

# Specialist English: Final Examination

Rebecca J. Stones

Examination date/time: 8 January 2018, 10am

This is the final examination for Specialist English (worth 50% of the final mark).

The examination is open to paper notes (e.g., your own notes, paper copies of lecture notes and slides, dictionaries), not electronic devices (no computers, mobile phones, etc.).

Your mark  $m$  will be out of 65, and I'll scale the marks according to  $m \mapsto \lceil \min(m, 60)/1.2 \rceil$  for Master's students and  $m \mapsto \lceil m/1.3 \rceil$  for Ph.D. students.

Please write answers as if you're communicating with a human, i.e., use complete sentences, and go beyond the absolute minimum answer, and explain *why* and *how*.

## 1 Titles and beyond

[section total: 8 marks]

### Parallel Frequent Subgraph Mining Algorithm

He Yanshan

Wang Ting

XieJianli

Zhang Ming

— He et al., ICSCA, 2017.

The following questions relate to the snippet by Yu et al. (2016) above.

**P1** Which two words in the title should be combined into a compound adjective? [1 mark]

**P2** Rewrite the title to fix a grammar error: “algorithm” is a countable noun (we have one algorithm, two algorithms, etc.). [1 mark]

**P3** Which name(s) are written incorrectly, and what’s wrong with it? [1 mark]

### Mobile Underwater Acoustic Communication Based on Hyperbolic Frequency Modulation Signal

Juntao Yu

Yuanxin Xu

Yifan Gu

Rui Yu

Tongchen Wang

Zhejiang University Zhejiang University Zhejiang University Zhejiang University Zhejiang University

College of information College of information College of information College of information College of information

science & Electronic science & Electronic science & Electronic science & Electronic science & Electronic

Engineering, Zhejiang Engineering, Zhejiang Engineering, Zhejiang Engineering, Zhejiang Engineering, Zhejiang

University, Hangzhou University, Hangzhou University, Hangzhou University, Hangzhou University, Hangzhou

310027, 310027, 310027, 310027, 310027,

China China China China China

yujuntao@zju.edu.cn xuyx@zju.edu.cn guyifan@zju.edu.cn 287646082@qq.com talentwtc@163.com

— Yu et al., WUWNET, 2016.

The following questions relate to the snippet by Yu et al. (2016) above.

**P4** Identify a capitalization problem in the above snippet. [1 mark]

**P5** Describe how we can write the authors’ affiliation more succinctly. [1 mark]

**P6** Rewrite the title to correct the grammar. [1 mark]

## The Research and Implementation of Super-Resolution Reconstruction for Multi-Frame Blurring Images

— Wang and Yang, ICCC, 2011.

The following questions relate to the title by Wang and Yang (2011) above. This title has multiple grammar errors, so we will rewrite it.

**P7** Which words in this title are important, and which can we throw away? [1 mark]

**P8** Rewrite this title to be succinct and grammatically correct. [2 marks]

## 2 Abstracts

[section total: 11 marks]

**As is widely known**, the electricity industry is closely connected to the economic development. **Lots of work** has focused on the regression relationship between the economic metrics and electricity-related indexes. Studies on this relationship have great significance on economic **predication** and regulation. **Obviously**, the economic and electricity data can be treated as time series data, which is called **trajectory** in this paper. The similarity of two trajectories could reflect how much two metrics change in **the same tendency**. We **proposed** a new trajectory similarity, **which** based on the area enclosed by two trajectories. **Different to the** traditional polygon area, we keep one trajectory **stay**, and move another trajectory vertically seeking for the least area. **Meanwhile**, some economic activities would last for some time. So, there exists some time lagging between **the** economic activities and **the** electricity consumption. We **could** move the economic indicators data horizontally so that the most similarity could be found. **Therefore**, we **could** get **time offset** between the economic consumption and economic indicators. **In the experiment**, we **used** three similarity functions to compare the results. Our algorithm outperforms state-of-the-art techniques in terms of both effectiveness and efficiency.

— Xu, Xie, and Feng, *Correlation Analysis between Electricity Consumption and Economic Development*, ACM TUR-C, 2017.

The following questions relate to the abstract by Xu, Xie, and Feng (2017) above. In this abstract, I've highlighted in bold places where I would edit this abstract.

