Counting Votes and the Attempt to Replicate Human Interpretation

LABELNOISE'2017
December 1, 2017

Professor Daniel Lopresti
Computer Science and Engineering
Lehigh University
Bethlehem, PA, USA

Motivation

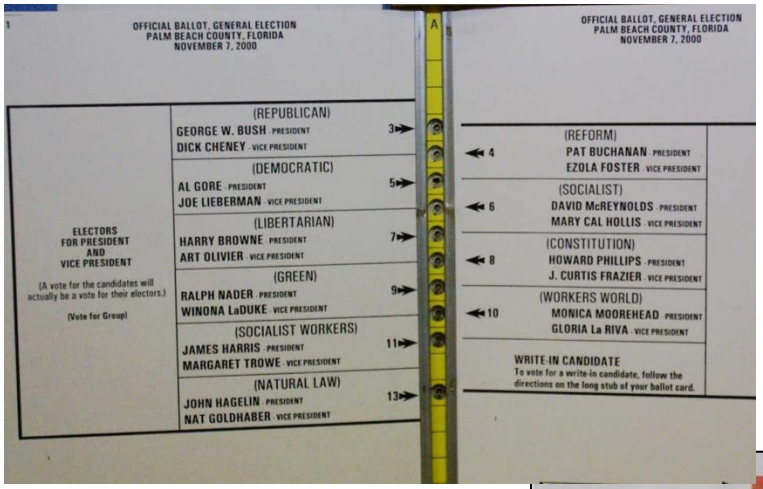
Attempt to interest you in an noisy label application area:

- Counting votes – replicating human interpretation

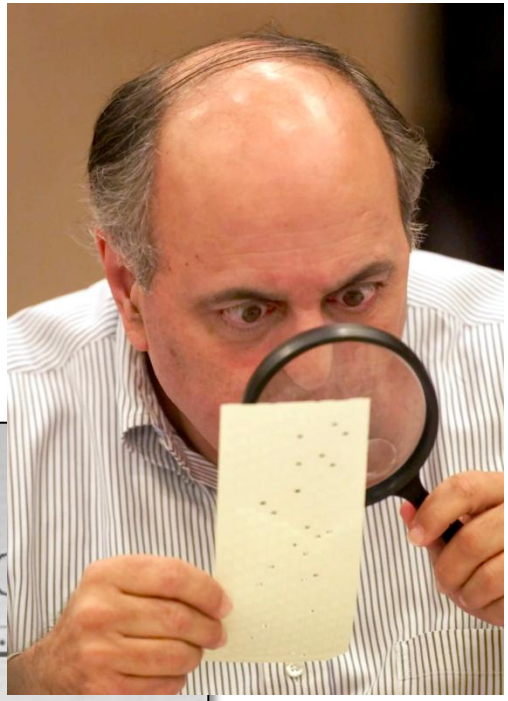
Key features:

- Vitally important application, drawn from real life
- Valuable lessons to learn that can be applied elsewhere
- Wonderful opportunity to apply our talents
- Work still needed to frame problem, outline next steps

How did US get where we are today?

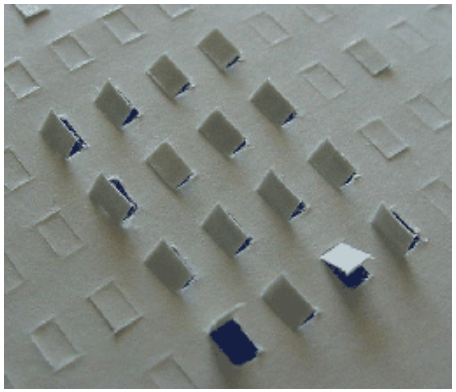


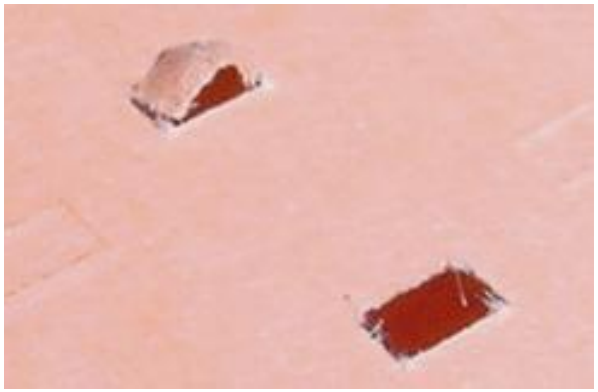
The infamous butterfly ballot from the 2000 Presidential election.



Bush v. Gore.

Hanging Chads and Voter Intent





Votomatic technology used in Florida was prone to paper jams. This led to hanging and dimpled chads, making it hard to determine voter intent.

http://www.cs.uiowa.edu/~jones/cards/chad.html
http://www.pushback.com/justice/votefraud/DimpledChadPictures.html

Next Big Step … Backward

Voting in the News: Take 3

The Voting Technology We Really Need? Paper

Software-independent backup systems are more important than ever.

LAWRENCE NORDEN | MAY 10, 2017 | TECHNOLOGY

Share Tweet ...

TEXT SIZE – +

In January, America's main intelligence agencies issued a report concluding that Russia interfered in the 2016 election, using a combination of cyber-intrusion, espionage, and propaganda. In addition to the details provided in this account, media outlets have since reported that several election databases were hacked before and after the election. While the Department of Homeland Security found no evidence any of these efforts manipulated vote tallies, the assaults have left many Americans asking: Just how safe are voting machines from cyberattack?

The answer is not reassuring.

Russian hacking fuels return to paper ballots

By Jenni Bergal Oct 03, 2017

This article originally appeared in Stateline, an initiative of the Pew Charitable Trusts.

After the "hanging chad" fiasco during the 2000 presidential recount, many states and counties switched to electronic-only voting machines to modernize their systems.

Now, amid security concerns over Russian hackers targeting state voting systems in last year's election, there's a renewed focus on shifting to paper ballots.

In Virginia, election officials decided last month to stop using paperless touch-screen machines, in an effort to safeguard against unauthorized access to the equipment and improve the security of the state's voting system.

In Georgia, which uses electronic voting machines with no paper record, legislators are discussing getting rid of their aging equipment and using

MORE INFO

Virginia considers decertifying touchscreen voting machines
The state could accelerate the move to systems that provide a paper trail for election audits.
Read more.

https://www.theatlantic.com/technology/archive/2017/05/the-voting-technology-we-really-need-paper/524820/
https://gcn.com/articles/2017/10/03/return-paper-ballots.aspx

Voting in the News: Take 3

Virginia Politics

Paper ballots make a comeback in Virginia this fall

In face of hacks and attempted hacks, paper ballots make a comeback in Virginia this fall. (Paul J. Richards/Agence France-Presse Via Getty Images)

By Patricia Sullivan October 7

The Computer Scientist Who Prefers Paper

Barbara Simons believes there is only one safe voting technology.

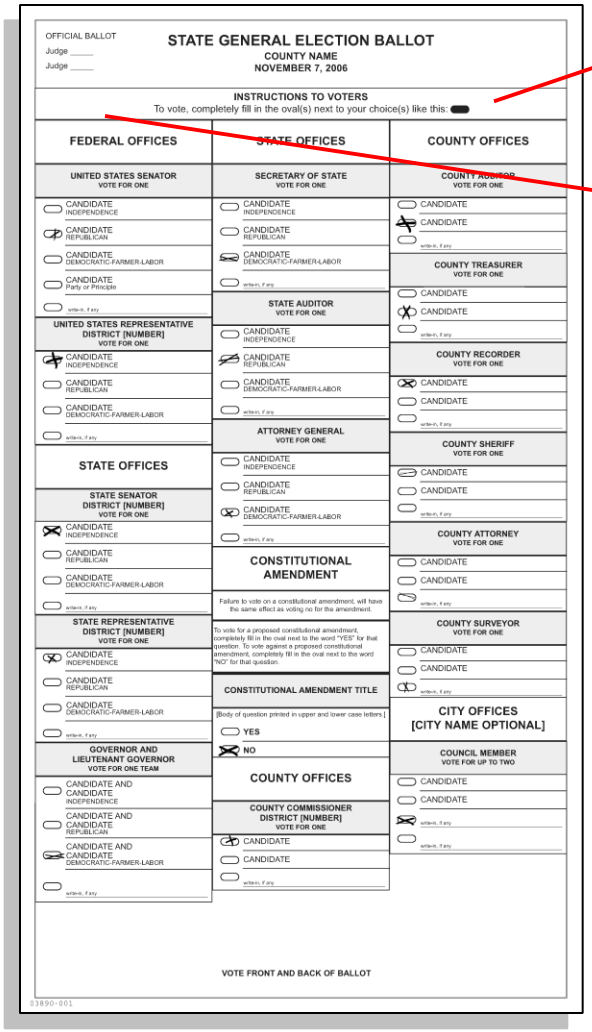
JILL LEOVY | DECEMBER 2017 ISSUE | TECHNOLOGY

https://www.washingtonpost.com/local/virginia-politics/paper-ballots-make-a-comeback-in-virginia-this-fall/2017/10/07/
https://www.theatlantic.com/magazine/archive/2017/12/guardian-of-the-vote/544155/

A Simple Yet Vexing Case Study: Counting Votes Recorded on Paper

Topic of current interest where the legal need to respect voter intent transforms a seemingly trivial pattern recognition problem into much more complex task.

Counting Votes Not So Easy


OFFICIAL BALLOT
Judge _____
Judge _____

STATE GENERAL ELECTION BALLOT
COUNTY NAME
NOVEMBER 7, 2006

INSTRUCTIONS TO VOTERS
To vote, completely fill in the oval(s) next to your choice(s) like this: ●

FEDERAL OFFICES

UNITED STATES SENATOR — VOTE FOR ONE
CANDIDATE — INDEPENDENCE
CANDIDATE — REPUBLICAN
CANDIDATE — DEMOCRATIC-FARMER-LABOR
CANDIDATE — Party or Principle
write-in, if any

UNITED STATES REPRESENTATIVE DISTRICT [NUMBER] — VOTE FOR ONE
CANDIDATE — INDEPENDENCE
CANDIDATE — REPUBLICAN
CANDIDATE — DEMOCRATIC-FARMER-LABOR
write-in, if any

STATE OFFICES

STATE SENATOR DISTRICT [NUMBER] — VOTE FOR ONE
CANDIDATE — INDEPENDENCE
CANDIDATE — REPUBLICAN
CANDIDATE — DEMOCRATIC-FARMER-LABOR
write-in, if any

STATE REPRESENTATIVE DISTRICT [NUMBER] — VOTE FOR ONE
CANDIDATE — INDEPENDENCE
CANDIDATE — REPUBLICAN
CANDIDATE — DEMOCRATIC-FARMER-LABOR
write-in, if any

GOVERNOR AND LIEUTENANT GOVERNOR — VOTE FOR ONE TEAM
CANDIDATE AND CANDIDATE — INDEPENDENCE
CANDIDATE AND CANDIDATE — REPUBLICAN
CANDIDATE AND CANDIDATE — DEMOCRATIC-FARMER-LABOR
write-in, if any

STATE OFFICES

SECRETARY OF STATE — VOTE FOR ONE
CANDIDATE — INDEPENDENCE
CANDIDATE — REPUBLICAN
CANDIDATE — DEMOCRATIC-FARMER-LABOR
write-in, if any

STATE AUDITOR — VOTE FOR ONE
CANDIDATE — INDEPENDENCE
CANDIDATE — REPUBLICAN
CANDIDATE — DEMOCRATIC-FARMER-LABOR
write-in, if any

ATTORNEY GENERAL — VOTE FOR ONE
CANDIDATE — INDEPENDENCE
CANDIDATE — REPUBLICAN
CANDIDATE — DEMOCRATIC-FARMER-LABOR
write-in, if any

CONSTITUTIONAL AMENDMENT

Failure to vote on a constitutional amendment, will have the same effect as voting no for the amendment.

