# IST Seminar 4

# Technical Writing

Marco CUTURI

# Purpose of technical writing

- Transmit technical information accurately using written symbols.

- Different from

  - Popular fiction/non-fiction (entertainment)

  - Advertising (sell)

- Sacrifice **style, grace, and technique** for **clarity, precision, and organization**

# Principles of Technical Writing

- Use the active voice.
    - Passive: Dolphins were taught by researchers in Hawaii to learn new behavior.
    - Active: Researchers in Hawaii taught dolphins to learn new behavior.

# Principles of Technical Writing

- Use plain rather than elegant or complex language.

- Elegant: The corporation deemed it necessary to terminate Joseph Smith.

- Plain: Joseph Smith was fired.

# Principles of Technical Writing

- Delete words, sentences, and phrases that do not add to your meaning.

  - Wordy: It is most useful to keep in mind that the term diabetes mellitus refers to a whole spectrum of disorders.

  - Concise: Diabetes mellitus refers to a whole spectrum of disorders.

# Principles of Technical Writing

- Know your reader. How much do you have to explain?
    - Can you say “operating system” or do you need to define it?
    - Do not invent words.
- Break your writing into short sections.

# When will you need to use Technical Writing?

- a thesis
- a scientific paper
- a letter to editor / collaborator / reviewer
- a grant proposal
- a cover letter/research statement
- a patent submission
- a blog post, news article

# Writing a Scientific Paper

## Dynamic Programming Algorithm Optimization for Spoken Word Recognition

HIROAKI SAKOE AND SEIBI CHIBA

*Abstract*—This paper reports on an optimum dynamic programming (DP) based time-normalization algorithm for spoken word recognition. First, a general principle of time-normalization is given using time-warping function. Then, two time-normalized distance definitions, called symmetric and asymmetric forms, are derived from the principle. These two forms are compared with each other through theoretical discussions and experimental studies. The symmetric form algorithm superiority is established. A new technique, called slope constraint, is successfully introduced, in which the warping function slope is restricted so as to improve discrimination between words in different categories. The effective slope constraint characteristic is qualitatively analyzed, and the optimum slope constraint condition is determined through experiments. The optimized algorithm is then extensively subjected to experimental comparison with various DP-algorithms, previously applied to spoken word recognition by different research groups. The experiment shows that the present algorithm gives no more than about two-thirds errors, even compared to the best conventional algorithm.

### I. Introduction

It is well known that speaking rate variation causes nonlinear fluctuation in a speech pattern time axis. Elimination of this fluctuation, or time-normalization, has been one of the central problems in spoken word recognition research. At an early stage, some linear normalization techniques were examined, in which timing differences between speech patterns were eliminated by linear transformation of the time axis. Reports on these efforts indicated that any linear transformation is inherently insufficient for dealing with highly complicated fluctuation nonlinearity as well as that time-normalization significantly improves recognition accuracy.

DP-matching, discussed in this paper, is a pattern matching algorithm with a nonlinear time-normalization effect. In this algorithm, the time-axis fluctuation is approximately modeled with a nonlinear warping function of some carefully specified properties. Timing differences between two speech patterns are eliminated by warping the time axis of one so that the maximum coincidence is attained with the other. Then, the time-normalized distance is calculated as the minimized residual distance between them. This minimization process is very efficiently carried out by use of the dynamic programming (DP) technique. The basic idea of DP-matching has been reported in several publications [1]-[3], where it has been shown by preliminary experiment on Japanese digit words that a recognition accuracy as high as 99.8 percent has been achieved, indicating the DP-matching effectiveness.

This paper reports an optimum algorithm for DP-matching through theoretical discussions and experimental studies. Investigations were made, based on the assumption that speech patterns are time-sampled with a common and uniform sampling period, as in most general cases. One of the problems discussed in this paper involves the relative superiority of either a symmetric form of DP-matching or an asymmetric one. In the asymmetric form, time-normalization is achieved by transforming the time axis of a speech pattern onto that of the other. In the symmetric form, on the other hand, both time axes are transformed onto a temporarily defined common axis. Theoretical and experimental comparisons show that the symmetric form gives better recognition than the asymmetric one. Another problem discussed concerns slope constraint technique. Since too much of the warping function flexibility sometimes results in poor discrimination between words in different categories, a constraint is newly introduced on the warping function slope. Detailed slope constraint condition is optimized through experimental studies. As a further investigation, the optimized algorithm is experimentally compared with several varieties of the DP-algorithm, which have been applied to spoken word recognition by some research groups [3]-[6]. The optimized algorithm superiority is established, indicating the validity of this investigation.