**P9** Describe a problem with writing “As is widely known, ...” and “Obviously, ...”. [1 mark]

**P10** “Lots of work ...” sounds like something children would say. How can we change this sentence to sound more formal? [1 mark]

**P11** How can we better rephrase “Different to the ...”? [1 mark]

**P12** Describe three other ways this abstract could be improved. [3 marks]

### ABSTRACT

With a bandwidth reservation mechanism, we propose that LTE can assist existing WiFi networks to improve the Quality of Experience (QoE) of wireless communication in enterprise and meanwhile increase the available spectrum. To address the associated network selection problem between WiFi and LTE, in this paper based on the 3GPP standard Access Network Discovery and Selection Function (ANDSF) framework, a context-aware solution named Extended Dynamic Enterprise ANDSF (EDE-ANDSF) is presented. It can select interfaces according to real time network conditions and cater for specific enterprise requirements.

— Xu et al., *Enterprise LTE and WiFi Interworking System ...*, ANCS, 2016.

**P13** Xu et al. (2016) introduce three acronyms in this abstract; which of these acronyms are subsequently used? [2 marks]

Experimental results demonstrated the superior performance of our method to state-of-the-art visual tracking methods [5, 12] on smartphones.

— Yang et al., *Vision-Inertial Hybrid Tracking for Robust and Efficient Augmented Reality on Smartphones*, MM, 2015.

The following questions relate to this sentence by Yang et al. (2015) above; this sentence describes the experimental results of their paper.

**P14** In describing their experiments, what are the authors specific about? [1 mark]

**P15** In describing their experiments, what is an example of something the authors **not** specific about? [1 mark]

**P16** In an abstract, what use are the reference numbers “[5, 12]”? [1 mark]

### 3 General writing

[section total: 10 marks]

All the methods were implemented using C# multithreaded programming, with the thread number being 24.

— Wang et al., *Group Matrix Factorization for Scalable Topic Modeling*, SIGIR, 2012.

**P17** Rewrite this sentence by Wang et al. (2012) to fix “the thread number”. [1 mark]

These background modeling methods can be divided into the following categories: statistical models [34, 43], cluster models [15], neural-network models [8, 25], sparse models [19], robust subspace models [7, 33], etc.

— Xie et al., *A Survey of Background Modeling Based on Robust Subspace Learning via Sparse and Low-rank Matrix Decomposition*, ICIMCS, 2016. (Modified.)

**P18** The above sentence by Xie et al. (2016) is an example of a reference-dense sentence. What do reference-dense sentences achieve? [1 mark]

to encode a frequent length. So the cost required to encode the lengths of all segments is

$$C_s'' = \sum_{k=1}^{|I|} (|I_k| - 1) \cdot (-\log p_{\sigma(k)\sigma(k)})$$

The total cost of encoding the segmentation is the summation of  $C_s'$  and  $C_s''$ :

$$C_s = C_s' + C_s''$$

We use a set of probabilities,  $p_i$  and  $p_{ij}$ ,  $1 \leq i, j \leq |\mathcal{M}|$ , to describe the dynamics of the system. These probabilities

— Wang, Wang, Liu, and Wang, *An Algorithmic Approach to Event Summarization*, SIGMOD, 2010.

The following questions relate to the snippet by Wang, Wang, Liu, and Wang (2010) above.

**P19** Sentences end in a full stop. Where should we put the full stop in sentence beginning “So the cost ...”? [1 mark]

**P20** Rewrite the sentence beginning “The total cost ...” to be grammatically correct and succinct. [2 marks]

**P21** Identify two (or more) problems with the sentence beginning “We uses a ...”. [2 marks]

As shown in Figure 1, this approach is composed of two parts: classifier building and re-ranking.

— Ye, Xie, Zou, and Chen, *Interrogative-Guided Re-Ranking for Question-Oriented Software Text Retrieval*, ASE, 2014.

**P22** Rewrite the above sentence by Ye, Xie, Zou, and Chen (2014) to avoid using “as shown” (which has multiple meanings). [1 mark]

The prediction model using only last used App also performs dramatically well, indicating the high correlation between two sequentially used Apps.

— Zhu et al., *Exploiting User Context and Network Information for Mobile Application Usage Prediction*, HotPOST, 2015.

**P23** Which word in the above sentence by Zhu et al. (2015) is over-the-top? [1 mark]

People under different network situations may launch different kinds of apps. For example, a person may start several download tasks when connected to Wi-Fi, while this hardly ever happens in mobile networks as the network volume is limited and its cost is relatively high. Meanwhile, people only watch videos when the network speed is fast enough, otherwise the videos may get stuck and result in the termination of the app.