To vote for a proposed constitutional amendment, completely fill in the oval next to the word "YES" for that question. To vote against a proposed constitutional amendment, completely fill in the oval next to the word "NO" for that question.

CONSTITUTIONAL AMENDMENT TITLE

[Body of question printed in upper and lower case letters.]

YES
NO

COUNTY OFFICES

COUNTY COMMISSIONER DISTRICT [NUMBER] — VOTE FOR ONE
CANDIDATE
CANDIDATE
write-in, if any

COUNTY OFFICES

COUNTY AUDITOR — VOTE FOR ONE
CANDIDATE
CANDIDATE
write-in, if any

COUNTY TREASURER — VOTE FOR ONE
CANDIDATE
CANDIDATE
write-in, if any

COUNTY RECORDER — VOTE FOR ONE
CANDIDATE
CANDIDATE
write-in, if any

COUNTY SHERIFF — VOTE FOR ONE
CANDIDATE
CANDIDATE
write-in, if any

COUNTY ATTORNEY — VOTE FOR ONE
CANDIDATE
CANDIDATE
write-in, if any

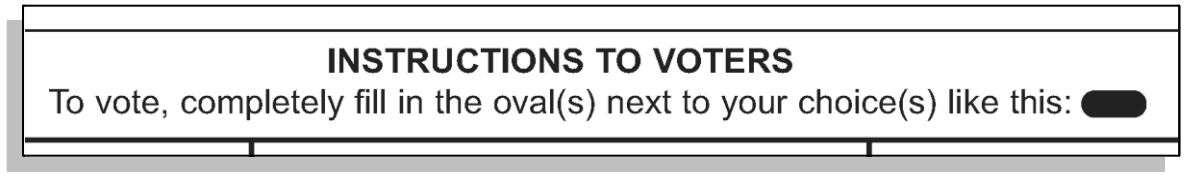
COUNTY SURVEYOR — VOTE FOR ONE
CANDIDATE
CANDIDATE
write-in, if any

CITY OFFICES [CITY NAME OPTIONAL]

COUNCIL MEMBER — VOTE FOR UP TO TWO
CANDIDATE
CANDIDATE
write-in, if any
write-in, if any

VOTE FRONT AND BACK OF BALLOT

03890-001



Is this a legal vote?

- Courts would probably say so …
- … but op-scan readers might not count it.

Increasing demands that machine's interpretation match a human's.

Research Questions

Issues that arise from using paper ballots in elections:

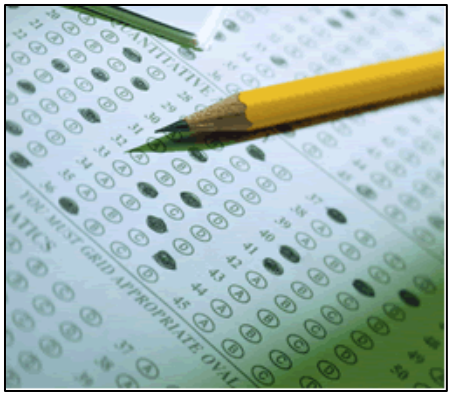
- Accurate interpretation of marginal markings.
- Human cost, error rate, and bias in performing manual recounts.
- Failure modes in ballot imaging (e.g., paper jams).
- Systematic errors due to ballot layout (one candidate may be disadvantaged over another based on physical location on page).

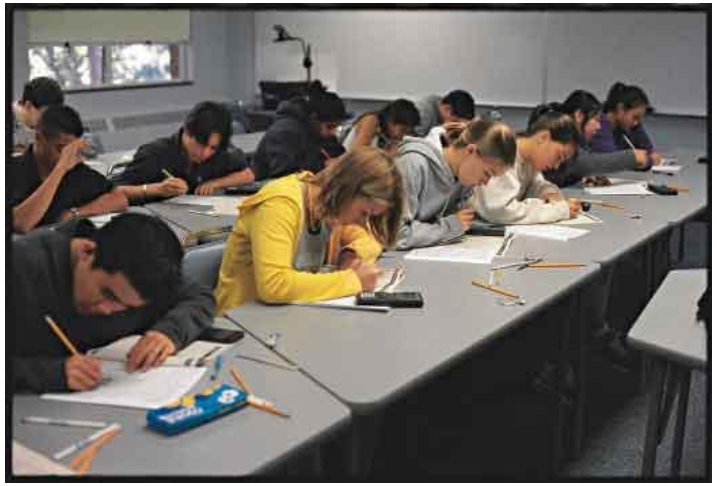
Also keep in mind:

- U.S. elections can be complex (10's to 100's of choices).
- Impact of "voter error" (e.g., improper markings, erasures).
- Potential for traditional ballot-box stuffing.
- Computer hackers attempting to manipulate the vote.

Why isn't this a solved problem?

Students have been taking standardized tests using op-scan answer sheets for decades …





- While accuracy rates are very high, problems do occur.
- Compared to voters, students are a much more homogeneous (and well-educated) population.
- Standardized testing is NOT anonymous. Students can (and do) complain when they receive a lower score than they expect.

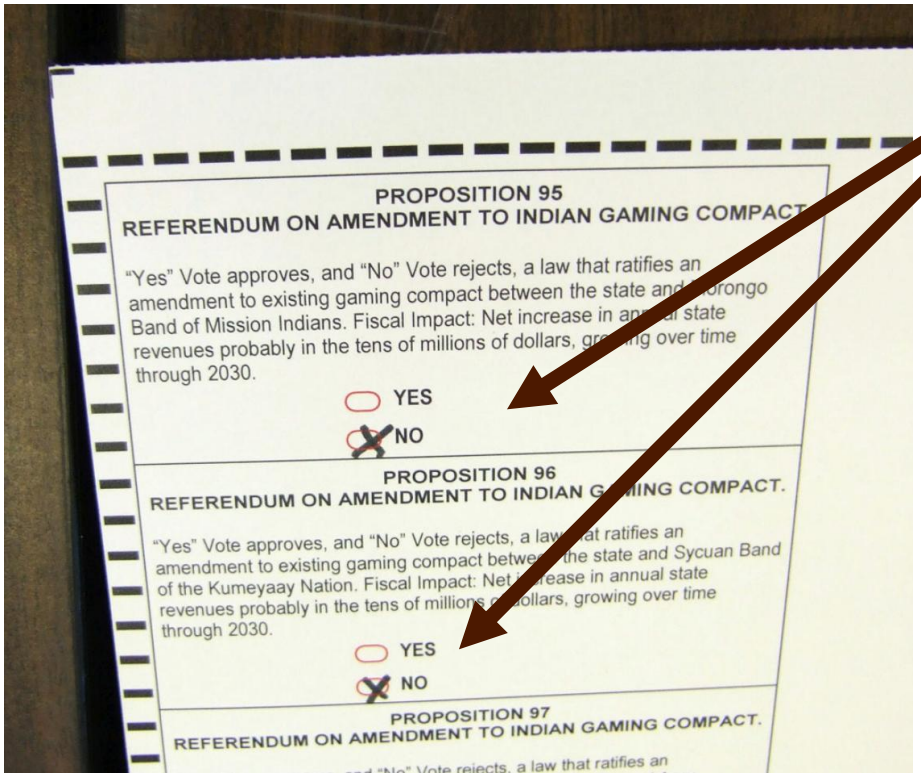
Connection to Forms Processing

Similarities to forms processing, but also some key differences:

- Much broader range of users (education level, literacy, etc.) than for traditional forms applications.
- Ballots must preserve a voter's anonymity.
- Demand to count votes and report results quickly.
- Elections are held infrequently, so voting equipment sits unused for long periods in storage.
- Poll workers often lack technical expertise.
- Maintaining chain-of-custody is a critical security requirement.
- No *financial* interest in making sure votes are counted accurately, but there is tremendous *public* interest.

Counting Votes Not So Easy

Real ballot from an election in California:

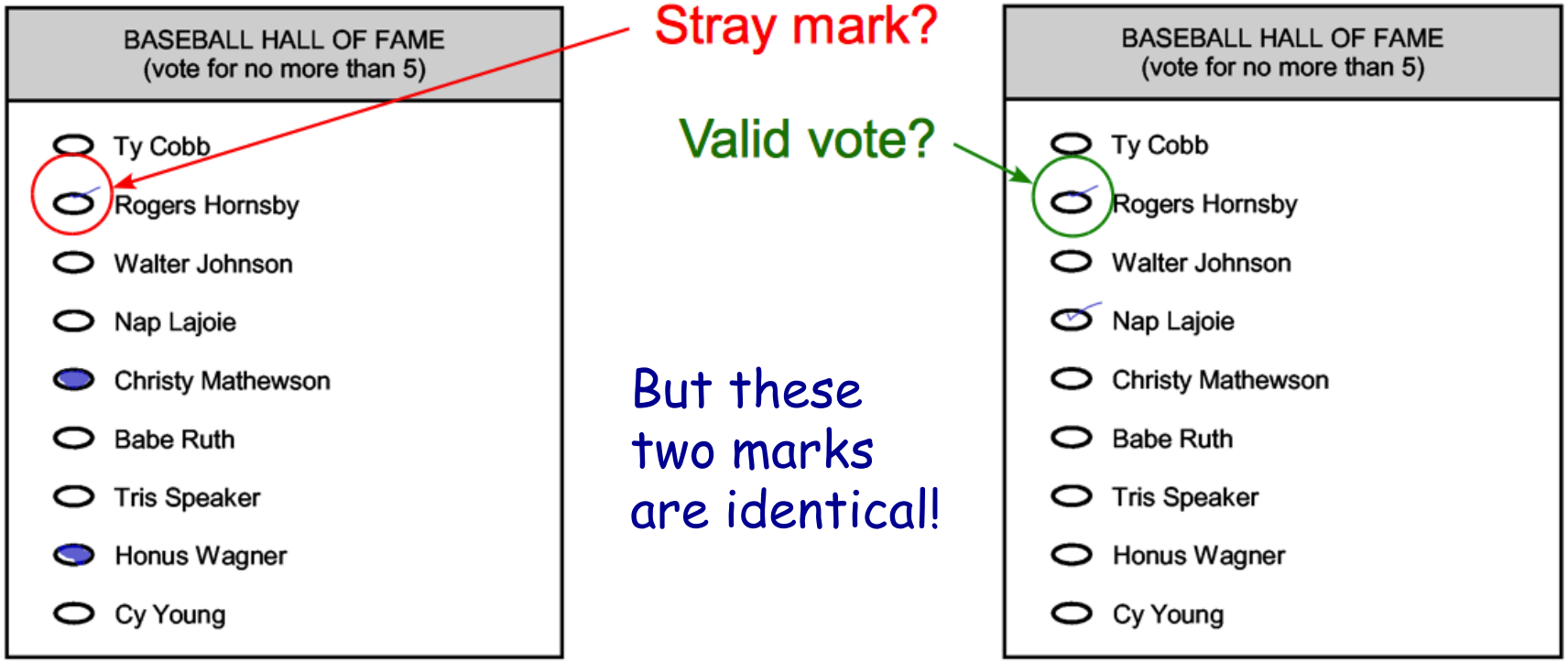


One of these votes was counted correctly by the op-scan equipment, the other was not.

