Specialist English: Assignment 8 (solutions)

Rebecca J. Stones
rebecca.stones82@nbjl.nankai.edu.cn

December 19, 2017

Here's my solutions; your solutions needn't be identical.

Problem 1

1. *Rephrase the following sentence to remove "As shown in Figure 1(a) ...", which I recommend to avoid as it's difficult for the reader to parse.*

> As shown in Figure 1(a), the baseline CMP system consists of multiple cores, where the L1 caches are private to each core and the lower level caches are shared by the cores.
>
> — Zhao, Xu, and Xie, *Bandwidth-Aware Reconfigurable Cache Design with Hybrid Memory Technologies*, ICCAD, 2011.

Looking at the paper, Figure 1 was introduced prior to this sentence, thus the authors only want to refer back to it (so it's not critically important).

One possibile way to revise this snippet is by not saying anything. I consider this the best option: the figure gives less information than the sentence, so looking at the figure does not help understand the sentence, and is more distracting than helpful.

> The baseline CMP system consists of multiple cores, where the L1 caches are private to each core and the lower level caches are shared by the cores.

However, if we wish to include it, we should keep it simple and separate, to ensure that it does not sound like Figure 1(a) determines the design.

> The baseline CMP system consists of multiple cores, where the L1 caches are private to each core and the lower level caches are shared by the cores; see Figure 1(a).

2. *Rewrite the following snippet to eliminate "evidences" and "besides".*

> A comparative summary should contain as many comparative evidences as possible. Besides, it should convey important information in the original document.
>
> — Huang, Wan, and Xiao, *Comparative News Summarization Using Linear Programming*, HLT, 2011.

The problem with this sentence is that "evidence" is uncountable, and thus "as many comparative evidences" is a grammar error. The word "besides" has several meanings, and in this case it basically means "and".

Keeping things simple, we fix these problems as follows.

A comparative summary should contain as **much** comparative **evidence** as possible, **and** it should **also** convey **the** important information in the original document.

(I suspect "convey" is the wrong choice of word; it should instead be "contain".)

3. *While the following sentence is grammatically correct, "apart from" has multiple commonly used meanings, which slows down the reader who has to identify which meaning the authors intend. Replace "apart from" with a more direct synonym.*

> In this paper, apart from horizontal handoff and vertical handoff, terminal handoff is also considered, which is performed between two different terminals.
>
> — Ma et al., *Research on Load Balancing Mechanism ...*, ICHIT, 2009.

We can change this to the following.

In this paper, **in addition to** horizontal handoff and vertical handoff, **we also consider** terminal handoff, which is performed between two different terminals.

In addition to removing "apart from" which has multiple meanings, this changes the passive "terminal handoff is also considered" to the active "we also consider terminal handoff". This change also gives better visual separation of "horizontal handoff" and "vertical handoff" from the newly considered "terminal handoff".

Problem 2 *Rewrite the above pseudocode algorithm to optimize its presentation. Explain at least four ways you've made the pseudocode more readable, more consistent, or have otherwise improved the pseudocode.*

We modify the pseudocode as follows:

Algorithm 1 EliminatePoints

Input: Stream $(p_1, \ldots, p_n)$; desired number of skyline points m.
Output: Set S of skyline points.

1: **for** $j \in \{1, \ldots, 24m\}$ **do** ▷ Pass 1
2: Define q_j as a point chosen uniformly at random from the stream.
3: **end for**
4: **for** $i \in \{1, \ldots, n\}$ **do** ▷ Pass 2
5: If point p_i dominates any point q_j, then set $q_j \leftarrow p_i$.
6: **end for**
7: Set $S \leftarrow \{q_1, \ldots, q_{24m}\}$.
8: Delete from the stream all points in S and all points dominated by any point in S. ▷ Pass 3
9: **return** S.

How has the pseudocode been improved:

- The parameter m shouldn't be part of the caption.
- We replace n' with n. (There is some argument for using n', since n is meant to be reserved for the "number of points", but I consider the benefit of using simpler notation outweighs this.)
- We refer to $(p_1, \ldots, p_n)$ as the stream; this has two added benefits: (a) it's self-explanatory that n is the number of points in the stream, and therefore this no longer needs stating, and (b) in mathematical notation, the brackets "(" and ")" imply it's ordered.
- The algorithm does not necessarily return m points, so I describe m as the "desired number of skyline points".