— Zhu et al., *Exploiting User Context and Network Information for Mobile Application Usage Prediction*, HotPOST, 2015. (Modified.)

**P24** Which word in the above snippet by Zhu et al. (2015) is incorrect? [1 mark]

## 4 Introductory material

[section total: 9 marks]

path of data requests, which are referred to as coherence overhead in this work. In particular data blocks both shared and written (referred to as *shared written blocks*) must be handled with care to guarantee correctness and consistency among cores. A write to a shared written block produces

— Hu et al., *Exploiting Transmission Lines on Heterogeneous Networks-on-Chip to Improve the Adaptivity and Efficiency of Cache Coherence*, NOCS, 2015.

**P25** Why is “shared written blocks” written in italics by Hu et al. (2015) above? [1 mark]

**P26** Rewrite the sentence beginning “In particular ...” to be grammatically correct. [2 marks]

timing for one manufactured die.

Previous works have shown that post-silicon measurements can significantly help reduce the variance in predicting the circuit timing in the presence of process variations. For example, adding generic on-chip test structures such as ring-oscillators as well as custom

— Li, Davoodi, and Xie, *Custom On-Chip Sensors for Post-Silicon Failing Path Isolation in the Presence of Process Variations*, DATE, 2012.

**P27** How can we fix the grammar problem “Previous works ...” in this snippet by Li, Davoodi, and Xie (2012)? [1 mark]

## 1.2 Limitations of Prior Art

Prior work on energy saving of smart phones can be classified into the following three parts: energy consumption of hardware (Fan et al. 2007; Bellosa et al. 2003; Rajan et al. 2006; Balasubramanian et al. 2009), power consumption models, and energy-saving schemes for specific applications. *For hardware*, Chen et al. (2013a) analyze the power consumption of AMOLED displays in multimedia applications and reveal that camera recording incurs high power cost. LiKamWa et al. (2013) report the experimental and analytical characterization of CMOS image sensors and reveal two energy-proportional mechanisms for energy saving. *For models*, Dong and Zhong (2011) propose Sesame, with which a mobile system constructs an energy model of itself without any external assistance. Xu et al. (2013) propose a new way called V-edge to generate power models based on battery voltage dynamics. *For specific applications*, Han et al. (2013) study the energy cost made by human-screen interaction, such as scrolling on the screen. Dietrich and Chakraborty (2013) detect the game’s current state and decrease the processor’s voltage and frequency whenever possible to save energy. Hu et al. (2013) propose a Mobility-Assisted User Contact detection algorithm (MAUC), which utilizes the accelerometer of the phone to detect user movements for energy-saving. The Bluetooth scans only when user movements have a high possibility to cause contact changes. LiKamWa et al. (2013) improve the energy efficiency of image sensors based on hardware modifications. There are fewer energy-saving schemes for photographing.

Being different from these prior work, we aim to propose an energy-saving scheme for photographing. We aim to recognize the user’s activity and reduce unnecessary energy cost when the user is not taking photos. The scheme does not need hardware modifications and user interaction, to guarantee a good user experience.

— Yin et al., *Tracking Human Motions in Photography: A Context-Aware Energy-Saving Scheme for Smart Phones*, ACM Trans. Sens. Netw., 2017.

The following questions relate to the snippet by Yin et al. (2017) above.

**P28** How do Yin et al. (2017) categorize relevant references? [1 mark]

**P29** Where do Yin et al. (2017) compare these references to the present paper? [1 mark]

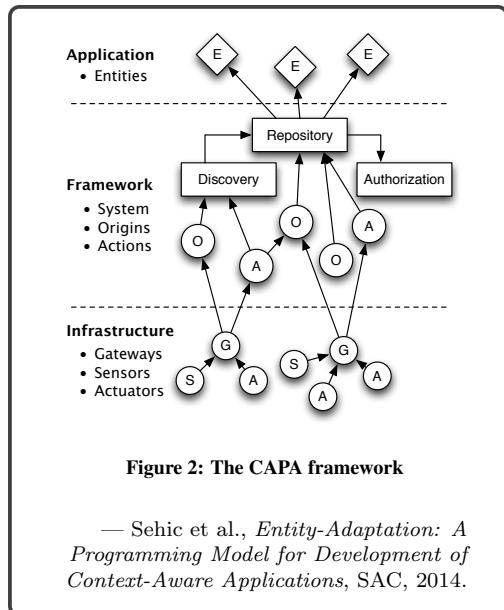
**P30** Is the tense used to refer to past work appropriate? [1 mark]