Note: this does not mean voting on paper ballots is bad, just (1) manual audits should be mandatory, and (2) more research is needed.

"Improving California's 1% Manual Tally Procedure," Joseph Lorenzo Hall, UC Berkeley School of Information, EVT Workshop 2008.

Whole-Ballot Recognition



Can we capture voter intent via style-based techniques?

Style-Based Mark Recognition

Traditional Forms Processing

Style-Base Ballot Mark Recognition

Can the system interpret the voter's intent? (If a human judge would interpret a marking as an intended vote, then the voting machine should do the same.)

Can fail to record some votes simply because they do not satisfy an arbitrary criterion (e.g., a fixed threshold on the number of black pixels).

Assume a voter is self-consistent when marking his/her ballot.

Create a style-based classifier from a set of style-specialized classifiers to improve recognition accuracy.

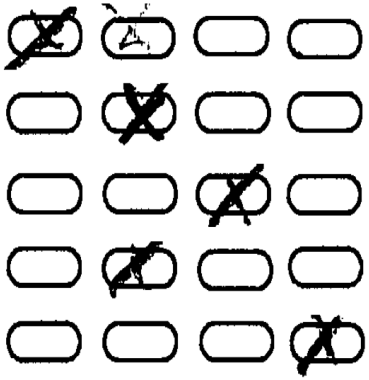
Limiting

Promising

"Style-Based Ballot Mark Recognition," P. Xiu, D. Lopresti, H. Baird, G. Nagy, and E. Barney Smith, *Proceedings of the Tenth International Conference on Document Analysis and Recognition*, July 2009, Barcelona, Spain, pp. 216-220.

Challenging Cases

A Style-Consistent Field

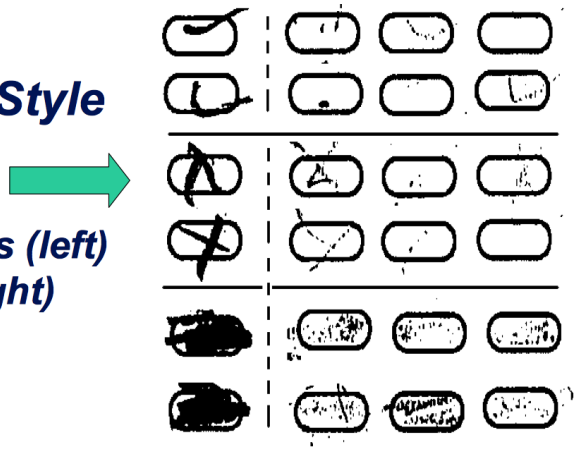


Human Interpretation

VNNN,NVNN,NNVN,NVNN,NNNV

(V for Vote, N for Non-vote)

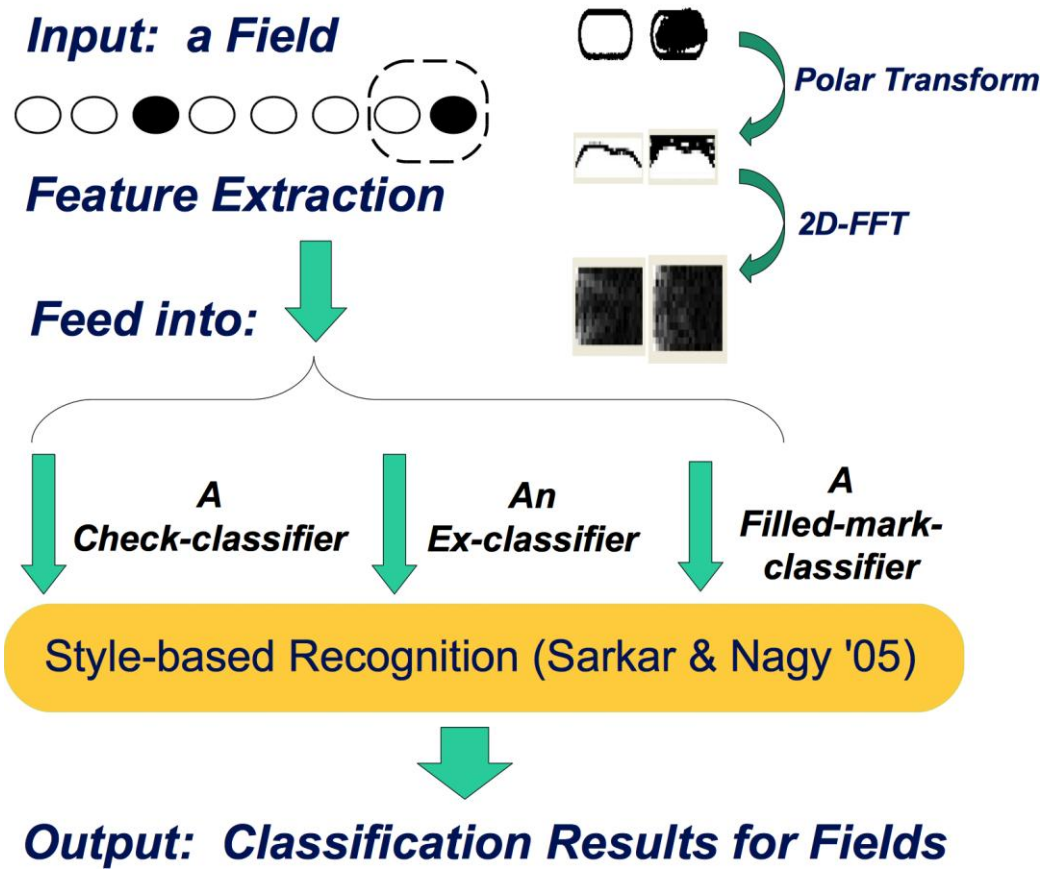
Variations in Marking Style



Check, ex, and filled marks (left) vs. noisy non-votes (right)

"Style-Based Ballot Mark Recognition," P. Xiu, D. Lopresti, H. Baird, G. Nagy, and E. Barney Smith, *Proceedings of the Tenth International Conference on Document Analysis and Recognition*, July 2009, Barcelona, Spain, pp. 216-220.

System Design



"Style-Based Ballot Mark Recognition," P. Xiu, D. Lopresti, H. Baird, G. Nagy, and E. Barney Smith, *Proceedings of the Tenth International Conference on Document Analysis and Recognition*, July 2009, Barcelona, Spain, pp. 216-220.

Style-Based Performance

Table 3. Target-level error rates (top) and field-level error rates (bottom).

	Classifier					
Sample Set	Check	Ex	Filled	Blend	Separate	Style-based
Check	2.36%	7.46%	25.00%	1.97%	4.35%	2.78%
Ex	0.40%	0.34%	16.16%	0.40%	0.40%	0.35%
Filled	2.75%	2.38%	1.10%	2.75%	2.50%	1.09%
Average	1.84%	3.39%	14.09%	1.70%	2.42%	1.41%

	Classifier					
Sample Set	Check	Ex	Filled	Blend	Separate	Style-based
Check	38.30%	83.25%	100.00%	33.43%	61.08%	42.85%
Ex	7.77%	6.70%	99.30%	7.77%	7.77%	6.75%
Filled	53.18%	46.07%	20.75%	53.18%	48.55%	20.63%
Average	33.08%	45.34%	73.35%	31.46%	39.13%	23.41%

"Style-Based Ballot Mark Recognition," P. Xiu, D. Lopresti, H. Baird, G. Nagy, and E. Barney Smith, *Proceedings of the Tenth International Conference on Document Analysis and Recognition*, July 2009, Barcelona, Spain, pp. 216-220.

A Bit of Good Luck

But what we'd like to have is ballots from a real election. Even better, the ballots would be from an important election where the voter markings present serious pattern recognition challenges.

Extremely close U.S. Senate race in State of Minnesota: six days after election, unofficial results showed Republican Norm Coleman leading Democratic challenger Al Franken by 206 votes out of nearly 3 million cast, a difference of less than 0.01%.

"Document Analysis Issues in Reading Optical Scan Ballots," D. Lopresti, G. Nagy, and E. Barney Smith, *Proceedings of the Ninth IAPR International Workshop on Document Analysis Systems*, June 2010, Boston, MA, pp. 105-112.

A Bit of Good Luck

- Minnesota uses op-scan ballots.
- Closeness of election triggers a manual recount.
- Both sides are allowed to challenge validity of "questionable" ballots.
- Openness laws make challenged ballots a matter of public record.
- Ballot images made available on MN public radio website.
- PDF files contain 300 dpi TIF images!

http://minnesota.publicradio.org/features/2008/11/19_challenged_ballots/

Minnesota Statutes

Remember that the guiding principle is *voter intent*. Here are a few key points to keep in mind when interpreting ballot markings:

- "A ballot shall not be rejected for a technical error that does not make it impossible to determine the voter's intent."
- "If a mark (X) is made out of its proper place, but so near a name or space as to indicate clearly the voter's intent, the vote shall be counted."
- "Misspelling or abbreviations of the names of write-in candidates shall be disregarded if the individual for whom the vote was intended can be clearly ascertained from the ballot."