Manuscript received February 17, 1977; revised September 7, 1977.
The authors are with Central Research Laboratories, Nippon Electric Company, Limited, Kawasaki, Japan.

### II. DP-Matching Principle

#### A. General Time-Normalized Distance Definition

Speech can be expressed by appropriate feature extraction as a sequence of feature vectors.

$$A = a_1, a_2, \cdots\cdots, a_i, \cdots\cdots, a_I$$
$$B = b_1, b_2, \cdots\cdots, b_j, \cdots\cdots, b_J. \quad (1)$$

Consider the problem of eliminating timing differences between these two speech patterns. In order to clarify the nature of time-axis fluctuation or timing differences, let us consider an $i$-$j$ plane, shown in Fig. 1, where patterns $A$ and $B$ are developed along the $i$-axis and $j$-axis, respectively. Where these speech patterns are of the same category, the timing differences between them can be depicted by a sequence of points $c = (i, j)$:

$$F = c(1), c(2), \cdots\cdots, c(k), \cdots\cdots, c(K), \quad (2)$$

where

$$c(k) = (i(k), j(k)).$$

This sequence can be considered to represent a function which approximately realizes a mapping from the time axis of pattern $A$ onto that of pattern $B$. Hereafter, it is called a warping function. When there is no timing difference between these