**P31** Is the tense used to refer to past work consistent? [1 mark]

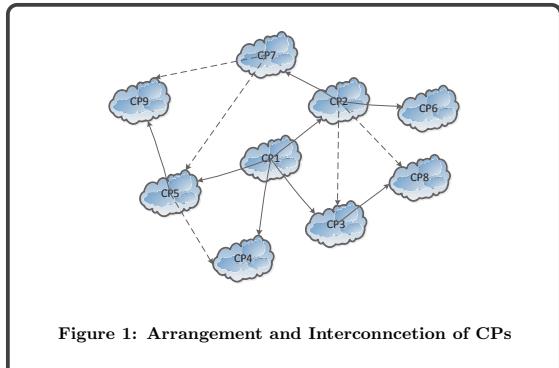
**P32** What’s wrong with writing “aim to propose”? [1 mark]

## 5 Visualizations

[section total: 7 marks]



— Sehic et al., *Entity-Adaptation: A Programming Model for Development of Context-Aware Applications*, SAC, 2014.



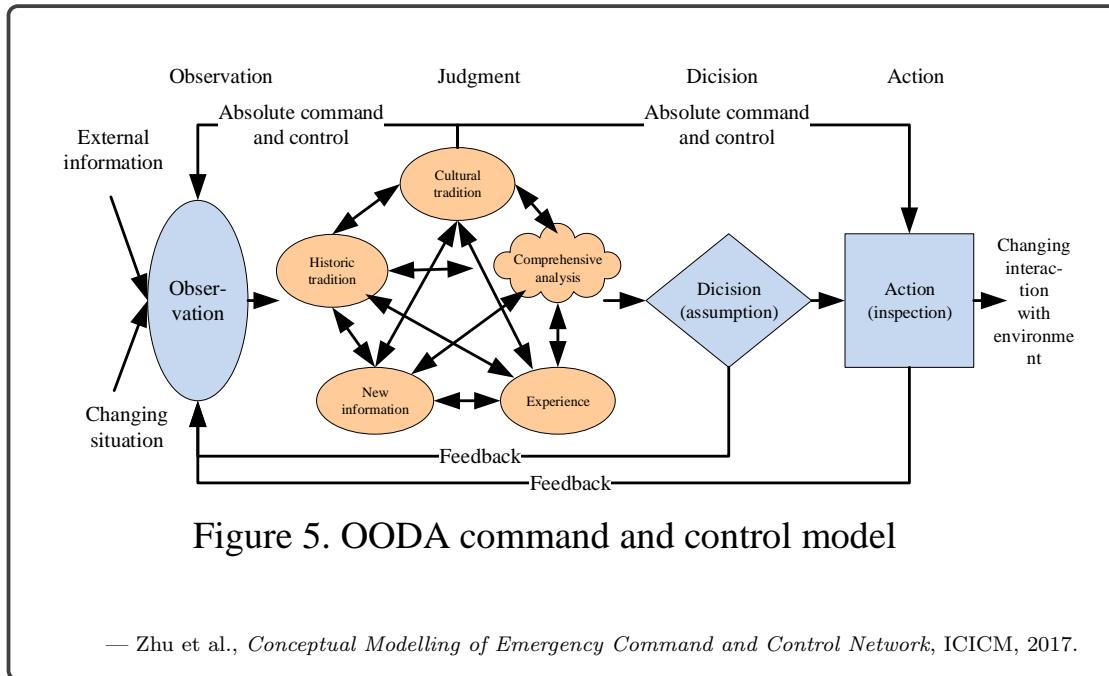
— Das et al., *A QoS and Profit Aware Cloud Confederation Model for IaaS Service Providers*, IMCOM, 2014.

These questions relate both figures above, one by Sehic et al. (2016), and one by Das et al. (2014).

**P33** Which of these figures is “self-contained”? [1 mark]

**P34** Which of these figures is “oversimplified” and which is “overcomplicated”? How do you come to that conclusion? [2 marks]

**P35** Describe one way Sehic et al. (2014) minimizes the words in their figure? [1 mark]



— Zhu et al., *Conceptual Modelling of Emergency Command and Control Network*, ICICM, 2017.