https://www.revisor.mn.gov/statutes/?id=204C.22

Minnesota Statutes

… and …

- "If a voter uniformly uses a mark other than (X) which clearly indicates an intent to mark a name or to mark yes or no on a question, and the voter does not use (X) anywhere else on the ballot, a vote shall be counted for each candidate or response to a question marked.
- If a voter uses two or more distinct marks, such as (X) and some other mark, a vote shall be counted for each candidate or response to a question marked, unless the ballot is marked by distinguishing characteristics that make the entire ballot defective …"

https://www.revisor.mn.gov/statutes/?id=204C.22

Minnesota Statutes

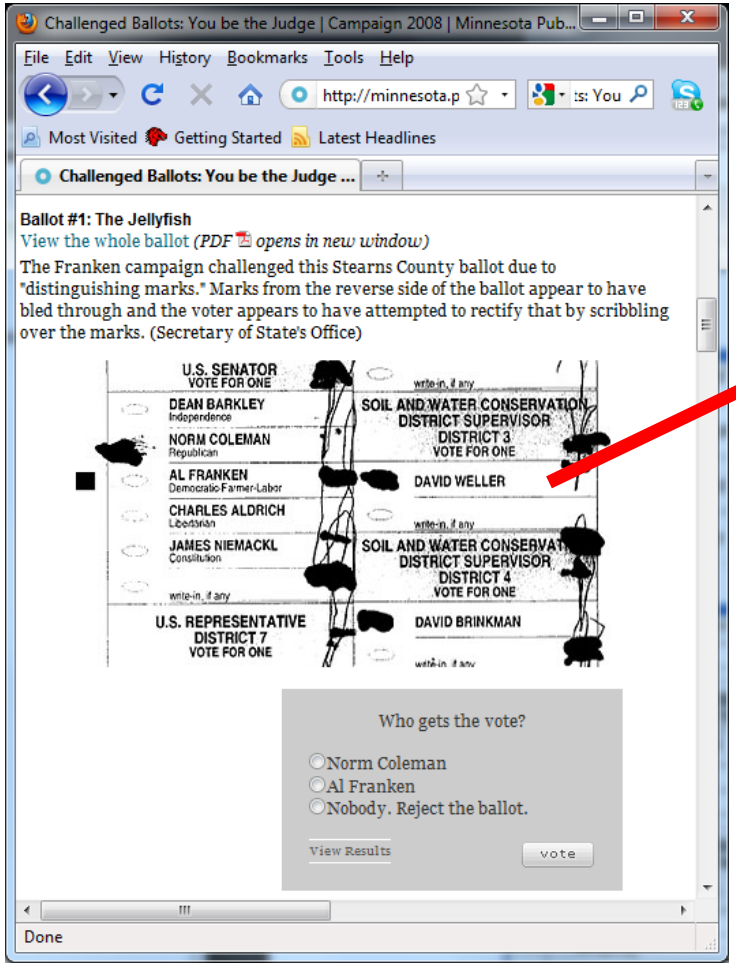
… and …

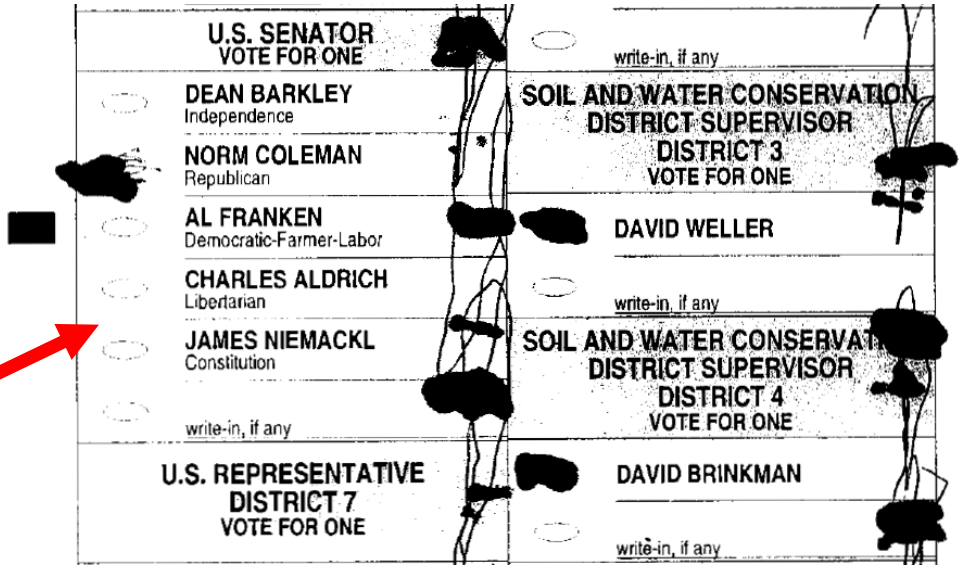
- "If the names of two candidates have been marked, and an attempt has been made to erase or obliterate one of the marks, a vote shall be counted for the remaining marked candidate."
- "A ballot shall not be rejected merely because it is slightly soiled or defaced."
- "If a ballot is marked by distinguishing characteristics in a manner making it evident that the voter intended to identify the ballot, the entire ballot is defective."

Goal here is to prevent coercion or vote selling.

https://www.revisor.mn.gov/statutes/?id=204C.22

Challenge: you be the judge

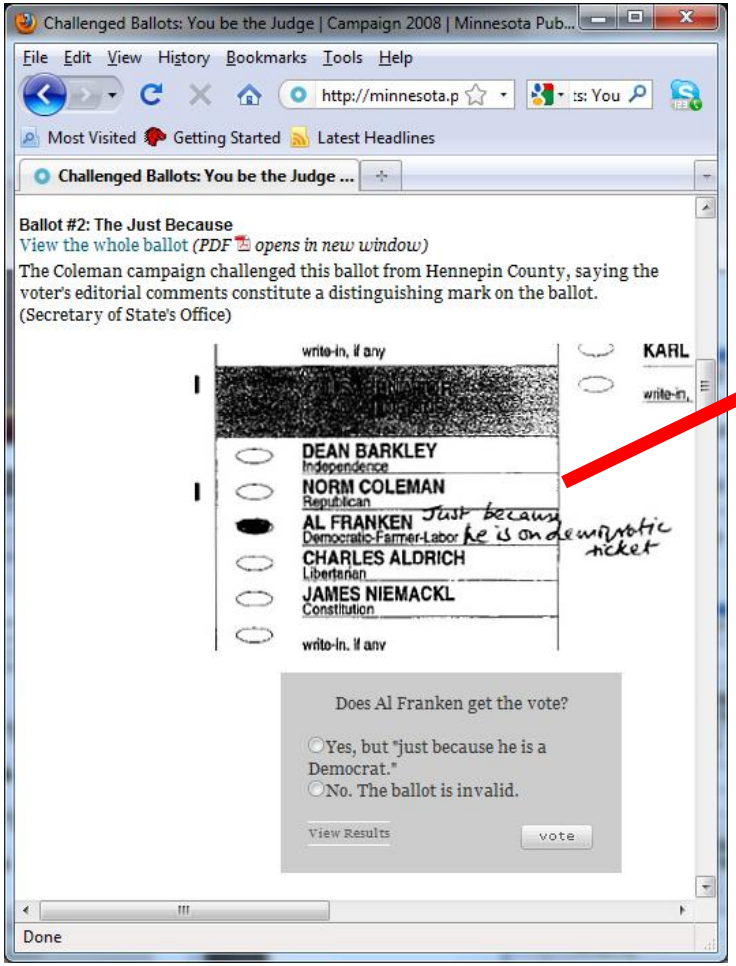


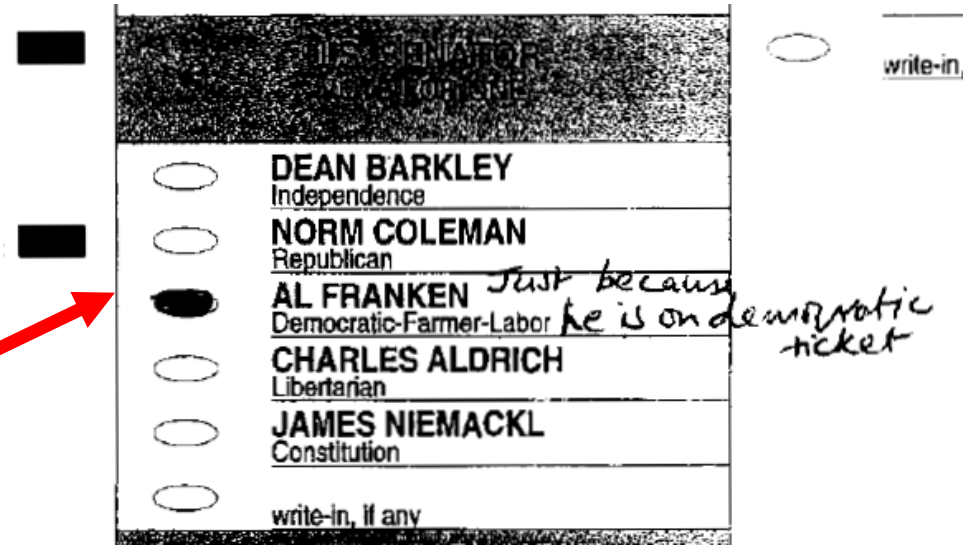


Who gets vote? Public opinion:

- Norm Coleman: 63% (7,626 votes)
- Al Franken: 4% (474 votes)
- Nobody: 33% (4,050 votes)