# Writing a Scientific Paper

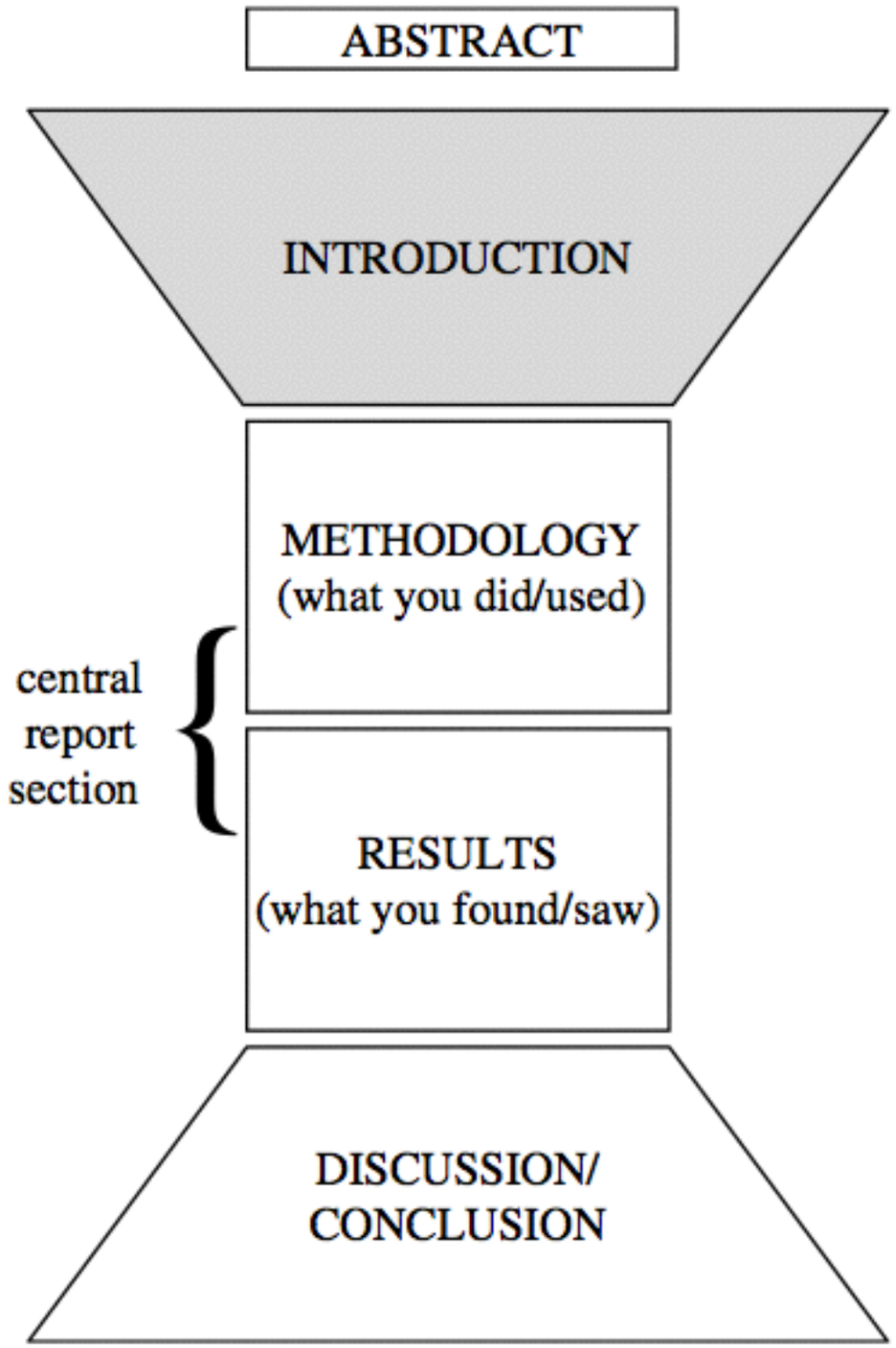



Fig. 1. The shape of a research article or thesis.

# Writing a Scientific Paper

- Writing a Scientific Paper is an Exercise in **communication**
- The goal is to **communicate** a complex idea, and **convince the reader that this new idea is worth the reader's time.**
- To **write** a paper, the **only thing you need** is become an expert in writing each of these sections:
- **abstract / introduction / methodology / experiments / discussion / conclusion.**
- **Unlike "writing" the goal of scientific writing is well defined.**

# The Abstract

# Elevator Talk

- **What is an elevator talk, and what does it have to do with writing a paper? A lot.**
- Imagine you are the president of the nonprofit *Light Is The Solution Foundation*.
- The board of directors is meeting at the New York Hilton, and you are waiting to ride the elevator from the 31st floor to the lobby.

The Abstract and the Elevator Talk: A Tale of Two Summaries Thomas M. Annesley

# Elevator Talk

- The doors open, and you find yourself standing with **Bill Gates**, head of the **philanthropic** (博愛) Gates Foundation

- The **Gates Foundation** meeting at the same hotel.

- Gates notices the logo on your shirt of a small child reading a book by the light of a lantern and asks you, *"What is that? What do you do?"*

The Abstract and the Elevator Talk: A Tale of Two Summaries Thomas M. Annesley

# Elevator Talk

- Indeed, **what do you do now?** You have 30 floors, or ~1 minute, to give him your message and explain what you do.
- This might be a **unique chance** to get his attention

The Abstract and the Elevator Talk: A Tale of Two Summaries Thomas M. Annesley

# Elevator Talk

- So you explain that **normal living activities** cease in many countries in the world **after the sun goes down**. Children **have no light to read textbooks**, mothers no light to cook, fathers no light to earn income.
- With this background, you then explain that the ***Light Is The Solution Foundation*** has addressed this problem by developing rechargeable LED lanterns that are low-cost, have a battery life of 30 hours, and put out light equivalent to three 60-W bulbs.
- You have given away **4500 lanterns in one country** and have results showing that more children now share books, study together, and graduate at a higher rate. **In fact, average incomes have risen by 20% for families who have received a lantern.**
- You have concluded that this unique program could be expanded to any country that has even the crudest electrical grid or generators for recharging the lanterns.

# Elevator Talk

- This is the **elevator talk**.
- Your **1-minute** opportunity to summarize what you do, how you do it, the results you produce, and the impact you make.
- A **well-developed elevator talk** entices the listener to want to learn more.
- In many professions, **entire careers are made and lost as a result of elevator talks**.

- **ABSTRACT = ELEVATOR TALK**

The Abstract and the Elevator Talk: A Tale of Two Summaries Thomas M. Annesley

# Abstract

**Table 1. Characteristics of a well-written abstract.**

| |
|---|
| Stands on its own without need to read the paper |
| States the hypothesis, question, or objective of the study |
| Completes the story by answering the hypothesis, question, or objective |
| Contains the same key words and terms as the title and the introduction |
| Follows the correct style and format |
| Follows the order of the main text (e.g., IMRAD) |
| Stays within the allowed word count |
| Does not contain information absent in the paper |
| Does not make conclusions unsupported by the data |
| Limits the use of abbreviations |
| Does not include references |
| Does not cite tables or figures |

IMRAD format (Introduction, Methods, Results, and Discussion)

