

# Visualisierung verteilter Systeme

---

## End-User Software Visualizations for Fault Localization

Referentin Maren Settekorn

# Die Autoren

---

- Margaret M. Burnett, Professorin am Institut für Informatik der Oregon State University
- Curtis Cook Professor Emeritus
- Joseph R. Ruthruff Student
- Shrinu Prabhakararao Student
- Marc Fisher II Student
- Eugene Rogan Creswick Student
- Martin Main Student

# Inhalt

---

- Motivation
- WYSIWYT
- Forms/3
- Techniken zur Visualisierung
- Das Experiment
- Ergebnisse

# Motivation

---

- 2005 ca 2,75 Millionen professionelle Programmierer in den USA
- dagegen ca 55 Millionen Endbenutzer die programmieren
- Für Profis gibt es viele Tools, die bei der Fehlersuche helfen
- Endbenutzer haben weder die Tools noch die Ausbildung

# Ziel der Gruppe

---

- Eine Software, die mit Spreadsheets-programmierenden Benutzern bei der Fehlersuche hilft
- Benutzer soll mögliche Fehler durch die Farbe der entsprechenden Felder finden

# WYSIWYT(I)

---

- What You See Is What You Test
- An Oregon State University entwickelte Testmethode
- Coloriert die Umrandung einer Zelle je nachdem, ob diese getestet ist oder nicht
- Relativ billig und schnell

# WYSIWYT(II)

---

- Annahme: Tabellen wachsen nach und nach  
-> Benutzer testen auch nach und nach
- In Forms/3 kann Benutzer eine Zelle abhaken, wenn sie korrekt ist, oder x-en
- Ungetestete Zellen sind rot, abgehakte und deren Erzeuger und Verbraucher blau

# Forms/3

*Forms/3*

- An der Oregon State University entwickelt
- Zur Zeit in Version 1.1 (June 19, 1998)
- Läuft nur auf Sun/Solaris und Hewlett-Packard/HP-UX
- Website:

<http://web.engr.oregonstate.edu/~burnett/Forms3/forms3.html>



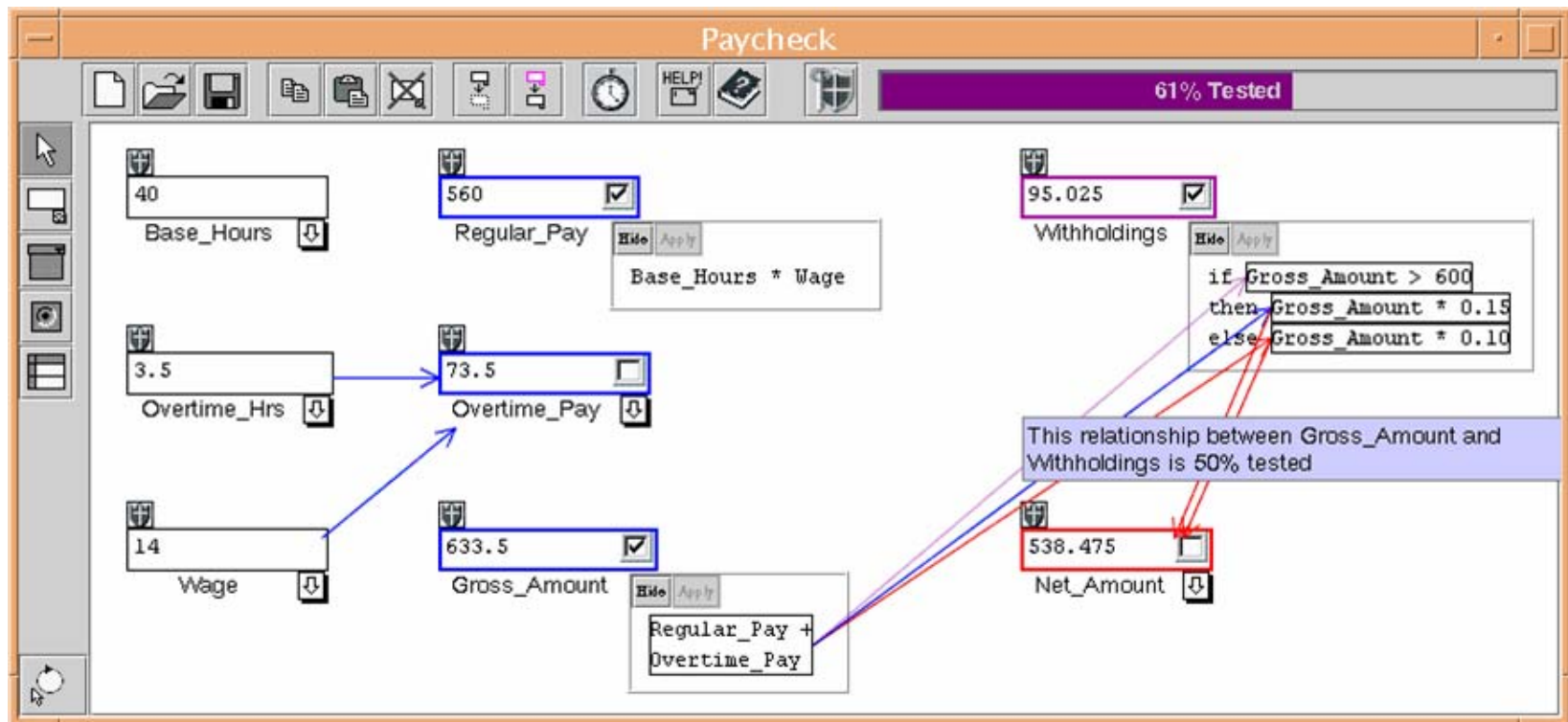
# Features von Forms/3

- Darstellung nicht als Tabelle (wie in Excel) sondern als einzelne Zellen
- Man kann durch einen Klick eine neue Zelle hinzufügen
- Formel einer Zelle festlegen, unter anderem
  - Konstante, Strings, booleans
  - algebraische Ausdrücke mit + - \* / und mod,
  - If-then-else

# Features

- Einfärben von Zelumrandungen
  - Rot ungetestet, blau fertig getestet
- Darstellung von farbigen Datenflusspfeilen
  - Definition : einer Zelle wird explizit ein Wert zugewiesen
  - Use : der Wert einer Zelle wird von einer anderen benutzt
  - eine Beziehung ist erst fertig getestet, wenn alle definition-use-Paare getestet sind

# Beispiel



# Beurteilung von Forms/3

---

- + Zellen bei „kleinen“ spreadsheets vielleicht übersichtlicher als Tabellenkalkulator
- + Sensibilisiert Benutzer für Fehlersuche
- Wird aber schnell unübersichtlicher wenn es viele Zellen gibt
- Bedienung relativ unkomfortabel
- Keine neue Version zum download

# Ziel

---

- Entwicklung eines Programms, das das Auffinden von Fehlern erleichtert
- Benutzer kann Zellen abhaken oder als fehlerhaft markieren
- Jede Markierung beeinflusst die Farbe dieser und der verbundenen Zellen (Produzenten und Konsumenten)

# Test-Count Technik

---

- Produzenten einer Zelle:
  - tragen zu dem Wert der Zelle bei
  - gehören zu vorhergehender Zeitscheibe
- Konsumenten einer Zelle:
  - nutzen den Wert der Zelle
  - gehören zu späterer Zeitscheibe

# Eigenschaft 1

---

- Wenn Zelle  $Z$  oder einer der Konsumenten von  $Z$  als fehlerhaft markiert ist, ist die Wahrscheinlichkeit dafür, dass  $Z$  fehlerhaft ist, größer als 0.
- Eine Zelle, die zu einem falschen Wert beigetragen hat, kann fehlerhaft sein.

# Eigenschaft 2

---

- Die Fehlerwahrscheinlichkeit von Zelle  $Z$  ist proportional zu der Anzahl der fehlerhaften Tests von  $Z$
- Wenn eine Zelle zu mehr fehlerhaften Werten beigetragen haben kann, als eine andere, ist es wahrscheinlicher, dass sie fehlerhaft ist.



# Eigenschaft 3

---

- Die Fehlerwahrscheinlichkeit von Zelle  $Z$  ist umgekehrt proportional zu der Anzahl der erfolgreichen Tests von  $Z$
- Wenn eine Zelle zu mehr korrekten Werten beigetragen haben kann, als eine andere, ist es wahrscheinlicher, dass sie korrekt ist.