Challenge: you be the judge

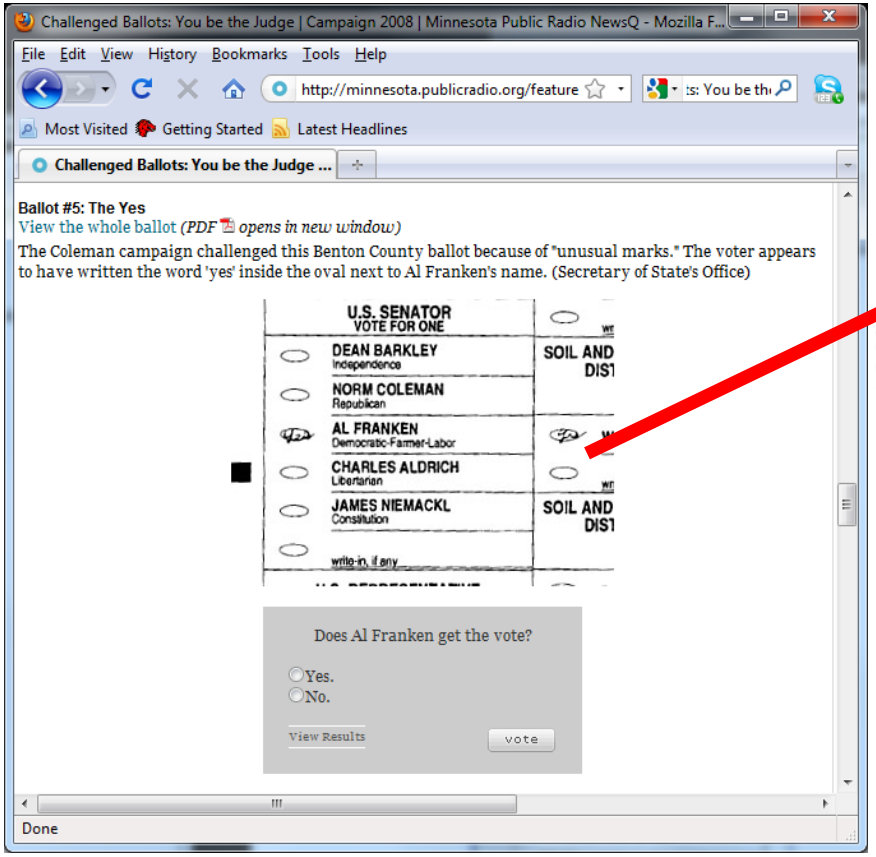


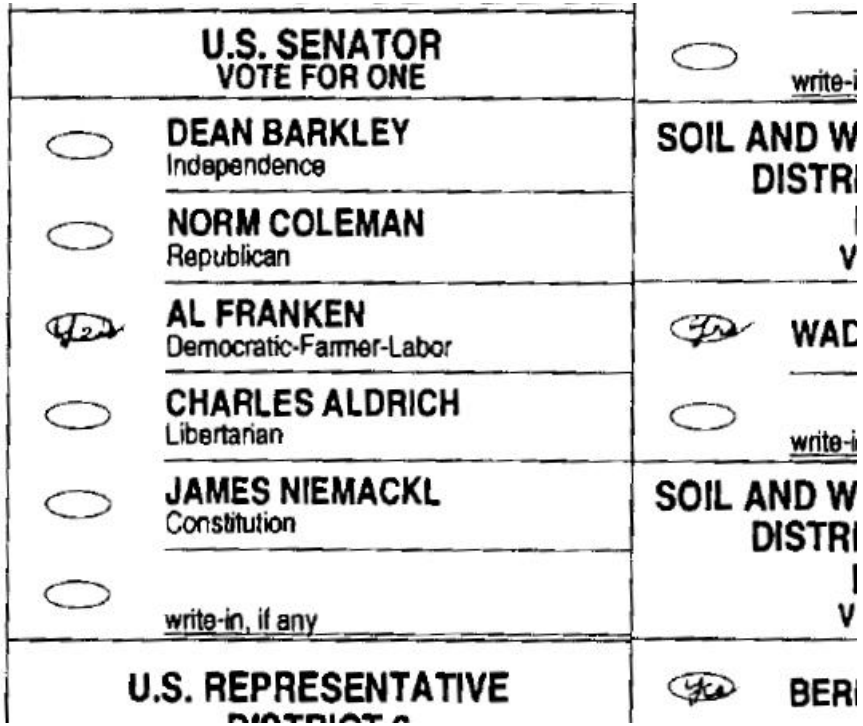


Vote for Franken? Public opinion:

- Yes: 92% (11,069 votes)
- No: 8% (1,012 votes)

Challenge: you be the judge

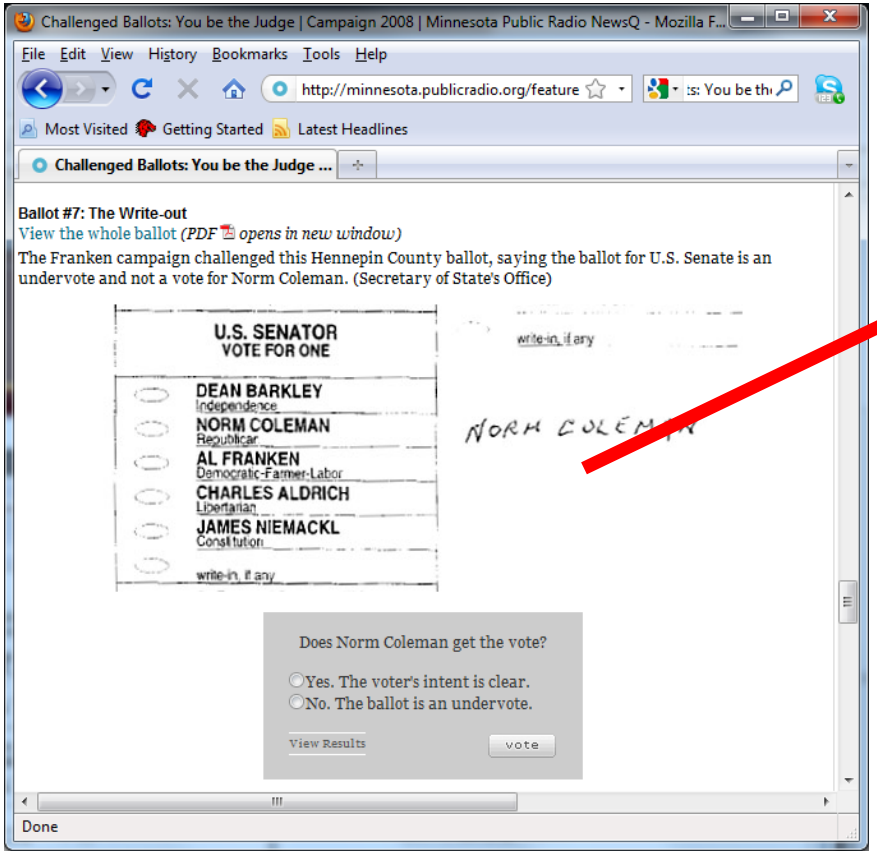


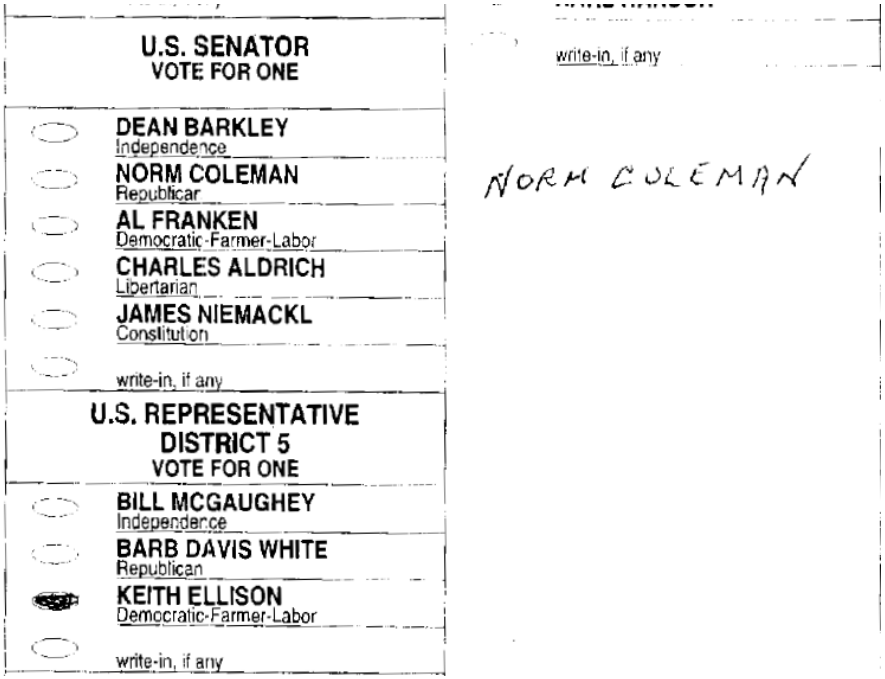


Vote for Franken? Public opinion:

- Yes: 96% (11,250 votes)
- No: 4% (452 votes)

Challenge: you be the judge





Vote for Coleman? Public opinion:

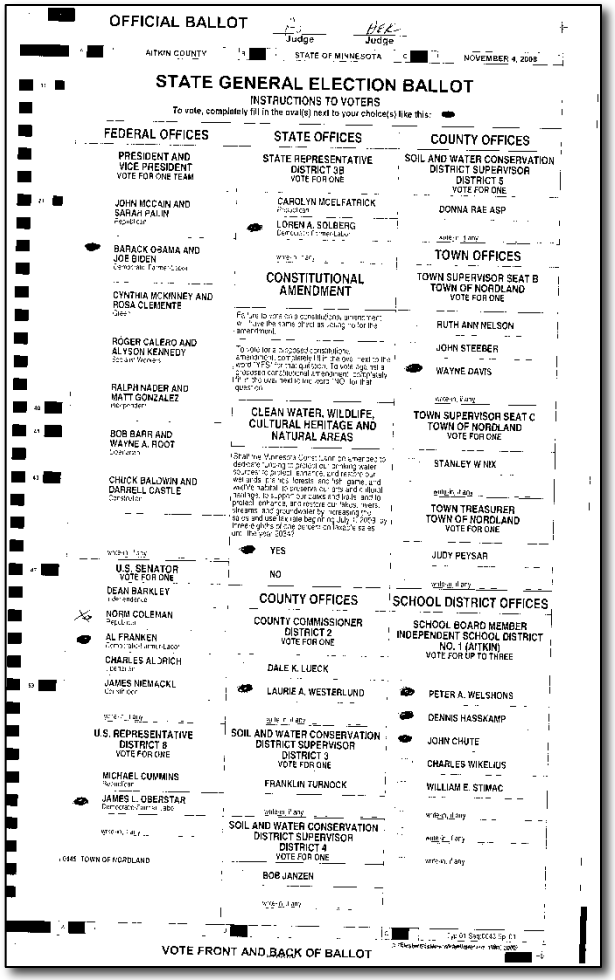
- Yes: 54% (6,080 votes)
- No: 46% (5,203 votes)

MN Challenged Ballot Collection

How the ballot collection was generated and harvested:

- Ballots photocopied and originals stored in a secure location.
- Copies scanned to PDF using auto-feeder flatbed scanner.
- Ballot was two-sided, with both sides scanned simultaneously.
- I wrote a simple web "crawler" that automatically downloaded all the files and extracted TIF images from PDF.
- A total of 6,737 ballots in the set.
- Examination of the TIF suggests that ballots were scanned at 300 dpi bitonal, and that lossy compression was never used.
- Hence, they form an ideal dataset for research purposes.

Minnesota Ballot Front and Back



OFFICIAL BALLOT

Judge

Judge

AITKIN COUNTY | STATE OF MINNESOTA | NOVEMBER 4, 2008

STATE GENERAL ELECTION BALLOT

INSTRUCTIONS TO VOTERS

To vote, completely fill in the oval(s) next to your choice(s) like this:

FEDERAL OFFICES

PRESIDENT AND VICE PRESIDENT
VOTE FOR ONE TEAM

JOHN MCCAIN AND SARAH PALIN
Republican

BARACK OBAMA AND JOE BIDEN
Democratic Farmer Labor

CYNTHIA MCKINNEY AND ROSA CLEMENTE
Green

ROGER CALERO AND ALYSON KENNEDY
Socialist Workers

RALPH NADER AND MATT GONZALEZ
Independent

BOB BARR AND WAYNE A. ROOT
Libertarian

CHUCK BALDWIN AND DARRELL CASTLE
Constitution

U.S. SENATOR
VOTE FOR ONE

DEAN BARKLEY
Independence

NORM COLEMAN
Republican

AL FRANKEN
Democratic-Farmer-Labor

CHARLES ALDRICH
Libertarian

JAMES NIEMACKI
Constitution

U.S. REPRESENTATIVE DISTRICT 8
VOTE FOR ONE

MICHAEL CUMMINS
Republican

JAMES L. OBERSTAR
Democratic-Farmer-Labor

write-in

0045 TOWN OF NORDLAND

STATE OFFICES

STATE REPRESENTATIVE DISTRICT 9B
VOTE FOR ONE

CAROLYN MCELFATRICK
Republican

LOREN A. SOLBERG
Democratic-Farmer-Labor

write-in

CONSTITUTIONAL AMENDMENT

To vote for or against a constitutional amendment, completely fill in the oval next to your choice.