The Abstract and the Elevator Talk: A Tale of Two Summaries Thomas M. Annesley

# Abstract


Google
support vector networks
Scholar
About 1,760,000 results (0.04 sec)
Articles
Case law
My library
Support-vector networks
C Cortes, V Vapnik - Machine learning, 1995 - Springer
Abstract The support-vector network is a new learning machine for two-group classification problems. The machine conceptually implements the following idea: input vectors are non-linearly mapped to a very high-dimension feature space. In this feature space a linear ...
Cited by 13476 Related articles All 40 versions Web of Science: 5016 Cite Saved

# Abstract

## SUPPORT-VECTOR NETWORKS

Corinna Cortes [1] and Vladimir Vapnik [2]
AT&T Labs-Research, USA

**Abstract.** The *support-vector network* is a new learning machine for two-group classification problems. The machine conceptually implements the following idea: input vectors are non-linearly mapped to a very high-dimension feature space. In this feature space a linear decision surface is constructed. Special properties of the decision surface ensures high generalization ability of the learning machine. The idea behind the support-vector network was previously implemented for the restricted case where the training data can be separated without errors. We here extend this result to non-separable training data.

High generalization ability of support-vector networks utilizing polynomial input transformations is demonstrated. We also compare the performance of the support-vector network to various classical learning algorithms that all took part in a benchmark study of Optical Character Recognition.

# Introduction

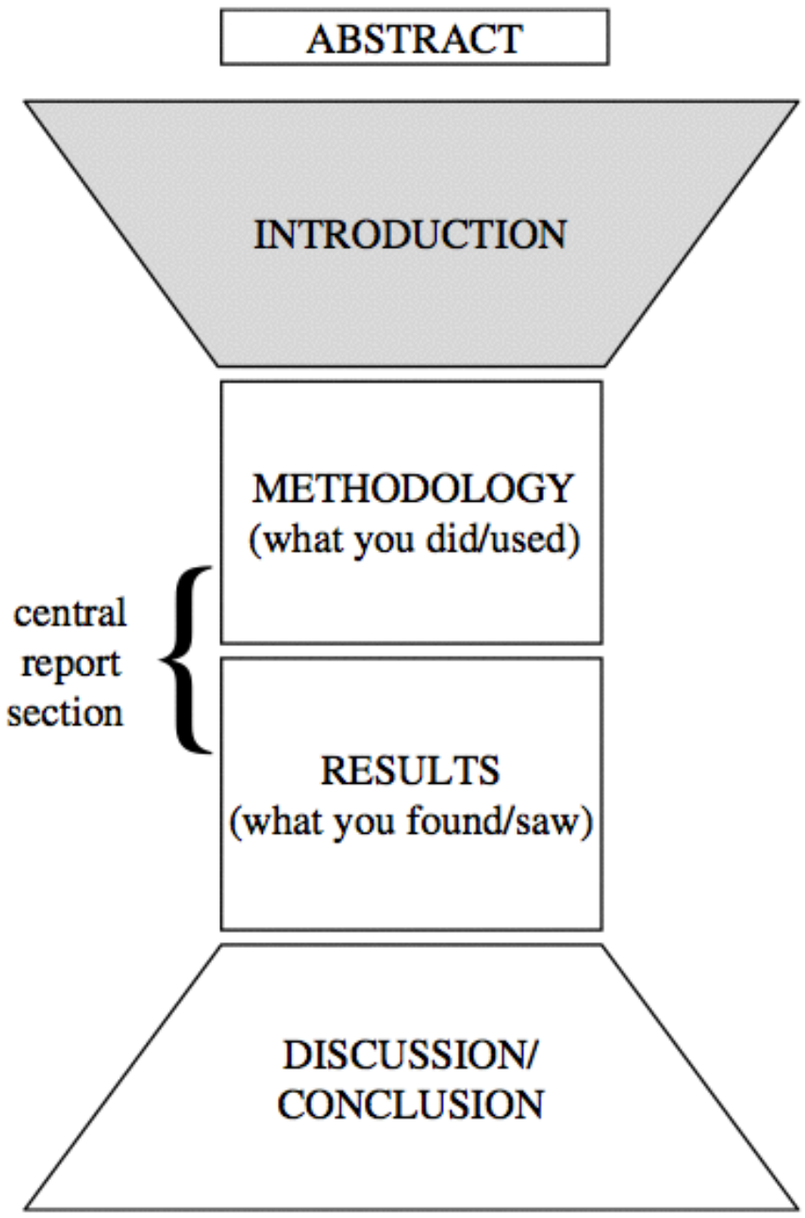



Fig. 1. The shape of a research article or thesis.

*Science research writing for non-native speakers of English* by Hilary Glasman-Deal.