# Berechnung der Fehlerwahrscheinlichkeit

- Anzahl fehlerhafter Tests  $a_f$
- Anzahl korrekter Tests  $a_k$
- Fehlerwahrscheinlichkeit  $f$
- $f(Z) = \max(1, 2 * a_f - a_k)$
- Fehlerwahrscheinlichkeit niedrig 1 bis 2, mittel 3 bis 4, hoch 5 bis 9 und sehr hoch 10 und drüber

# Beispiel Test-Count Technik

Paycheck

15% Tested

40	560	57.925
Base_Hours	Regular_Pay	Withholdings
	This cell's formula is somewhat likely to have an error. If the formula is wrong, fix it. Otherwise, try looking at other highlighted cells.	
3.5	19.25	
Overtime_Hrs	Overtime_Pay	
	Hide Apply	
	Overtime_Hrs * 1.5 + Wage	
14	579.25	521.325
Wage	Gross_Amount	Net_Amount
	Hide Apply	
	Regular_Pay + Overtime_Pay	

# Kosten für Test-Count(I)

- Benutzer kann konstante Werte oder Formeln ändern, oder Zellen markieren
  - Erfordert Updates in  $O(1)$  Zeit
- Beziehungen zwischen Zellen werden in beide Richtungen gespeichert
  - Zeit zum aktualisieren der Datenstruktur nach hinzufügen einer Testmarke:  $O(u \cdot p)$ 
    - $u$  maximale Anzahl an Referenzen einer Zelle
    - $p$  Anzahl der Produzenten von  $Z$

# Kosten für Test-Count(II)

- Zeit zum Ändern einer Formel :  
 $O(mc * \max(u, \text{Preis der Änderungsoperation}))$ 
  - $m$  # der Markierungen die  $Z$  erreichen
  - $c$  # der Konsumenten die durch die Änderung der Zelle betroffen sind
  - $u$  maximale # an Referenzen einer Zelle

# Blocking Technik Eigenschaft 4

---

- Wenn Zelle  $Z$  als fehlerhaft markiert ist, wird die Verbreitung der Wirkung eines korrekten Tests an einem von  $Z$ 's Konsumenten zu den Produzenten von  $Z$  abgeblockt

# Eigenschaft 5

---

- Wenn Zelle  $Z$  als korrekt markiert ist, wird die Verbreitung der Wirkung eines fehlerhaften Tests an einem von  $Z$ 's Konsumenten zu den Produzenten von  $Z$  abgeblockt
- Ausnahme: Wenn Zelle  $Z$  oder einer der Konsumenten von  $Z$  als fehlerhaft markiert ist, ist die Wahrscheinlichkeit dafür, dass  $Z$  fehlerhaft ist, größer als 0.

# Berechnung der Fehlerwahrscheinlichkeit

- Hat Zelle Z keine fehlerhaften Tests ist die Fehlerwahrscheinlichkeit  $f$  null.
- Hat Z fehlerhafte Tests, die aber nicht erreichbar sind, ist  $f$  niedrig
- $f = \max(1, aef - \downarrow(aek/2))$ 
  - $aef$  Anzahl erreichbarer fehlerhafter Tests
  - $aek$  Anzahl erreichbarer korrekter Tests



# Beispiel Blocking Technik

Paycheck

15% Tested

40	550	57.925
Base_Hours	Regular_Pay	Withholdings
	This cell's formula could have an error. If the formula is wrong, fix it. Otherwise, try looking at other highlighted cells.	
3.5	19.25	
Overtime_Hrs	Overtime_Pay	
	Overtime_Hrs * 1.5 + Wage	
14	579.25	521.325
Wage	Gross_Amount	Net_Amount
	Regular_Pay + Overtime_Pay	

# Kosten der Blocking Technik

- Kosten für Änderung einer Formel:  
 $O(p+em^2)$
- Kosten für Zufügen einer Markierung  
 $O((p+e)m^2)$ 
  - $p$  # der Produzenten von  $Z$
  - $e$  # der Abhängigkeiten in der  $Z$  vorhergehenden Zeitscheibe
  - $m$  # der Markierungen im ganzen Programm

# Nearest Consumers Technik

---

- Greedy Algorithmus
- Soll alle 5 genannten Eigenschaften mit geringen Kosten erfüllen
- Änderungen nach einer Markierung berücksichtigen nur die direkten Konsumenten

# Nearest Consumers Technik

---

- Wenn Zelle Z markiert wird, errechnet man die Fehlerwahrscheinlichkeit von Z aus der Markierung der Zelle und den Fehlerwahrscheinlichkeiten von Z's direkten Konsumenten.
- Dann wird die Fehlerwahrscheinlichkeit von Z's direkten Produzenten errechnet