To vote for a proposed constitutional amendment, completely fill in the oval next to the word "YES." To vote against a proposed constitutional amendment, completely fill in the oval next to the word "NO." for that question.

CLEAN WATER, WILDLIFE, CULTURAL HERITAGE AND NATURAL AREAS

Shall the Minnesota Constitution be amended to dedicate funding to protect drinking water sources; to protect, enhance, and restore our wetlands, prairies, forests, and fish, game, and wildlife habitat; to preserve our parks and cultural heritage; to support our arts and history; and to protect, enhance, and restore our lakes, rivers, streams, and groundwater by increasing the sales and use tax rate beginning July 1, 2009, by three-eighths of one percent to take effect until the year 2034?

YES

NO

COUNTY OFFICES

COUNTY COMMISSIONER DISTRICT 2
VOTE FOR ONE

DALE K. LUECK

LAURIE A. WESTERLUND

SOIL AND WATER CONSERVATION DISTRICT SUPERVISOR DISTRICT 3
VOTE FOR ONE

FRANKLIN TURNOCK

SOIL AND WATER CONSERVATION DISTRICT SUPERVISOR DISTRICT 4
VOTE FOR ONE

BOB JANZEN

write-in

COUNTY OFFICES

SOIL AND WATER CONSERVATION DISTRICT SUPERVISOR DISTRICT 5
VOTE FOR ONE

DONNA RAE ASP

write-in

TOWN OFFICES

TOWN SUPERVISOR SEAT B TOWN OF NORDLAND
VOTE FOR ONE

RUTH ANN NELSON

JOHN STEEBER

WAYNE DAVIS

write-in

TOWN SUPERVISOR SEAT C TOWN OF NORDLAND
VOTE FOR ONE

STANLEY W NIX

write-in

TOWN TREASURER TOWN OF NORDLAND
VOTE FOR ONE

JUDY PEYSAR

write-in

SCHOOL DISTRICT OFFICES

SCHOOL BOARD MEMBER INDEPENDENT SCHOOL DISTRICT NO. 1 (AITKIN)
VOTE FOR UP TO THREE

PETER A. WELSHONS

DENNIS HASSKAMP

JOHN CHUTE

CHARLES WIKELIUS

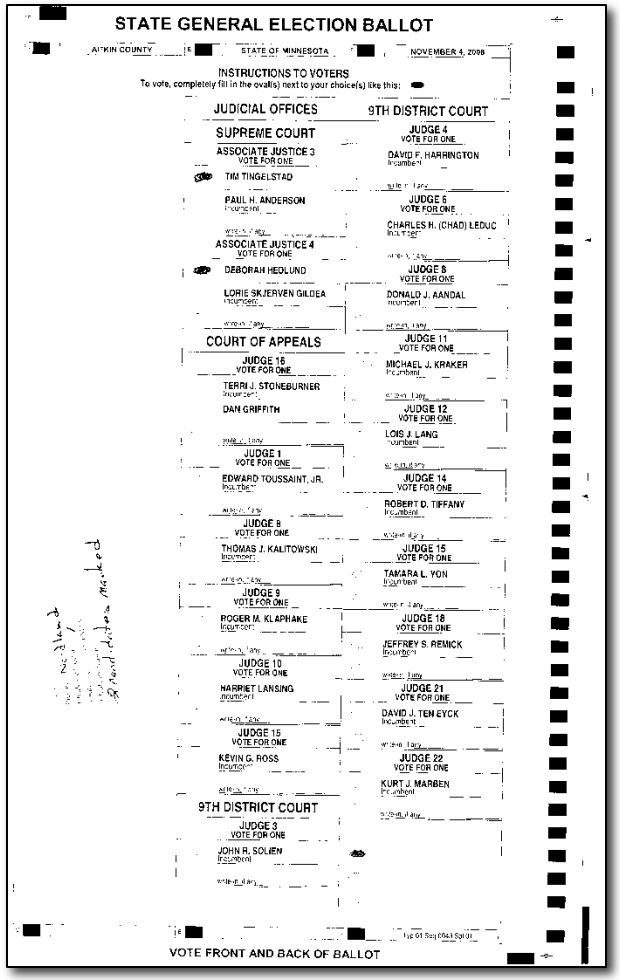
WILLIAM E. STIMAC

write-in

write-in

write-in

VOTE FRONT AND BACK OF BALLOT



STATE GENERAL ELECTION BALLOT

AITKIN COUNTY | STATE OF MINNESOTA | NOVEMBER 4, 2008

INSTRUCTIONS TO VOTERS

To vote, completely fill in the oval(s) next to your choice(s) like this:

JUDICIAL OFFICES

SUPREME COURT

ASSOCIATE JUSTICE 3
VOTE FOR ONE

TIM TINGELSTAD

PAUL H. ANDERSON
Incumbent

write-in

ASSOCIATE JUSTICE 4
VOTE FOR ONE

DEBORAH HEDLUND

LORIE SKJERVEN GILDEA
Incumbent

write-in

COURT OF APPEALS

JUDGE 16
VOTE FOR ONE

TERRI J. STONEBURNER
Incumbent

DAN GRIFFITH

write-in

JUDGE 1
VOTE FOR ONE

EDWARD TOUSSAINT, JR.
Incumbent

write-in

JUDGE 8
VOTE FOR ONE

THOMAS J. KALITOWSKI
Incumbent

write-in

JUDGE 9
VOTE FOR ONE

ROGER M. KLAPHAKE
Incumbent

write-in

JUDGE 10
VOTE FOR ONE

HARRIET LANSING
Incumbent

write-in

JUDGE 15
VOTE FOR ONE

KEVIN G. ROSS
Incumbent

write-in

9TH DISTRICT COURT

JUDGE 3
VOTE FOR ONE

JOHN R. SOLIEN
Incumbent

write-in

9TH DISTRICT COURT

JUDGE 4
VOTE FOR ONE

DAVID F. HARRINGTON
Incumbent

write-in

JUDGE 6
VOTE FOR ONE

CHARLES H. (CHAD) LEDUC
Incumbent

write-in

JUDGE 8
VOTE FOR ONE

DONALD J. AANDAL
Incumbent

write-in

JUDGE 11
VOTE FOR ONE

MICHAEL J. KRAKER
Incumbent

write-in

JUDGE 12
VOTE FOR ONE

LOIS J. LANG
Incumbent

write-in

JUDGE 14
VOTE FOR ONE

ROBERT D. TIFFANY
Incumbent

write-in

JUDGE 15
VOTE FOR ONE

TAMARA L. YON
Incumbent

write-in

JUDGE 18
VOTE FOR ONE

JEFFREY S. REMICK