**P36** Identify three (or more) ways to improve the presentation of Figure 5 in Zhu et al. (2017). [3 marks]

## 6 Algorithms

[section total: 4 marks]

---

**Algorithm 1** Interference-Aware Scheduling

---

```

1: When a heartbeat is received from a free node  $n$ :
2: Collect system information  $N_{info}$ ;
3: Given a job  $j$ 
4: Fetch task's profile  $Tasks_j$ ;
5: Predict the slow down rate  $\hat{S}_j = Model(N_{info}, Tasks_j)$ ;
6: Get the number of running slots  $Z_R$  and idle slots  $Z_I$ 
7: if  $Z_I > 1$  for  $k$  seconds then
8:   // use dynamic threshold
9:   if  $(Z_R + 1)/\hat{S}_j > Z_R/H_d$  and  $\hat{S}_j > H_d$  then
10:    update  $H_d = \hat{S}_j$  and return the accepted job  $j$ ;
11:   else
12:     if  $\hat{S}_j \leq H_d$  then
13:       update  $H_d = \min(H_d - 1, \hat{S}_j)$  and return the accepted
           job  $j$ ;
14:     else
15:       reject job  $j$ 
16:     end if
17:   end if
18: else
19:    $H_d = H$ ; //use the predefined threshold
20:   if  $\hat{S} \leq H$  then
21:     return the accepted job  $j$ 
22:   else
23:     reject job  $j$ ;
24:   end if
25: end if

```

---

— Bu, Rao, and Xu, *Interference and Locality-Aware Task Scheduling ...*, HPDC, 2013.

**P37** Identify four ways Algorithm 1 by Bu, Rao, and Xu (2013) can be improved. [4 marks]

## 7 Tables

[section total: 5 marks]

Variable	Interpretation
$\mathcal{U}, \mathcal{V}$	the set of users and POIs
$\mathcal{R}, \mathcal{T}, \mathcal{W}$	the set of regions, time slots and words
$l_v, \mathcal{W}_v,$	POI $v$ 's location and content word set
$\tau, t,$	the timestamp of check-in and its discretized time slot
$\mathcal{D}_u$	the profile of user $u$
$\mathbb{R}^d$	$d$ dimensional latent space
$\vec{u}_\tau$	time-aware user preference embedding
$\vec{v}, \vec{r}, \vec{t}, \vec{w}$	embeddings of POI $v$ , region $r$ , time slot $t$ and word $w$ , respectively
$\Delta T$	the time interval

**Table 1: Notations used in this paper.**

— Xie et al., *Learning Graph-based POI Embedding ...*, CIKM, 2016.

These questions relate to the figure by Xie et al. (2016) above.

**P38** How does Table 1 help the reader? [1 mark]

**P39** Identify a (common) grammar error. [1 mark]

### 3. EXPERIMENTS

We crawled data of Movie, Music and Book from Douban website, and extracted the registered users who perform sequential behaviors on at least two domains. Finally we constructed 6 transfer learning recommendation problems (i.e., 3 pairs of data sets). We conduct extensive experiments to demonstrate the effectiveness of the proposed CDNST, and find that the temporal property of sequential data affects the performance of CDNST.

**Table 1: Recommendation Performance**

		OF(OF_U)	MC(MC_U)	NSM(NSM_U)	CDNST
$A \rightarrow B$	MRR	0.1601(0.1522)	0.2015(0.1779)	0.3128(0.3017)	<b>0.3623</b>
	nDCG@15	0.2153(0.2047)	0.2677(0.2299)	0.3821(0.3673)	<b>0.4363</b>
	p@3	0.1044(0.0937)	0.1409(0.1203)	0.2822(0.2736)	<b>0.3325*</b>
$B \rightarrow A$	MRR	0.3982(0.2413)	0.4135(0.2575)	<b>0.5644</b> (0.3180)	0.5014
	nDCG@15	0.4998(0.3279)	0.5125(0.3715)	<b>0.6489</b> (0.3945)	0.5687
	p@3	0.3373(0.2100)	0.3649(0.2241)	<b>0.5488</b> (0.2992)	0.4797

OF (Order by Frequency) gives a recommendation list according to the frequency in the individual's historical behavior sequence, while for OF\_U we compute the frequency in both source and target domains. MC (Markon Chain) [2] models sequential behaviors in target domain by learning a transition graph and performing predictions (In our report, the factorization was set at 20 for comparison), NSM (Novel Seeking Model) [4] is a data-driven model to predict the behaviour (In our report, the novelty seeking level was set at 9). MC\_U and NSM\_U are similarly defined with OF\_U.