# Berechnung der Fehlerwahrscheinlichkeit

- DK seien die direkten Konsumenten von Z
- $x$  # der X-Marken in DK
- $y$  # der korrekten Marken in DK
- $z = 1$  falls
  - $x > 1$  und  $y = 0$  oder
  - $x > y$  und  $y > 0$  oder
  - $x > y$  und Z fehlerhaft
- $z = 0$  sonst

# Berechnung der Fehlerwahrscheinlichkeit

- Fehlerwahrscheinlichkeit  $f$  von Zelle  $Z$ :  
 $f(Z) = \text{durchschnitt}(f(DK)) + z$
- Ausnahmen:
  - $Z$  als korrekt markiert  $\rightarrow f(Z)$  niedrig
  - $Z$  als fehlerhaft markiert und  $\text{durchschnitt}(f(DK)) < \text{mittel} \rightarrow f(Z)$  mittel
  - $Z$  als fehlerhaft markiert  $\rightarrow f(Z)$  und  $f(DK)$  können nicht kleiner werden und umgekehrt

# Beispiel Nearest Consumers

Paycheck

15% Tested

40	560	57.925
Base_Hours	Regular_Pay	Withholdings
	Error Likelihood: VERY LOW -- Check for formula errors, OR look at other tinted cells, OR check and X other cells to get more feedback.	
3.5	19.25	
Overtime_Hrs	Overtime_Pay	
	Overtime_Hrs * 1.5 + Wage	
14	579.25	521.325
Wage	Gross_Amount	Net_Amount
	Regular_Pay + Overtime_Pay	

# Experiment

---

- Ziel:

Beantwortung der Frage, welche der für Endbenutzer in interaktiver Umgebung gedachten vorgestellten Techniken soll man verfolgen und wie müssen sie verändert werden?



# Experiment

---

- Für das Experiment wurden Teilnehmer eines EDV-Kurses rekrutiert
- Den Testern wurden 2 verschiedene Forms/3-Dokumente, Change und Grades vorgesetzt, in denen sie die 3 festen Fehler finden sollten
- Tester durften keine Formeln ändern

# Change

---

- Errechnet die optimale Anzahl an Dollar, Vierteldollar, 10-Cent-Stücken, 5-Cent-Stücken und 1-Cent-Stücken, die man beim Wechseln eines Gefäßes voller Pennies zurückbekommt
- Begrenzter Datenfluss und nur eine mögliche Ausgabe

# Grades

---

- Berechnet eine Note (A, B, C, D, F) aus den bei 4 Quizfragen und einer Zusatzfrage erreichten Punkten.
- Es gibt mehrere relativ unabhängige Datenfluss-Pfade.

# Ergebnisse

---

- Effizienz – Wie gut werden korrekte Zellen visuell von fehlerbehafteten unterschieden?
- Robustheit – Wie gut ist die Unterscheidung, wenn der Benutzer korrekte Zellen als falsch markiert und umgekehrt?
- Verhältnis zwischen Effizienz und Kosten

# Effizienz

---

- Untersuchung der Unterscheidbarkeit zwischen falschen und korrekten Zellen
  - zu Beginn der Testphase, nachdem die erste Zelle als fehlerhaft markiert wurde
  - am Ende der Testphase

# Effizienz – Hypothese 1

---

- Zu dem Zeitpunkt, an dem die erste Zelle als fehlerhaft markiert wird, gibt es keinen signifikanten Unterschied zwischen der Effizienz der drei Techniken zur Fehlerlokalisierung.

# Effizienz – Hypothese 2

---

- Zu dem Zeitpunkt, da alle Tests abgeschlossen sind, gibt es keinen signifikanten Unterschied zwischen der Effizienz der drei Techniken.

# Effizienz - Change

	Erstes X			Letztes X		
	TC	B	NC	TC	B	NC
Change	0.212	-0.008	0.201	-0.466	-0.227	-0.152

- Beim ersten Feedback liefert die Blocking Technik das beste Ergebnis
- Am Ende des Testvorgangs ist die Test-Count Technik deutlich besser als die anderen



# Effizienz - Grades

	Erstes X			Letztes X		
	TC	B	NC	TC	B	NC
Grades	-0.016	-0.093	-0.190	-0.151	-0.031	0.081

- Am Anfang war die Nearest-Consumer Technik deutlich besser als die anderen
- Am Ende ist auch hier die Test-Count Technik am besten

