Writing in Computer Science: Style

Javier Larrosa

UPC Barcelona Tech

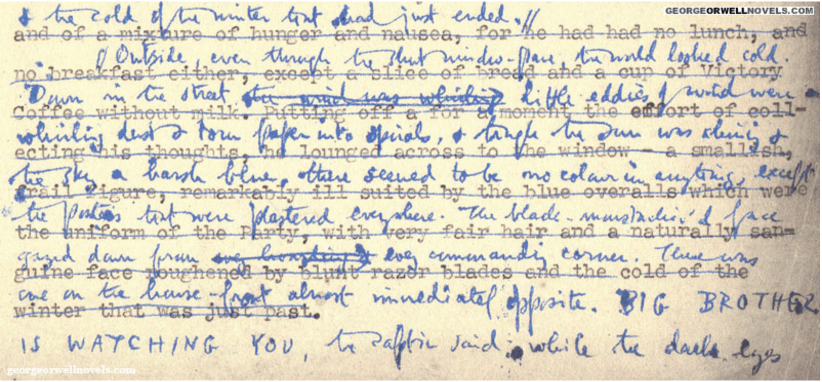
.

Do not mind revisions...

Revision

Most of the writing time is spent in revisions

1984's draft by George Orwell


GEORGEORWELLNOVELS.COM

& the cold of the winter that had just ended. //
~~and of a mixture of hunger and nausea, for he had had no lunch, and~~
// Outside, even through the shut window-pane, the world looked cold.
~~no breakfast either, except a slice of bread and a cup of Victory~~
Down in the street ~~the wind was whirling~~ little eddies of wind were
~~Coffee without milk. Putting off for a moment the effort of coll-~~
whirling dust & torn paper into spirals, & though the sun was shining &
~~ecting his thoughts, he lounged across to the window — a smallish,~~
the sky a harsh blue, there seemed to be no colour in anything, except
~~frail figure, remarkably ill suited by the blue overalls which were~~
the posters that were plastered everywhere. The black-moustachio'd face
~~the uniform of the Party, with very fair hair and a naturally san-~~
gazed down from ~~every~~ every commanding corner. There was
~~guine face roughened by blunt razor blades and the cold of the~~
one on the house-front almost immediately opposite. BIG BROTHER
~~winter that was just past.~~
IS WATCHING YOU, the caption said, while the dark eyes

georgeorwellnovels.com

...for the sake of clarity

Clarity

- Be concise and precise ... But not cryptic
- Guide the reader through the process (and remember selective reading)

Clear Structure

- One **idea** per **paragraph**, one **topic** per (sub) **section**
- Use **short sentences** with simple structure and **short paragraphs**
- Omit unnecessary material (unused definitions are very irritating)

Guide the Reader: spoilers

The logic and structure of the paper has to be clearly communicated.

- The **introduction** should give some indications of the organization

The structure of the paper is as follows. Section 2 revises basic concepts and properties of Discrepancy-based search. Section 3 makes the generalization to AND/OR trees and presents LDSAO. In the first part of Section 4 there are preliminaries on Graphical Models concepts and the description of their AND/OR search tree. In the second part, it is shown how LDSAO instantiates to Graphical Models and it is shown its theoretical advantage over LDS. Section 5 reports experimental results on the min-sum problem. Finally, in Section 6 we discuss some limitations of our approach and some directions of future work.

Guide the Reader: connections

The logic and structure of the paper has to be clearly communicated.

- The connection between one paragraph and the next should be obvious
- **Brief summaries** at the start and/or end of each section
- Sentences **linking** one section to the next

Together these results show that the hypothesis holds linear coefficients. In the next section we address the difficulties presented by non-linear coefficients.

Guide the Reader: Math and English

.

The connection between Math and English has to be clearly communicated. Exploit the best of each **language**

- Explain how a definition, theorem,... will be used or fits in the paper

> The following property shows why, in practice, it is only possible to run the first iterations of the algorithm.
>
> **Property 1.** *The complexity of the k-th iteration of LDS searching on a binary tree of height h, and assuming $k < h/2$, is $O(h^{k+1})$.*

Guide the Reader: On Algorithms

- Avoid as much as possible **programming language notation** in your writing (algorithms can often be described in plain English)
  - If the algorithm is important, describe it in words in the text, and give its pseudo-code as a Figure.

