is a shorthand for repeated multiplication.
- A factorial (eg, 10!) is a shorthand for repeated multiplication and subtraction.
- The bounded summation notation ('∑') and the bounded power notation ('∏') are shorthands for repeated addition and repeated multiplication, respectively.
- In both '∑' and '∏', there is normally a change in the value of a variable on each iteration, so these notations may be seen as a combination of run-length coding and schema-plus-correction.

FURTHER INFORMATION

- Chapters 2 and 10 in *Unifying Computing and Cognition*, CognitionResearch.org.
- www.cognitionresearch.org.
- *Contact:* jgw@cognitionresearch.org.