- We change the description of S' (now denoted S) to "set of skyline points", which is more accurate.
- The variable x is only used twice; so we replace it by its definition, i.e., $24m$.
- We write "Pass 1", "Pass 2", and "Pass 3" as comments, as they're not part of the code.
- We typeset the for loop in Pass 1 as a for loop, using `\For` and `\EndFor`; this eliminates the multi-line sentence.
- We change the for loop from "$j = 1, 2, ..., x$" which is mathematically incorrect and uses the wrong ellipsis to "$j \in \{1, 2, \ldots, x\}$".
- We replace "picked" in "picked uniformly at random" with "chosen", since "picked" sounds informal.
- We change p' to q, to get rid of the distracting notation.
- We rephrase the sentence "let p'_j be ..." to be more succinct. Moreover, "let" is not an appropriate word for assignment (it should be "set").
- The notation S ("set of such points") is unused and eliminated. So we also replace S' with S to get rid of the apostrophe.
- The notation "$i = 1..n'$" is incorrect, and changed to "$i \in \{1, \ldots, n'\}$". Also note that "$i = 1..n'$" (which has no increment) is inconsistent with "$j = 1, 2, ..., x$" (which has an increment).
- We rephrase "For any p'_j, if p_i dominates p'_j ..." to be more succinct.
- We change ":=" for assignment (which is used inconsistently) to "$\leftarrow$", which is less ambiguous. This is further emphasized by adding the word "set".
- The mathematics "$S' = \{p'_1, p'_2, ...p'_x\}$" is missing a comma, and the ellipsis is typeset incorrectly.
- We change "Delete from stream" to "Delete from the stream". (This is a kind of abbreviation where we omit "the" for succinctness, even though it's grammatically incorrect.)
- We use consistent punctuation; specifically, full stops to end lines. (It's okay not to, but we shouldn't end some lines with full stops, and others without full stops.)

Some students changed "For any p'_j, ..." into a for loop, with the aim of being more consistent. Bearing in mind that pseudocode is intended to be read by a human, I've chosen a different way of writing this.

If we're concerned about the order in which the for loops occur, we can write e.g. "for i from 1 to n'" instead of "for $i \in \{1, \ldots, n'\}$".

We should use semi-colons ";" to separate inputs. Using a comma to separate independent clauses is called a *comma splice*[1], and they're generally considered to be a mistake.

Problem 3 *The task here is to redraw this figure to make it look polished and worthy of publication. The purpose is not to make it look identical, but to fix the problems and make it visually more appealing.*

Figure 1 redraws the plot using `tikz` in LaTeX. Two important changes are: (a) getting rid of the legend, which is poorly placed in the top-left, and (b) replacing actual tweets with icons representing tweets. This reduces the amount of text in the figure. We also (a) change "promoting" to "promotion", (b) change "filter" to "filter out", (c) explain what's being classified, (d) minimize the number of user icons, and (e) get rid of the "jumps" where lines cross (including these suggests it's important, but it's not).

A lot of students simply redrew the figure with some minor polishing, but this misses the point of the exercise: students were meant to develop technical drawing skills, and to learn what to write (and what not to write) in a visualization, which is affected by your ability to draw. This, in turn, affects what's included in the caption and the main text.