# Write an Introduction

- Overall Structure
- Grammar and important language skills
- Build a Model

# Structure

- *How do writers usually start an introduction?*
- *What types of information should be in my introduction?*
- *How do writers usually end an introduction?*

# Structure

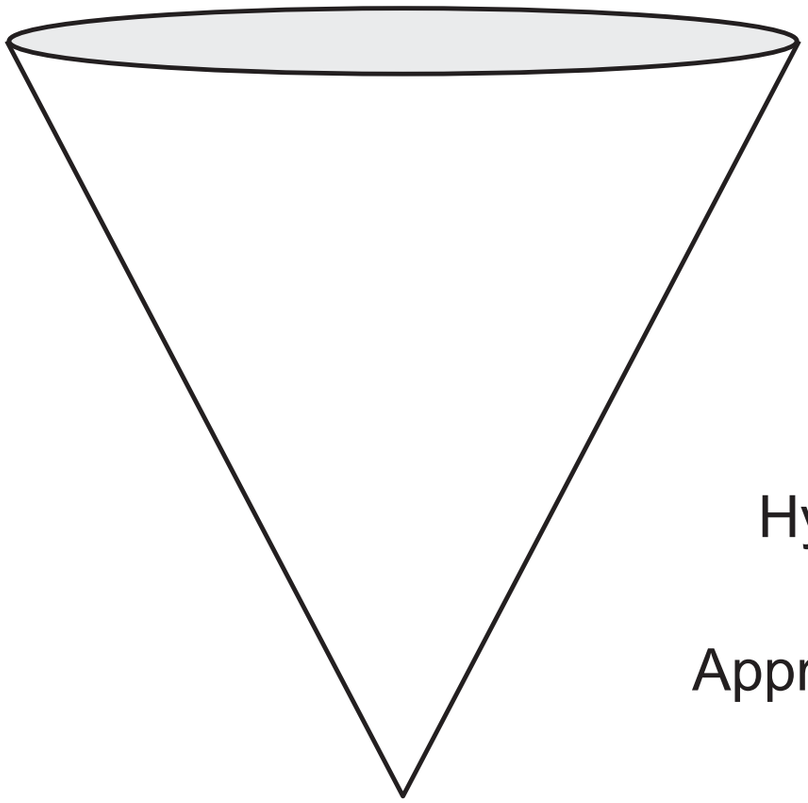
- Introduction & Conclusion are asymmetric (非対称)

- Start with general topics, converge to your contribution or point.

- Highlight the importance of the problem you discuss and explain as simply as possible your contribution

- Close by laying out the structure

# Structure



Background, known information

Knowledge gap, unknown information

Hypothesis, question, purpose statement

Approach, plan of attack, proposed solution

**Fig. 1.** The introduction: a cone or funnel.

**"It was a cold and rainy night":**
**Set the Scene with a Good Introduction** Thomas M. Annesley

# Introduction: Example

**The synthesis of flexible polymer blends from polylactide and rubber**

## Introduction

**1** *Polylactide (PLA) has received much attention in recent years due to its biodegradable properties, which offer important economic benefits.* **2** *PLA is a polymer obtained from corn and is produced by the polymerisation of lactide.* **3** *It has many possible uses in the biomedical field*[1] *and has also been investigated as a potential engineering material.*[2,3] **4** *However, it has been found to be too weak under impact to be used commercially.*[4]

**5** *One way to toughen polymers is to incorporate a layer of rubber particles*[5] *and there has been extensive research regarding the rubber modification of PLA.* **6** *For example, Penney et al. showed that PLA composites could be prepared using blending techniques*[6] *and more recently, Hillier established the toughness of such* composites.[7] **7** *However, although the effect of the rubber particles on the mechanical properties of copolymer systems was demonstrated over two years ago,*[8] *little attention has been paid to the selection of an appropriate rubber component.*

**8** *The present paper presents a set of criteria for selecting such a component.* **9** *On the basis of these criteria it then describes the preparation of a set of polymer blends using PLA and a hydrocarbon rubber (PI).* **10** *This combination of two mechanistically distinct polymerisations formed a novel copolymer in which the incorporation of PI significantly increased flexibility.*

# Introduction: Example

**Example 1**

Cardiovascular disease is a public health problem worldwide. Chronic inflammation has been linked to cardiovascular disease and sudden cardiac death *(1–3)*. Recent studies have demonstrated that a strong association exists between β-selectin, a recognized systemic marker of inflammation, and cardiovascular disease *(4–6)*, and that patients with peripheral vascular disease have increased serum concentrations of β-selectin that correlate with the degree of functional impairment *(7,8)*. Additionally, prospective studies have presented data regarding the prognostic value of β-selectin in predicting the severity of underlying cardiovascular disease and risk of mortality. The Vascular Inflammation Prediction (VIP) Study found a positive correlation between concentrations of β-selectin and the risk of developing cardiovascular disease *(9)*. The Canadian All-Cause Mortality Study revealed that individuals with serum β-selectin concentrations >90 μg/L are 4.5 times as likely to die within 5 years as those with concentrations ≤90 μg/L *(10)*.