Pseudo-code

ALGORITHM 1: BT algorithm

BTALGORITHM(G, $cliqueCost$, $mergeCost$)
$\Pi \leftarrow EnumeratePMCs(G)$ [7, 17];
$T \leftarrow \{\}$;
foreach $\Omega \in \Pi$ **do**
    foreach $D \in \mathcal{C}(G \setminus \Omega)$ **do**
        $S \leftarrow \mathrm{N}(D)$;
        $C \leftarrow$ The component of $G \setminus S$ such that $\Omega \subset S \cup C$;
        $T \leftarrow T \cup \{(\Omega, S, C)\}$;
    end
    $T \leftarrow T \cup \{(\Omega, \emptyset, V(G))\}$;
end
sort T in increasing order of $|S \cup C|$;
$dp[(S,C)] \leftarrow \infty$ for all (S, C);
foreach $(\Omega, S, C) \in T$ **do**
    $cost \leftarrow cliqueCost(\Omega, S)$;
    foreach $C' \in \mathcal{C}(G[C \setminus \Omega])$ **do**
        $S' \leftarrow \mathrm{N}(C')$;
        $cost \leftarrow mergeCost(cost, dp[(S',C')])$;
    end
    if $cost < dp[(S,C)]$ **then**
        $dp[(S, C)] \leftarrow cost$;
        $optChoice[(S, C)] \leftarrow \Omega$;

4.2 Detailed Description

Our implementation of the BT algorithm is presented in pseudocode as Algorithm 1. The implementation is mainly based on [7, 17, 23]. As mentioned, the BT algorithm works by decomposing the computation of $f(G) = f(R(\emptyset, V(G)))$ into the computation of $f(R(S, C))$ of all blocks of G. Furthermore, following Corollary 1, the value of $f(R(S, C))$ is computed as the minimum cost of $R(S, C)$ with respect to Ω over all $\Omega \in \Pi(G)$ satisfying $S \subsetneq \Omega \subset (S, C)$.

Algorithm 1 proceeds over triplets of form (Ω, S, C), where (S, C) is a block and $\Omega \in \Pi(G)$ satisfies $S \subsetneq \Omega \subset (S, C)$. The optimal cost of $R(S, C)$ with respect to Ω is computed on Lines 14–18. Whenever this cost is lower than the best known cost for $R(S, C)$ (Line 20), the value of $dp[(S, C)]$ is updated (Line 21) and Ω is stored in $optChoice[(S, C)]$ (Line 22). After processing all triplets (Lines 16–22), the value of each $dp[(S, C)]$ is equal to $f(R(S, C))$ for all blocks, and $optChoice[(S, C)]$ contains the PMC $\Omega \subset S \cup C$ that needs to be completed into a clique when constructing a optimal triangulation of $R(S, C)$. Specifically, the value of $f(G)$ is stored in $dp[(\emptyset, V(G))]$ (Line 25).

After running Algorithm 1, the optimal triangulation H can be reconstructed using a breadth-first search like procedure shown in Algorithm 2. Starting from $B = (\emptyset, V(G))$ (Line 3), the potential maximal clique Ω_B corresponding to $f(R(B))$ is completed into a clique (Line 7), and all blocks $B_i \in (B : \Omega_B)$ are added to the queue (Lines 8–10). Notice that $\Omega_B = optChoice[B]$.

We note that in order to compute the optimal cost of $R(S, C)$ for a PMC Ω, the optimal cost of all blocks in $(S, C : \Omega)$ needs to be computed. In our implementation of Algorithm 1, all triplets (Ω, S, C) are computed before the actual search (Lines 2–10), and then processed in order of increasing sizes of $S \cup C$ (Line 12). First all potential maximal cliques are enumerated using the procedure from [17] (Line 2) which in turn uses the procedure for enumerating all minimal separators from [7]. Then, for each potential maximal clique, all blocks (S, C) for which $S \subsetneq \Omega \subset (S, C)$ are initialized (Lines 4–10). The addition of the "dummy state" $(\Omega, \emptyset, V(G))$ on Line 10 is used for retrieving the optimal value $f(G)$.

We do not present pseudocode for enumerating potential maximal cliques here, as our implementation is directly based on the pseudocode of [17], using also the optimizations mentioned therein. Let G be a graph with n nodes, $v \in V(G)$ and $G' = G \setminus \{v\}$, i.e., G with the node v removed. The enumeration of potential maximal cliques is based on a characterization of $\Pi(G)$ in terms of $\Pi(G')$, $\Delta(G')$ and $\Delta(G)$. In other words, the set $\Pi(G) = \Pi(G[V_n])$ is computed by iteratively computing

Precision and Conciseness

- Be **precise**, *...less memory is likely to be required by the new*

  structure, depending on the magnitude of the numbers to be stored and the access pattern...
- **Do not dress up ideas**

  Sometimes the local network stalls completely for a few seconds. This is what we call the "Grimwade effect", discovered serendipitously during an experiment to measure the impact of server configuration on network traffic
- Shortcomings:
  - lack of economy: stalls completely
  - arrogance: "Grimwade effect", serendipitously
  - Aiming at entertainment.
  - "Paper as a diary" effect