[1] `https://en.wikipedia.org/wiki/Comma_splice`

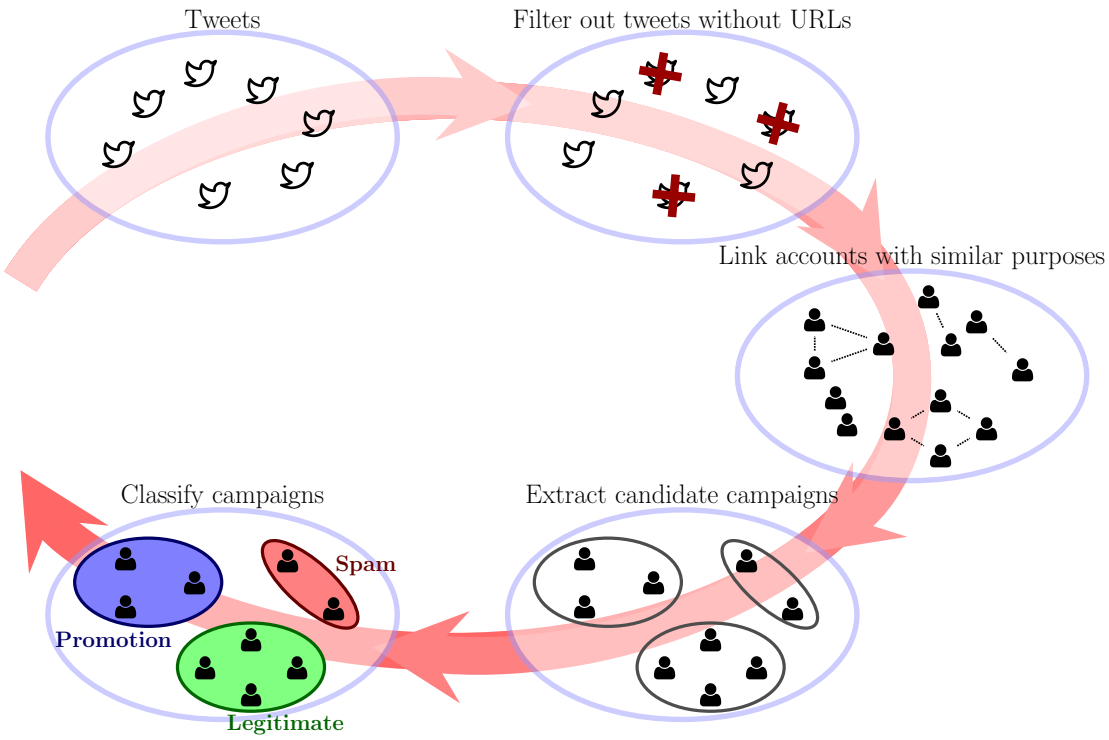

Figure 1: The work flow of the proposed detection framework.

Problem 4 *Explain how to edit this figure and/or caption to make it self contained, and to actually show an example of destination prediction.*

The figure is reasonably well-drawn, so changes should be minimal. I would simply add a small car icon along the road from l_1 to l_4, and explain in the caption that the car's predicted destination is the airport based on historical trajectories.

It might also be worthwhile adding "Historical trajectories" above the legend (i.e., above where T_1 through T_5 is described in the figure).

For those who would like my LaTeX code for Figure 1. The approximate process was:

1. Draw what I want on a piece of paper.
2. Typeset in it `tikz` one piece at a time.
3. I use `\documentclass[margin=1mm,crop]{standalone}` to make a separate .pdf file, and I use `\includegraphics[width=\textwidth]{Ass8Q3.pdf}` to resize the image appropriately.
4. Coordinates like `(2*360/6:10)` means (a) angle $2 \times 360/6 = 120$ degrees and (b) distance 10.
5. We can name coordinates such as with `\coordinate (c1) at (2*360/6:10);`.
6. We can add coordinates such as with `($(c1)+(t1)$)`.
7. I use simple `tikz` functions like `arc`, `circle`, and so on.
8. I downloaded some (free) icons from Wikimedia Commons.
9. I do some manual tweaking to make it look nice.