Whereas these association and prospective studies indicate that β-selectin is a predictor of cardiovascular disease and risk of mortality, they provide little information about the underlying pathophysiology of vascular inflammation and the contributory role, if any, of β-selectin.

We therefore investigated in an animal model whether β-selectin is a cause or just a marker of vascular inflammation associated with cardiovascular disease.

Using a herpes simplex virus type 2 infection protocol to stimulate continuous β-selectin production in mice, we investigated the effects of β-selectin production on the development of atherosclerotic lesions, life span, and potential mechanisms of β-selectin–induced inflammation.

# Passive / Active

- *I describe an algorithm to cluster sequences*
- *We describe an algorithm to cluster sequences*
- *An algorithm to cluster sequences is described*
- *This paper describes an algorithm to cluster sequences*

# Passive / Active

- ~~*I describe an algorithm to cluster sequences*~~
- *We describe an algorithm to cluster sequences*
- ~~*An algorithm to cluster sequences is described*~~
- *This paper describes an algorithm to cluster sequences*

# Only Use *We* for you!

- ~~*We know that algorithms to cluster sequences are important*~~
- In that case, Passive is useful:
- *It is well known that clustering sequences is an important algorithmic problem*

# Using Paragraphs

- What is a paragraph? a unit of text. Usually a few sentences

- Why use paragraphs? make it easier for your reader to parse your text

- Having said that: how do readers usually look at articles?

# Quick Paper Reading

1. READ THE TITLE
   and try to predict the type of information you expect to see
2. LOOK AT THE NAME OF THE AUTHOR
   What you know about the writer will help you predict and evaluate the content.
3. CHECK THE DATE
   and use it to help you assess the content.
4. READ THE ABSTRACT
   to find out what the researchers did and/or what they found
5. LOOK QUICKLY AT THE FIRST PARAGRAPH
   without trying to understand all the words.
6. LOOK QUICKLY AT THE FIRST SENTENCE OF EACH PARAGRAPH
   without trying to understand all the words
7. LOOK QUICKLY AT EACH FIGURE/TABLE AND READ ITS TITLE
   to try and find out what type of visual data is included
8. READ THE LAST PARAGRAPH
   especially if it has a subtitle like ‘Summary’ or ‘Conclusion’

# The Crime Novel

- *Suppose someone gives you a crime novel*
- *Suppose you read **first** the end*
  - *You know who did the crime, why, how*
- *Start reading the novel from the start now*
- **Will you read the novel faster?**

# Structure

- *What types of information should be in my introduction?*

# Building A Model

**The synthesis of flexible polymer blends from polylactide and rubber**

***Introduction***

**In this sentence, the writer:**

**1** Polylactide (PLA) has received much attention in recent years due to its biodegradable properties, which offer important economic benefits. **2** PLA is a polymer obtained from corn and is produced by the polymerisation of lactide. **3** It has many possible uses in the biomedical field[1] and has also been investigated as a potential engineering material.[2,3] **4** However, it has been found to be too weak under impact to be used commercially.[4]

**5** One way to toughen polymers is to incorporate a layer of rubber particles[5] and there has been extensive research regarding the rubber modification of PLA. **6** For example, Penney *et al.* showed that PLA composites could be prepared using blending techniques[6] and more recently, Hillier established the toughness of such composites.[7] **7** However, although the effect of the rubber particles on the mechanical properties of copolymer systems was demonstrated over two years ago,[8] little attention has been paid to the selection of an appropriate rubber component.

**8** The present paper presents a set of criteria for selecting such a component.

1_____

2_____

3_____

4_____

5_____

6_____

7_____

8_____

# Building A Model

**In Sentence 1** '*Polylactide (PLA) has received much attention in recent years due to its biodegradable properties, which offer important economic benefits.*' **the writer establishes the importance of this research topic.**

# Building A Model

**In Sentence 2** '*PLA is a polymer obtained from corn and is produced by the polymerisation of lactide*.' **the writer provides general background information for the reader**.