Precision and Conciseness

- A better writing:
  ...We observed that network traffic sometimes stalls for a few seconds.
- Even better:
  ...We observed that network traffic stalls around 2% of the time during 2 *to* 5 *seconds.*

Writing Tips (from *Writing for Computer Science*)

- **Avoid double negatives**
  - *There do* **not** *seem to be any reason* **not** *to adopt the new approach*
  - (correction) The new approach is at least as good as the old and should be adopted.
- **Avoid passive voice**
  - *the following theorem can now be proved*
  - (correction) Now we can prove the following theorem
- **Avoid unneeded verbs**
  - *Local packet transmission* **was performed** *to test error rates*
  - (correction) We tested the error rates on local packet transmissions
- Prefer **short** and **familiar** words (e.g. *begin* vs *initiate*, *use* vs *utilize*,...)

Check for Ambiguity

- **Ambiguity**
  - *The compiler did not accept the program because* **it** *contained errors*
  - A safe-sex guide included "*a table on which sexual practices are safe*"
- **Partial ambiguity**: Sentences should be readable from left to right without ambiguity,
  - **Bad**: Smith remarked in a paper about the scarcity of data
  - **Bad**: In the theory of rings, groups and other algebraic structures are treated

Avoid nested sentences

- *If the machine is lightly loaded then speed is acceptable whenever the data is on local disks*
- (correction) If the machine is lightly loaded and data is on local disk then speed is acceptable.

- *We explore the classic method of look-ahead in heuristic search in the context of AND/OR search for combinatorial optimization tasks in graphical models (e.g. MAP/MPE inference in Markov networks or minimization in weighted CSPs)*

Some hints from Donald Knuth

- The statement just preceding a theorem, algorithm, etc. should be a complete sentence or should end with a colon,
  - **Bad**:
    Consider the following

Theorem

$H(x), x \geq 0$ *is continuous...*

  - **Good**:
    We can now prove the following result.

Theorem

The function $H(x)$ *defined in (2) when* $x \geq 0$ *is continuous...*

    (note that it contains three additional improvements)
  - **Even better**: Replace the sentence before the theorem by a more suggestive motivation, linking the theorem with the previous discussion.

Some hints from Donald Knuth

- Use parallelism when parallel concepts are being discussed.
  - **Bad**: Formerly, science was taught by the textbook method, while now the laboratory method is employed.
  - **Good**: Formerly, science was taught by the textbook method; now it is taught by the laboratory method.

Definitions

- Emphasize the **first occurrence** of a new word or concept
  - it shows that it is important
  - facilitates posterior identification
- Use a **consistent format** for introducing new terminology
- If the name of the concept is **long** and it is going to be used very frequently, consider introducing an **acronym** (e.g. DF for *Depth-first*)
- If a concept is going to be used in a math environment consider introducing **notation** (e.g. *A Graphical Model* $\mathcal{G}$ *is a ...*)

Some hints from Donald Knuth

On formulas and text

Most readers will skim over formulas on their first reading. Therefore, your sentences should flow smoothly when all but the simplest formulas are replaced by *blah*

The total revenue, R, made from selling widgets is given by the equation $R = pq$, where p is the price at which each widget is sold and q is the number of widgets sold. Based on past experience, we know that when widgets are priced at \$15 each, 2000 widgets are sold. We also know that for every dollar increase in price, 150 fewer widgets are sold. Hence, if the price is increased by x dollars, then the revenue is,

$$R = (15 + x)(2000 - 150x)$$

which can be simplified to,

$$R = -150x^2 - 250x + 30000$$

Some hints from Donald Knuth

- Display **important formulas on a line** by themselves. If you need to refer to some of these formulas from remote parts of the text
- Give reference numbers to all of the important ones, even if they are not referenced.
- Capitalize names like Theorem 1, Lemma 2, Algorithm 3, Figure 4.

Some hints from Donald Knuth

- Not all formulas $A = B$ are **equations**. They may also be **definitions**, **theorems**, ... Use the right name in your text
- Some people call p and **element** of P (with P being a set), and p_r and **element** of p (with p being a vector or array). It is better to call p_r a **component** or **position** of p to distinguish the two kinds of relationship.

Some hints from Donald Knuth

A very common error is a misplaced **only**,

- **Only** I hit him in the eye yesterday.
- I **only** hit him in the eye yesterday.
- I hit **only** him in the eye yesterday.
- I hit him **only** in the eye yesterday.
- I hit him in the **only** eye yesterday.
- I hit him in the eye **only** yesterday.
- I hit him in the eye yesterday **only**.