My tikz code is the following (it might not be optimized).

```
\documentclass[margin=1mm,crop]{standalone}

\usepackage{graphicx}

\usepackage{tikz}
\usetikzlibrary{calc}
\usetikzlibrary{arrows}
\usetikzlibrary{shapes.misc}


\begin{document}

\begin{tikzpicture}[xscale=1.5,yscale=0.9]

% red arrows around the outside
\draw[-stealth,line width=35,color=red!60!white] (160:10) arc (160:-160:10);
\draw[-stealth,line width=35,color=red!50!white] (160:10) arc (160:-100:10);
\draw[-stealth,line width=35,color=red!40!white] (160:10) arc (160:-40:10);
\draw[-stealth,line width=35,color=red!30!white] (160:10) arc (160:20:10);
\draw[-stealth,line width=35,color=red!20!white] (160:10) arc (160:80:10);

% the centers for the blobs
\coordinate (c1) at (2*360/6:10);
\coordinate (c2) at (1*360/6:10);
\coordinate (c3) at (0*360/6:10);
\coordinate (c4) at (5*360/6:10);
\coordinate (c5) at (4*360/6:10);

% text above blobs
\node at ($(c1)+(0,4.3)$) (tweets) {\Huge Tweets};
\node at ($(c2)+(0,4.3)$) (tweets) {\Huge Filter out tweets without URLs};
\node at ($(c3)+(0,4.3)$) (tweets) {\Huge Link accounts with similar purposes};
\node at ($(c4)+(0,4.3)$) (tweets) {\Huge Extract candidate campaigns};
\node at ($(c5)+(0,4.3)$) (tweets) {\Huge Classify campaigns};

% drawing the blobs
\fill[white,opacity=0.5,draw=blue!40,line width=5] (c1) circle (3.8);
```

```
\fill[white,opacity=0.5,draw=blue!40,line width=5] (c2) circle (3.8);
\fill[white,opacity=0.5,draw=blue!40,line width=5] (c3) circle (3.8);
\fill[white,opacity=0.5,draw=blue!40,line width=5] (c4) circle (3.8);
\fill[white,opacity=0.5,draw=blue!40,line width=5] (c5) circle (3.8);


% random tweet coordinates
\coordinate (t1) at ($(0*360/7:2)+(0.5*rand,0.5*rand)$);
\coordinate (t2) at ($(1*360/7:2)+(0.5*rand,0.5*rand)$);
\coordinate (t3) at ($(2*360/7:2)+(0.5*rand,0.5*rand)$);
\coordinate (t4) at ($(3*360/7:2)+(0.5*rand,0.5*rand)$);
\coordinate (t5) at ($(4*360/7:2)+(0.5*rand,0.5*rand)$);
\coordinate (t6) at ($(5*360/7:2)+(0.5*rand,0.5*rand)$);
\coordinate (t7) at ($(6*360/7:2)+(0.5*rand,0.5*rand)$);

% draw random tweets in first blob
\node at ($(c1)+(t1)$) (foo) {\includegraphics[width=3em]{Twitter.pdf}};
\node at ($(c1)+(t2)$) (foo) {\includegraphics[width=3em]{Twitter.pdf}};
\node at ($(c1)+(t3)$) (foo) {\includegraphics[width=3em]{Twitter.pdf}};
\node at ($(c1)+(t4)$) (foo) {\includegraphics[width=3em]{Twitter.pdf}};
\node at ($(c1)+(t5)$) (foo) {\includegraphics[width=3em]{Twitter.pdf}};
\node at ($(c1)+(t6)$) (foo) {\includegraphics[width=3em]{Twitter.pdf}};
\node at ($(c1)+(t7)$) (foo) {\includegraphics[width=3em]{Twitter.pdf}};


% draw random tweets in second blob
\node at ($(c2)+(t1)$) (foo) {\includegraphics[width=3em]{Twitter.pdf}};
\node at ($(c2)+(t2)$) (foo) {\includegraphics[width=3em]{Twitter.pdf}};
\node at ($(c2)+(t3)$) (foo) {\includegraphics[width=3em]{Twitter.pdf}};
\node at ($(c2)+(t4)$) (foo) {\includegraphics[width=3em]{Twitter.pdf}};
\node at ($(c2)+(t5)$) (foo) {\includegraphics[width=3em]{Twitter.pdf}};
\node at ($(c2)+(t6)$) (foo) {\includegraphics[width=3em]{Twitter.pdf}};
\node at ($(c2)+(t7)$) (foo) {\includegraphics[width=3em]{Twitter.pdf}};

% eliminate random tweets in second blob
\draw ($(c2)+(t1)$) node[cross out,rotate=30,draw=red!60!black,minimum size=20,line width=8] {};
\draw ($(c2)+(t3)$) node[cross out,rotate=35,draw=red!60!black,minimum size=20,line width=8] {};
\draw ($(c2)+(t6)$) node[cross out,rotate=40,draw=red!60!black,minimum size=20,line width=8] {};


% user group coordinates
\coordinate (g1) at (30:2.2);
\coordinate (g2) at (2*360/5:2);
\coordinate (g3) at (4*360/5:2);

\coordinate (g4) at (1*360/5:2);
\coordinate (g5) at (3*360/5:2);

% user coordinates
\coordinate (v1) at ($(g1)+(120:1)$);
\coordinate (v2) at ($(g1)+(-60:1)$);

\coordinate (v3) at ($(g2)+(0*360/3:1)$);