# Building A Model

**In Sentence 3** '*PLA has many possible uses in the biomedical field*[1] *and has also been investigated as a potential engineering material* [2,3]' **the writer does the same as in Sentences 1 and 2, but in a more specific/detailed way, using research references to support both the background facts and the claim for significance.**

# Building A Model

**In Sentence 4** '*However, it has been found to be too weak under impact to be used commercially.*[4]' **the writer describes the general problem area or the current research focus of the field.**

# Building A Model

**In Sentence 5** ‘*One way to toughen polymers is to incorporate a layer of rubber particles.*[5]’ **the writer provides a transition between the general problem area and the literature review.**

# Building A Model

**In Sentence 6** *'For example, Penney et al. showed that PLA composites could be prepared using blending techniques*[6] *and more recently, Hillier*[7] *established the toughness of such composites.'* **the writer provides a brief overview of key research projects in this area.**

# Building A Model

**In Sentence** 7 '*However, although the effect of the rubber particles on the mechanical properties of copolymer systems was demonstrated over two years ago,*[8] *little attention has been paid to the selection of an appropriate rubber component.*' **the writer describes a gap in the research**.

# Building A Model

**In Sentence 8** '*The present paper presents a set of criteria for selecting such a component.*' **the writer describes the paper itself.**

# Building A Model

**In Sentence 9** '*On the basis of these criteria it then describes the preparation of a set of polymer blends using PLA and a hydrocarbon rubber(PI).*' **the writer gives details about the methodology reported in the paper**.

# Building A Model

**In Sentence 10** '*This combination of two mechanistically distinct polymerisations formed a novel copolymer in which the incorporation of PI significantly increased flexibility*.' **the writer announces the findings**.