# Ergebnisse Effizienz

---

- Beide Hypothesen sind falsch, es gibt große Unterschiede zwischen den Techniken
- Am Ende des Testens schnitt die Test-Count Technik am besten ab
- Für das erste Feedback scheint Blocking bei nur einem Pfad am geeignetsten zu sein, bei vielen Datenpfaden die Nearest-Consumer Technik

# Effizienz – dunkelste Zelle tatsächlich fehlerhaft?

	korrekt			fälschlich		
	TC	B	NC	TC	B	NC
Change	18.18%	13.64%	31.82%	11.36%	13.64%	9.09%
Grades	44.19%	51,16%	39.53%	6,98%	11.63%	16.28%

- Bei Changes mit Nearest-Consumers ist bei ca 30% der Tests die dunkelste Zelle fehlerbehaftet
- Bei Grades bei Blocking Technik in etwa 50% der Tests

# Robustheit

---

- Die Benutzer können Zellen, die korrekt sind, als fehlerhaft markieren und umgekehrt
- Perfekte Testläufe enthalten keine falschen Markierungen
- Unperfekte Testläufe enthalten fälschlich platzierte Markierungen

# Robustheit – Hypothese 3

---

- Wenn die erste Zelle als falsch markiert wird, gibt es für perfekte Testläufe keine signifikanten Unterschiede in der Effizienz der drei Techniken.

# Robustheit – Hypothese 4

---

- Wenn alle Tests komplett sind, gibt es bei perfekten Testläufen keine signifikanten Effizienzunterschiede zwischen den Techniken.

# Robustheit – Hypothese 5

---

- Zu der Zeit, da alle Tests beendet sind, gibt es bei unperfekten Testläufen keine signifikanten Unterschiede in der Effizienz der Techniken.

# Perfekte Testläufe

	Erstes X			Letztes X		
	TC	B	NC	TC	B	NC
Change	0.089	-0.078	0.078	-0.667	-0.389	-0.456
Grades	-0.296	-0.333	-0.722	-0.037	0	0.148

- Beim Testende ist die Test-Count Technik gut
- Zu Beginn der Testphase ist bei Changes die Blocking Technik am besten, bei Grades die Nearest-Consumer Technik



# Unperfekte Testläufe

	TC	B	NC
Change	-0.362	-0.144	0.006
Grades	-0.181	-0.039	0.064

- Die Test-Count Technik schneidet am besten ab

# Ergebnisse Robustheit

---

- Hypothese 4 könnte zutreffend sein, die anderen beiden müssen fallengelassen werden
- Für perfekte Testläufe gibt es zur Zeit, da die Testphase abgeschlossen ist, nur geringe Unterschiede zwischen den Techniken
- Die Test-Count Technik ist gegen Ende des Tests am robustesten

# Ergebnisse des Experiments

---

- Bei den letzten Tests liefert die Test-Count Technik stets das beste Ergebnis, die Historie, die dabei gespeichert wird, erscheint also sinnvoll
- Am Anfang der Testphase liefern die Blocking oder die Nearest-Consumer Technik bessere Ergebnisse

# Ergebnisse des Experiments

---

- Es scheint sinnvoll, die billige Nearest-Consumer Technik und die Blocking Technik für die frühen Phasen der Tests zu verwenden
- Die teure Test-Count Technik könnte dann in den späten Phasen des Tests gute Dienste leisten

# Beurteilung - Vorteile

---

- Einfärben möglicherweise fehlerbehafteter Zellen kann bei der Fehlersuche helfen
- Eingefärbte Datenflusspfeile und leicht einzublendende Formeln helfen bei der Lokalisierung des Fehlers

# Beurteilung - Nachteile

---

- Zellen mit konstantem Wert führen häufig zu unerwarteten Werten, werden aber nicht eingefärbt
- Benutzer können einen Schock bekommen, wenn plötzlich sehr viele Zellen rot angezeigt werden
- Die vorgestellten Techniken liefern noch keine wirklich guten Werte für Effizienz und Robustheit
- Fehlerhafte und korrekte Zellen werden eingefärbt

# Quellen

---

- J. Ruthruff et al.: End-user software visualizations for fault localization. Proc. 1. ACM Symp. on Software Visualization, 2003.
- Karen Rothermel's MS thesis (2000): "Empirical Studies of a WYSIWYT Testing Methodology". OSU readers:  
<http://web.engr.oregonstate.edu/~burnett/Forms3/theses.html>
- Die Homepage von Forms/3  
<http://web.engr.oregonstate.edu/~burnett/Forms3/forms3.html>



Das wars von mir

Noch Fragen?