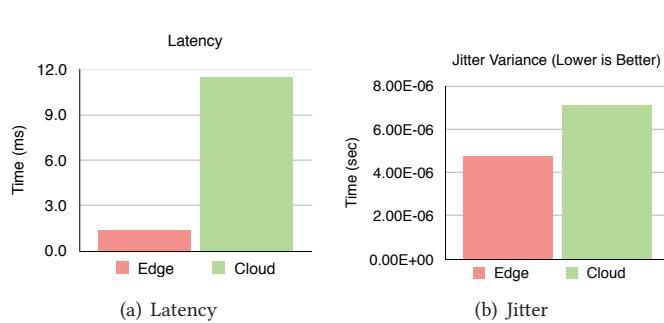
For all compared methods, they give a list of recommended choices with prediction probabilities, according to which we sort the candidate choices in descending order. In our experiments, the widely used evaluation metrics of nDCG, MRR and Precision are adopted to evaluate the performance of all algorithms.

— Zhuang et al., *Sequential Transfer Learning: Cross-domain Novelty Seeking Trait Mining for Recommendation ...*, WWW, 2017.

**P40** Identify three ways Table 1 by Zhuang et al. (2017) could be improved. [3 mark]

## 8 Experimental Results

[section total: 5 marks]



**Figure 2: EC2 Instance vs. Edge Server**

— Aske and Zhao, *An Actor-Based Framework for Edge Computing*, UCC, 2017.

These questions relate to the figure by Aske and Zhao (2017) above.

**P41** How many data points does the the above figure contain? [1 mark]

**P42** What would be a more appropriate way to represent this data? [1 mark]

**P43** Identify an inconsistency between the two subfigures. [1 mark]

**P44** How can we improve the mathematical presentation in subfigure “(a) Latency”? [1 mark]

**P45** How can we improve the mathematical presentation in subfigure “(b) Jitter”? [1 mark]

## 9 References

[section total: 5 marks]

Yang *et al* also leveraged the frequent-value locality to reduce the access energy of caches [17]. In their work, frequent values were read

— Sun, Niu, Ouyang, and Xie, *A Frequent-Value Based PRAM Memory Architecture*, ASPDAC, 2011.

[17] J. Yang and R. Gupta. Energy efficient frequent value data cache design. In *Proceedings of MICRO 2002*, pages 197–207.

— Sun, Niu, Ouyang, and Xie, *A Frequent-Value Based PRAM Memory Architecture*, ASPDAC, 2011.

**P46** Identify two problems in how Sun, Niu, Ouyang, and Xie (2011) reference Yang et al. (2012) above. [2 marks]

```
@INPROCEEDINGS{Ding09usinggraphics,  
author = {Shuai Ding and Josh Attenberg and Torsten Suel and Shuai Ding  
and Jinru He and Hao Yan and Torsten Suel and Hao Yan and Shuai Ding  
and Torsten Suel and Hao Yan and Shuai Ding and Torsten Suel  
and Work Experience},  
title = {Using graphics processors for high performance  
ir query processing},  
booktitle = {In WWW},  
year = {2009}  
}
```

[1] S. DING, J. ATTENBERG, T. SUEL, S. DING, J. HE, H. YAN, T. SUEL,  
H. YAN, S. DING, T. SUEL, H. YAN, S. DING, T. SUEL, AND W. EXPE-  
RIENCE, *Using graphics processors for high performance ir query processing*,  
in In WWW, 2009.

### Using Graphics Processors for High Performance IR Query Processing

Shuai Ding Polytechnic Inst. of NYU Brooklyn, NY 11201 sding@cis.poly.edu	Jinru He Polytechnic Inst. of NYU Brooklyn, NY 11201 jhe@cis.poly.edu	Hao Yan Polytechnic Inst. of NYU Brooklyn, NY 11201 hyan@cis.poly.edu	Torsten Suel <sup>*</sup> Yahoo! Research Sunnyvale, CA 94089 suel@poly.edu
--	--	--	--

The above three boxes respectively give (a) a BibTeX entry from CiteSeerX, (b) the compiled output, and (c) the header from the paper.

**P47** Identify three (or more) errors in the above reference.

[3 marks]