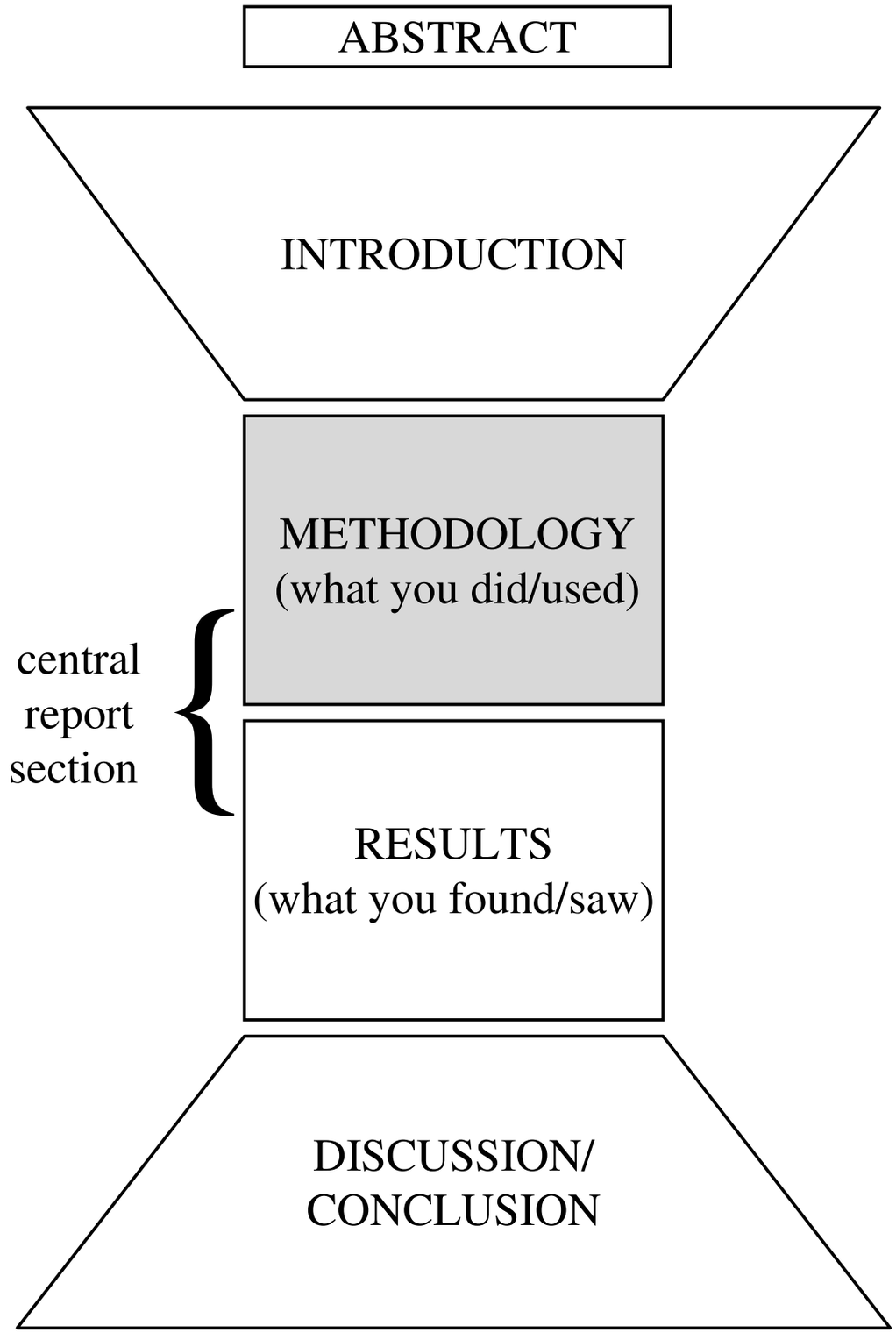
# Building A Model

| | |
|---|---|
| **In Sentence 1** | **the writer establishes the importance of this research topic.** |
| **In Sentence 2** | **the writer provides general background information.** |
| **In Sentence 3** | **the writer does the same as in Sentences 1 and 2, but in a more specific/detailed way.** |
| **In Sentence 4** | **the writer describes the general problem area or the current research focus of the field.** |
| **In Sentence 5** | **the writer provides a transition between the general problem area and the literature review.** |
| **In Sentence 6** | **the writer provides a brief overview of key research projects in this area.** |
| **In Sentence 7** | **the writer describes a gap in the research.** |
| **In Sentence 8** | **the writer describes the paper itself.** |
| **In Sentence 9** | **the writer gives details about the methodology reported in the paper.** |
| **In Sentence 10** | **the writer announces the findings.** |

# A Model For the Introduction

| | |
|---|---|
| 1 | ESTABLISH THE IMPORTANCE OF YOUR FIELD<br>PROVIDE BACKGROUND FACTS/INFORMATION (possibly from research)<br>DEFINE THE TERMINOLOGY IN THE TITLE/KEY WORDS<br>PRESENT THE PROBLEM AREA/CURRENT RESEARCH FOCUS |
| 2 | PREVIOUS AND/OR CURRENT RESEARCH AND CONTRIBUTIONS |
| 3 | LOCATE A GAP IN THE RESEARCH<br>DESCRIBE THE PROBLEM YOU WILL ADDRESS<br>PRESENT A PREDICTION TO BE TESTED |
| 4 | DESCRIBE THE PRESENT PAPER |

# Methodology: *Your Work*



# Your Work as a Student

- Up to now, you have mostly written about science to **prove that you have understood known concepts**.
- **Your professor knows how to fill the gaps,** understand what you write even if it's not clear.

# Work as a Scientist/Engineer

- Writing a scientific paper is different: **you will be the professor, the teacher**.

# Work as a Scientist/Engineer

- Writing a scientific paper is different: **you will be the professor, the teacher**.
- Becomes crucial that you can provide **all necessary** information

# who what when where how & why

| Who |
| --- |
| Who maintained the records? Who reviewed the data? Who collected the specimens? Who enrolled the study participants? Who supplied the reagents? Who made the primary diagnosis? Who did the statistical analyses? Who reviewed the protocol for ethics approval? Who provided the funding? |
| **What** |
| What reagents, methods, and instruments were used? What type of study was it? What were the inclusion and exclusion criteria for enrolling study participants? What protocol was followed? What treatments were given? What endpoints were measured? What data transformation was performed? What statistical software package was used? What was the cutoff for statistical significance? What control studies were performed? What validation experiments were performed? |

# who what when where how & why

## When

When were specimens collected? When were the analyses performed? When was the study initiated? When was the study terminated? When were the diagnoses made?

## Where

Where were the records kept? Where were the specimens analyzed? Where were the study participants enrolled? Where was the study performed?

# who what when where how & why

## How

How were samples collected, processed, and stored? How many replicates were performed? How was the data reported? How were the study participants selected? How were patients recruited? How was the sample size determined? How were study participants assigned to groups? How was response measured? How were endpoints measured? How were control and disease groups defined?

## Why

Why was a species chosen (mice vs rats)? Why was a selected analytical method chosen? Why was a selected experiment performed? Why were experiments done in a certain order?