\coordinate (v4) at ($(g2)+(1*360/3:1)$);
\coordinate (v5) at ($(g2)+(2*360/3:1)$);

\coordinate (v6) at ($(g3)+(0*360/4:1)$);
\coordinate (v7) at ($(g3)+(1*360/4:1)$);
```

```
\coordinate (v8) at ($(g3)+(2*360/4:1)$);
\coordinate (v9) at ($(g3)+(3*360/4:1)$);

\coordinate (v10) at ($(g4)+(0,0.7)+(0.5*rand,0.5*rand)$);
\coordinate (v11) at ($(g4)+(0,-0.7)+(0.5*rand,0.5*rand)$);

\coordinate (v12) at ($(g5)+(0,0.5)+(0.5*rand,0.5*rand)$);
\coordinate (v13) at ($(g5)+(0,-0.5)+(0.5*rand,0.5*rand)$);

% draw random users in third blob
\node at ($(c3)+(v1)$) (u1) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c3)+(v2)$) (u2) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c3)+(v3)$) (u3) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c3)+(v4)$) (u4) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c3)+(v5)$) (u5) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c3)+(v6)$) (u6) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c3)+(v7)$) (u7) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c3)+(v8)$) (u8) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c3)+(v9)$) (u9) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c3)+(v10)$) (u10) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c3)+(v11)$) (u11) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c3)+(v12)$) (u12) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c3)+(v13)$) (u13) {\includegraphics[width=2.4em]{User.pdf}};

% draw lines in third blob
\draw[ultra thick,densely dotted] (u1) -- (u2);
\draw[ultra thick,densely dotted] (u3) -- (u5) -- (u4) -- (u3);
\draw[ultra thick,densely dotted] (u6) -- (u7) -- (u8) -- (u9) -- (u6);
\draw[ultra thick,densely dotted] (u10) -- (u11);


% draw ellipses in fourth blob
\draw[line width=3,black!70,rotate=30] ($(c4)+(g1)$) ellipse (0.6 and 1.8);
\draw[line width=3,black!70] ($(c4)+(g2)$) ellipse (1.6 and 1.6);
\draw[line width=3,black!70] ($(c4)+(g3)$) ellipse (1.6 and 1.6);

% draw clustered users in fourth blob
\node at ($(c4)+(v1)$) (u1) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c4)+(v2)$) (u2) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c4)+(v3)$) (u3) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c4)+(v4)$) (u4) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c4)+(v5)$) (u5) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c4)+(v6)$) (u6) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c4)+(v7)$) (u7) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c4)+(v8)$) (u8) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c4)+(v9)$) (u9) {\includegraphics[width=2.4em]{User.pdf}};


% draw ellipses in fifth blob
\draw[line width=3,red!40!black,fill=red!50,rotate=30] ($(c5)+(g1)$) ellipse (0.6 and 1.8);
\draw[line width=3,blue!40!black,fill=blue!50] ($(c5)+(g2)$) ellipse (1.6 and 1.6);
\draw[line width=3,green!40!black,fill=green!50] ($(c5)+(g3)$) ellipse (1.6 and 1.6);

% draw clustered users in fifth blob
\node at ($(c5)+(v1)$) (u1) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c5)+(v2)$) (u2) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c5)+(v3)$) (u3) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c5)+(v4)$) (u4) {\includegraphics[width=2.4em]{User.pdf}};
```

```
\node at ($(c5)+(v5)$) (u5) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c5)+(v6)$) (u6) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c5)+(v7)$) (u7) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c5)+(v8)$) (u8) {\includegraphics[width=2.4em]{User.pdf}};
\node at ($(c5)+(v9)$) (u9) {\includegraphics[width=2.4em]{User.pdf}};

% add labels to classifications
\node at ($(c5)+(g1)+(30:1.4)$) (foo) {\huge\bf \color{red!40!black} Spam};
\node at ($(c5)+(g2)+(-110:2.2)$) (foo) {\huge\bf \color{blue!40!black} Promotion};
\node at ($(c5)+(g3)+(-70:2.2)$) (foo) {\huge\bf \color{green!40!black} Legitimate};

\end{tikzpicture}

\end{document}
```