Incumbent

write-in

JUDGE 21
VOTE FOR ONE

DAVID J. TEN EYCK
Incumbent

write-in

JUDGE 22
VOTE FOR ONE

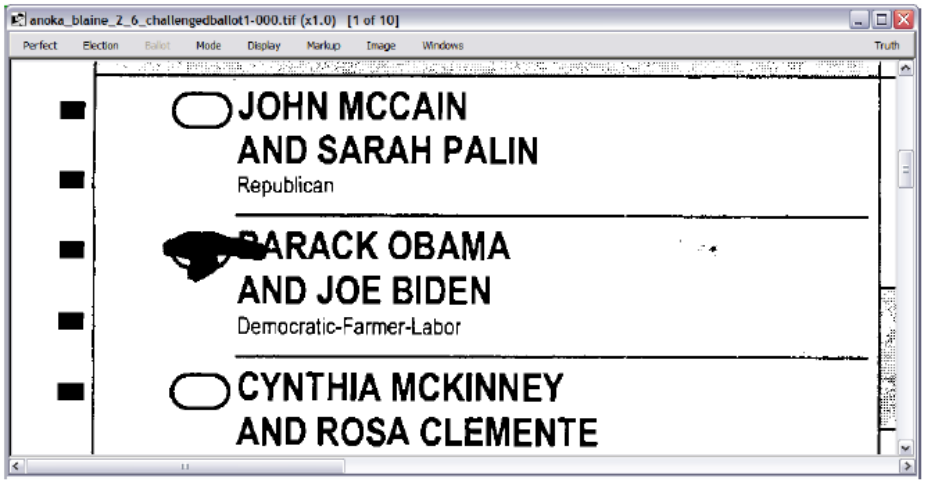
KURT J. MARBEN
Incumbent

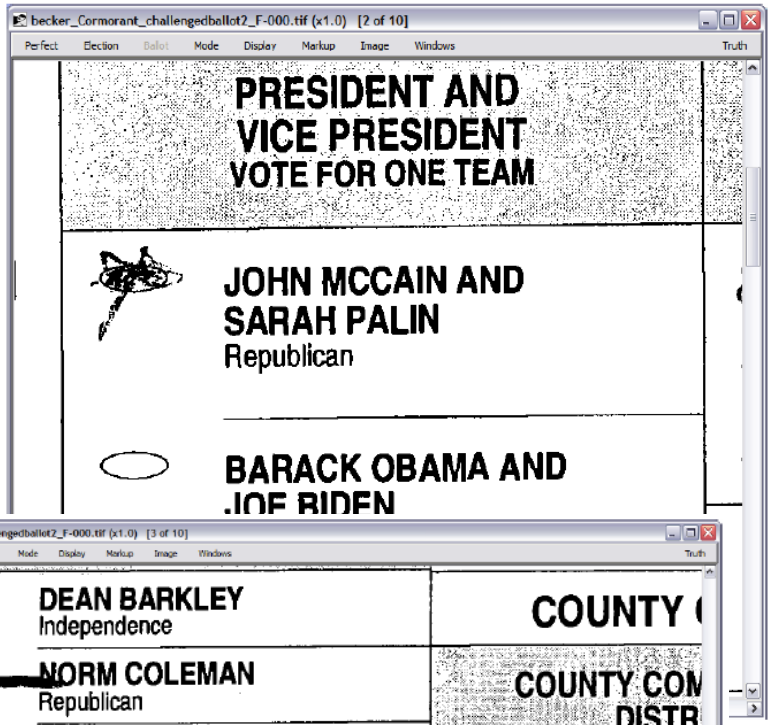
write-in

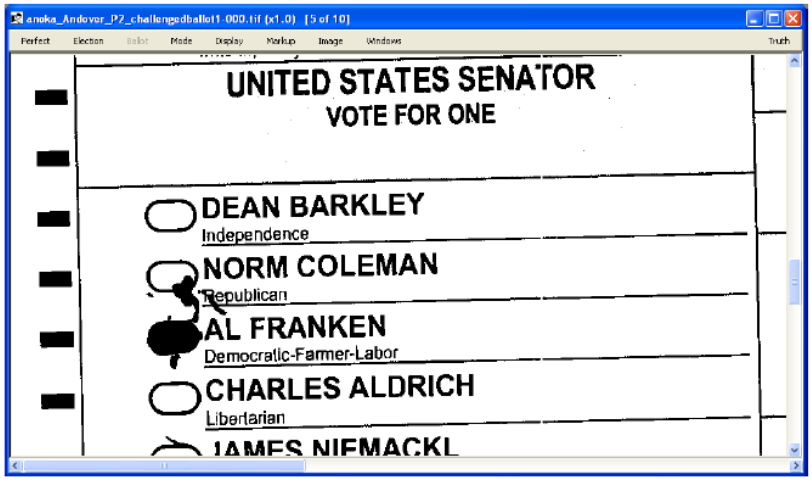
Northland Ground Pellets Marked

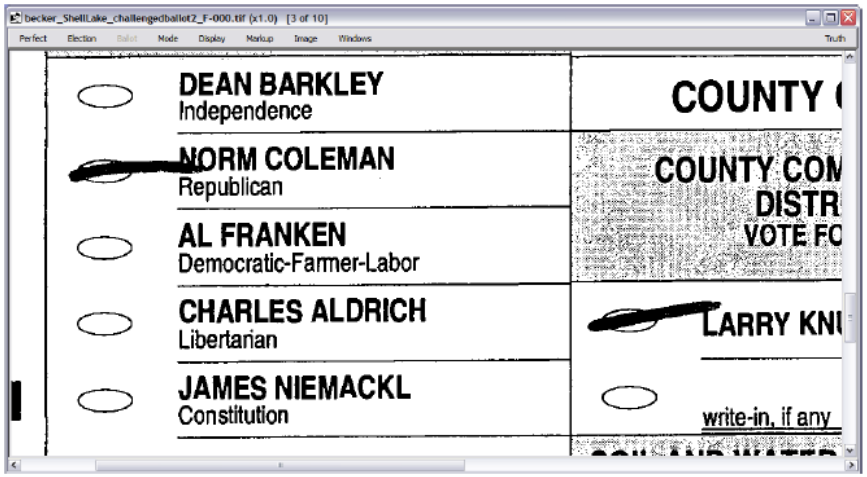
VOTE FRONT AND BACK OF BALLOT

Sloppy-But-Valid Marks

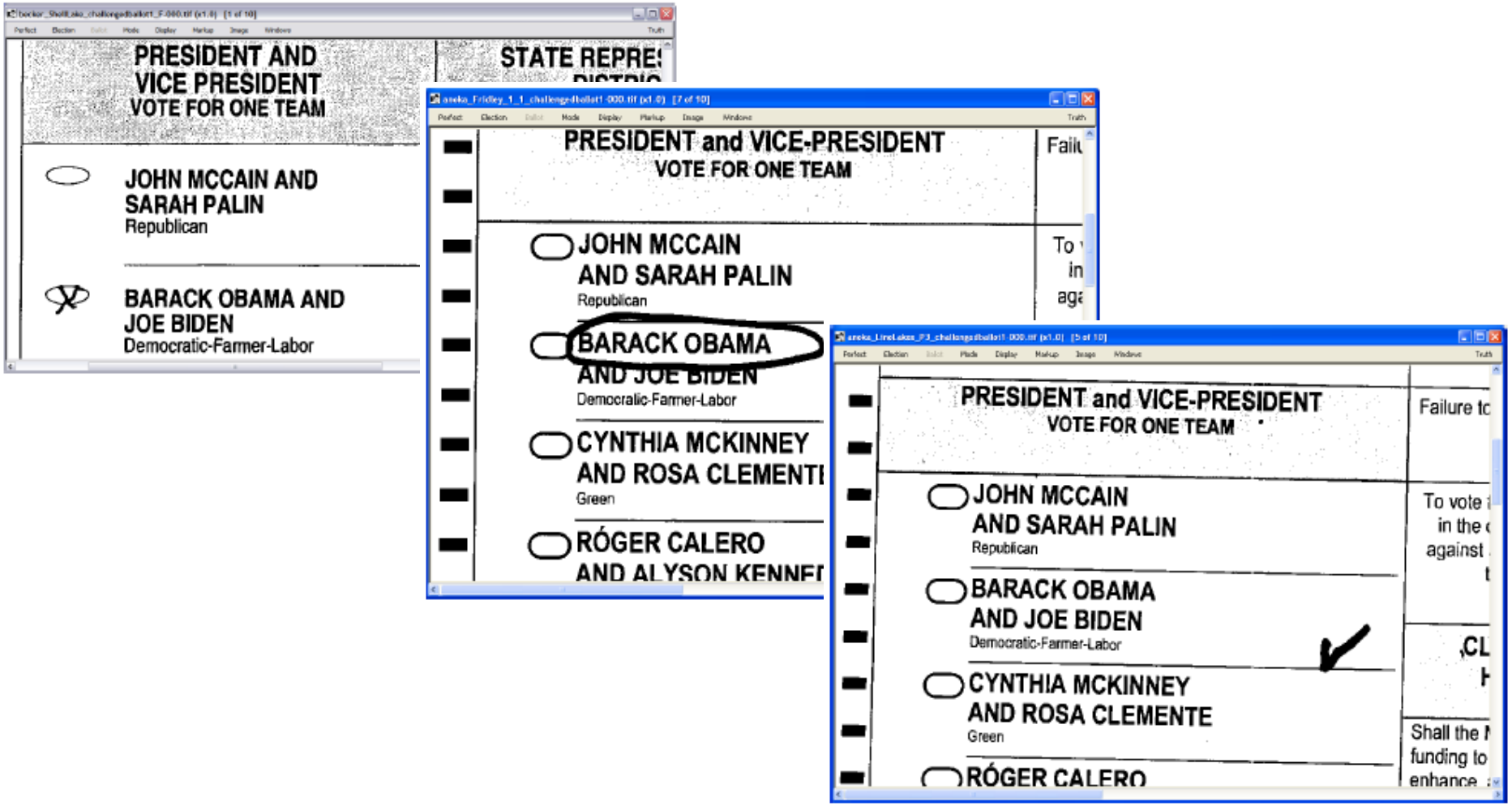




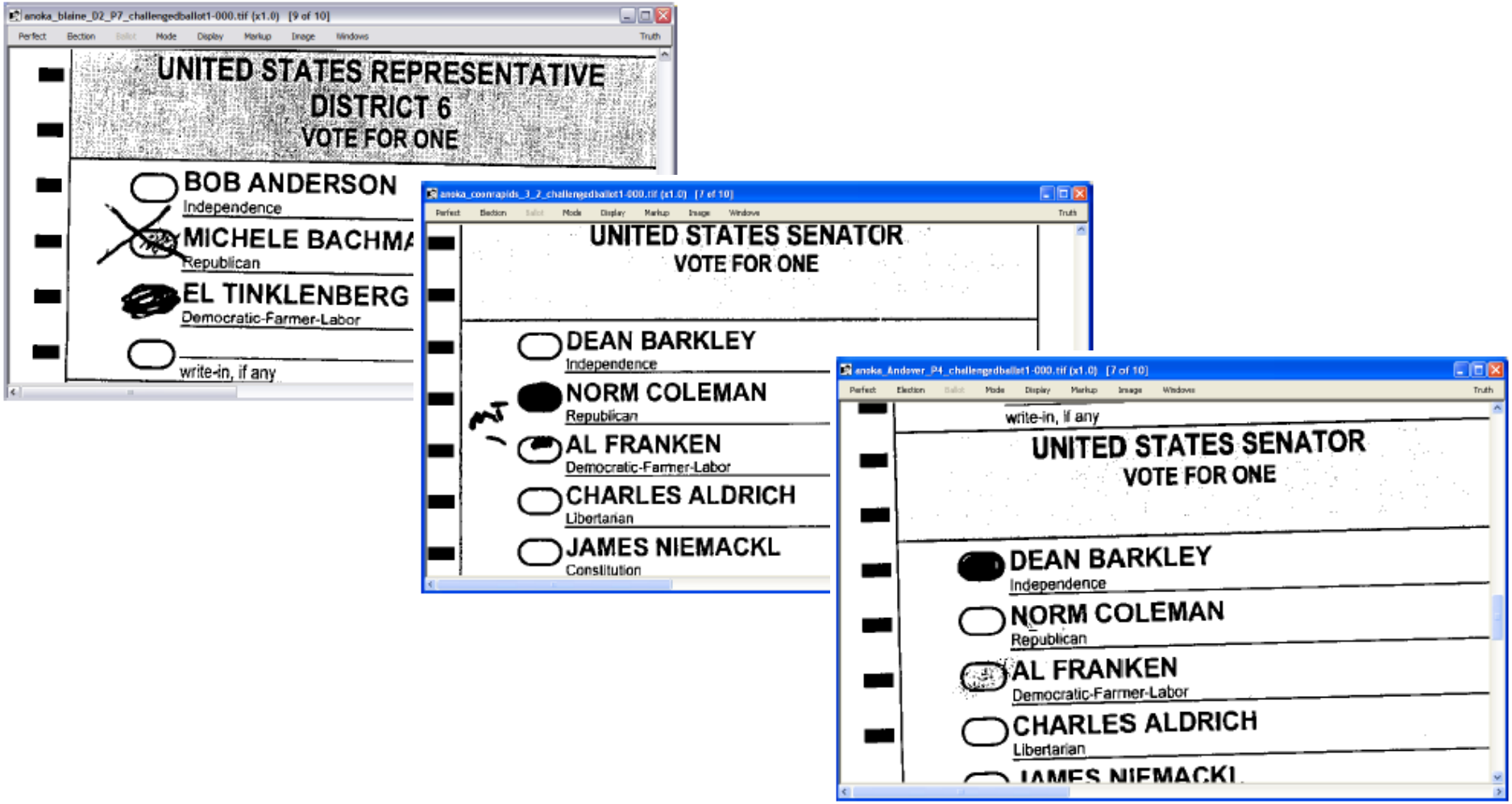




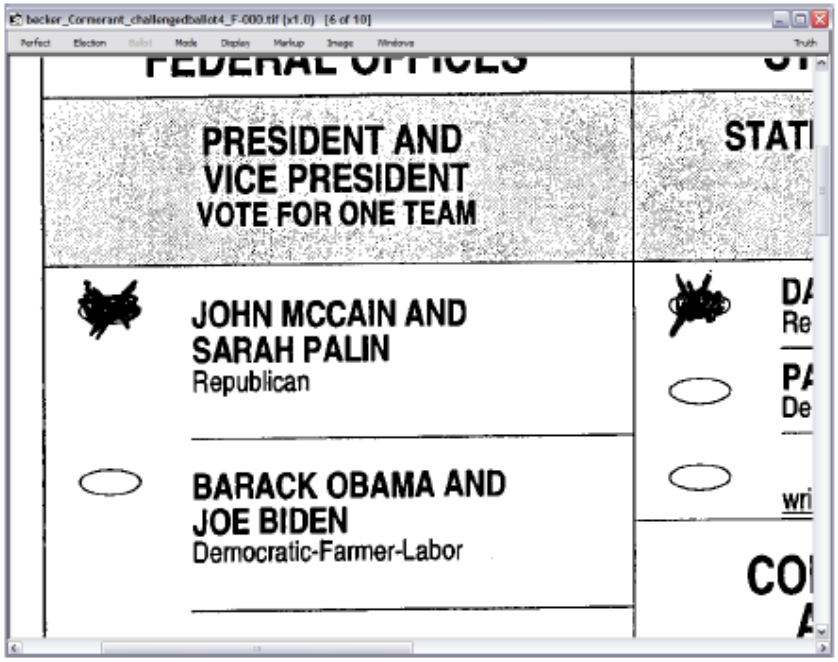
Non-Conforming Marking Styles

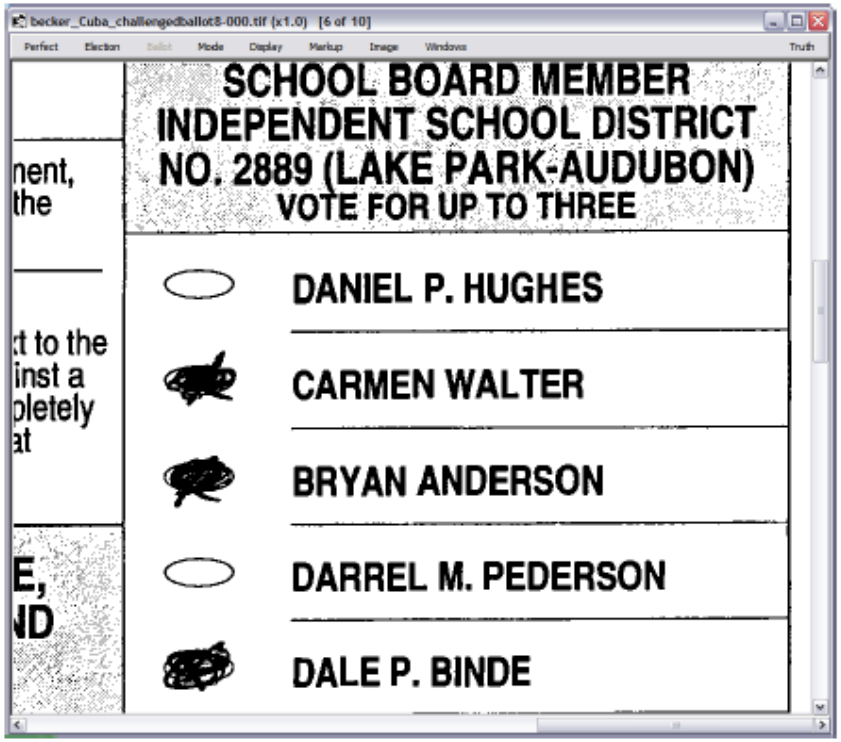


Attempts to Cancel a Vote

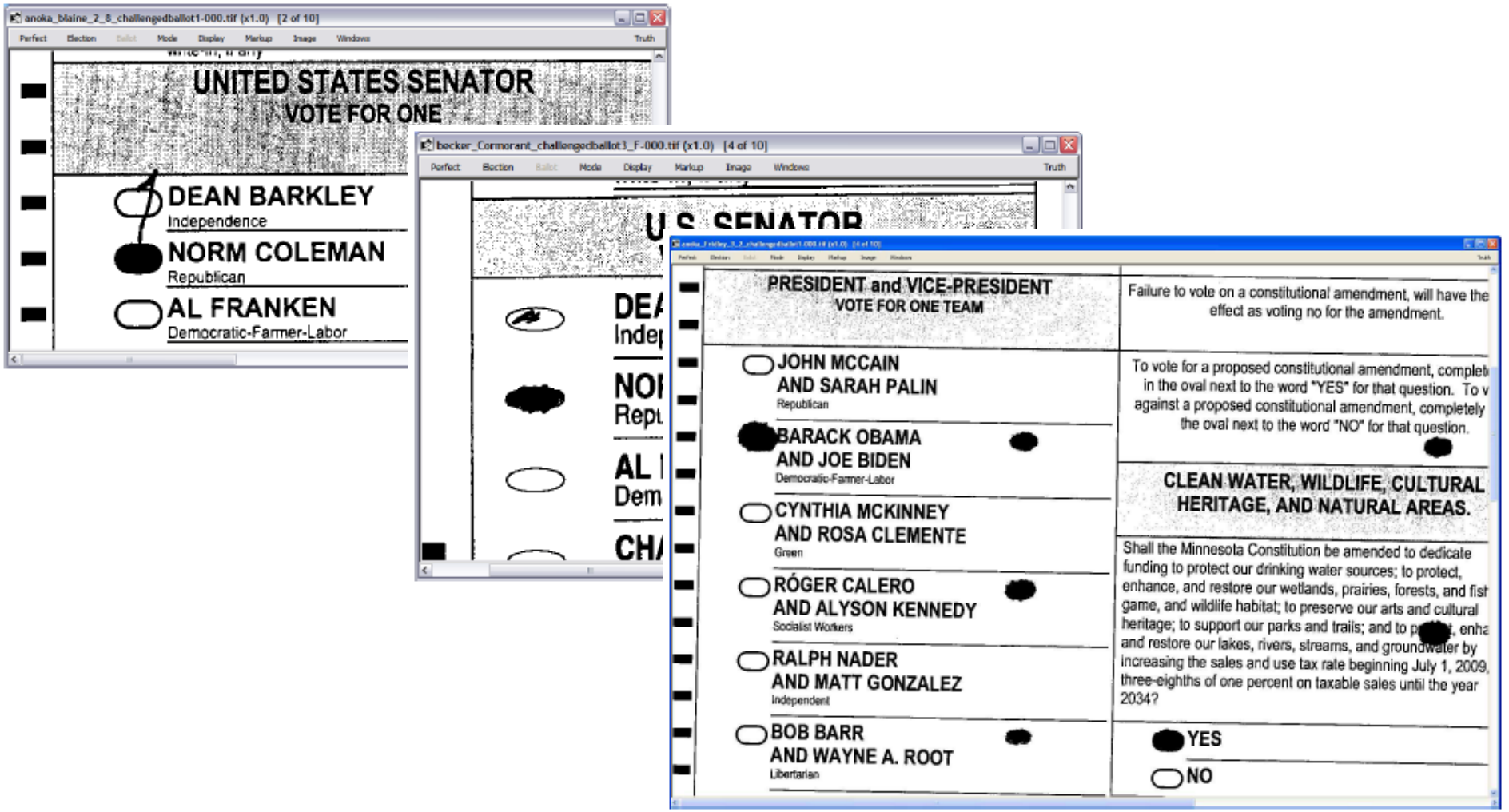


Votes that Look Cancelled

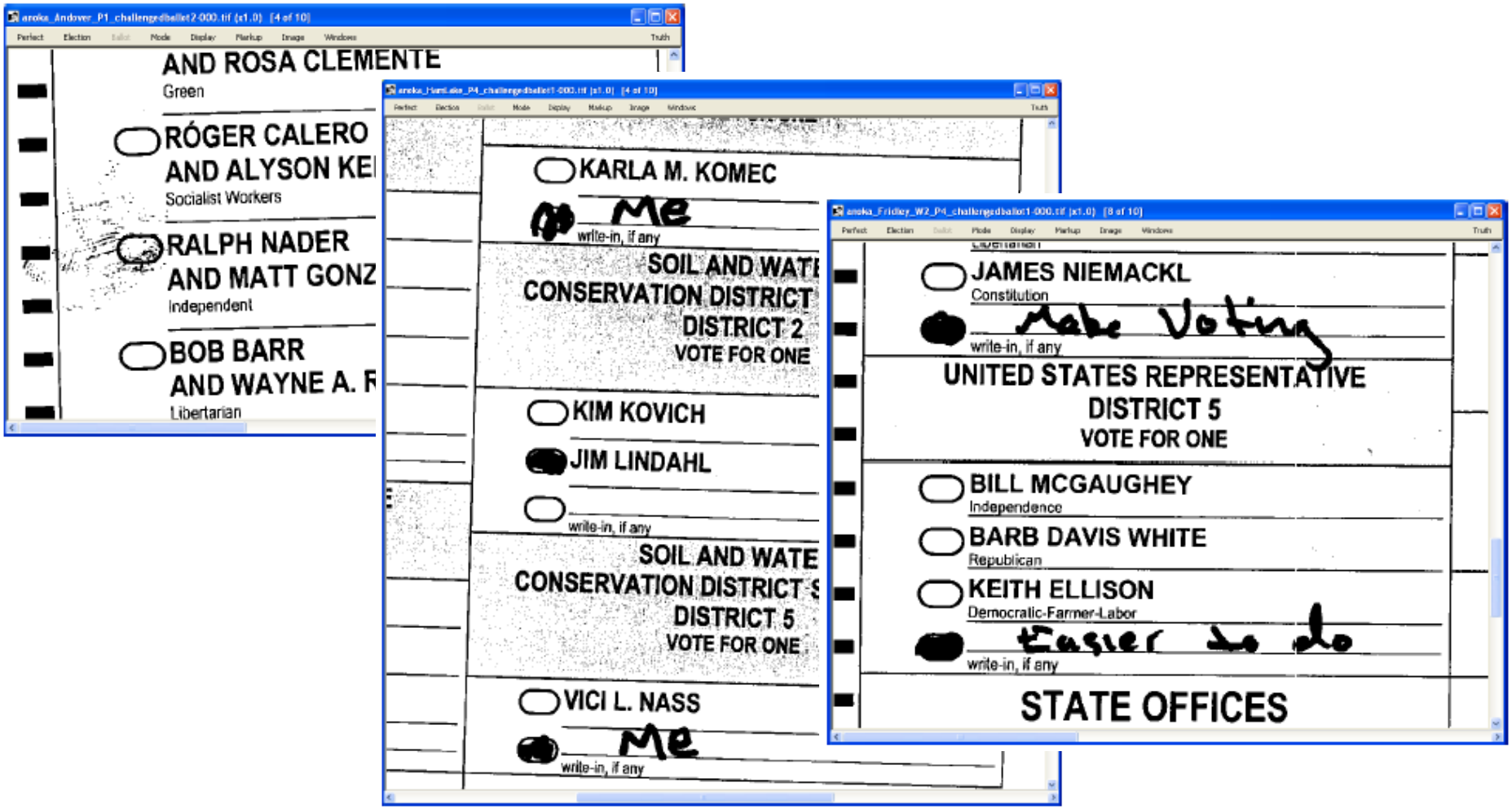




Stray Marks and Bleedthrough



Invalidating Markings



Another Example of Recent Interest

Note that ballots were counted by hand in this case.

See Dealing with doubtful paper ballots in GB :

http://www.electoralcommission.org.uk/__data/assets/pdf_file/0012/87699/UKPE-doubtfuls-booklet.pdf

Why isn't this an easy problem?

After all, ballots are just a simple type of form. We must read votes correctly, but we aren't expected to recognize write-ins.

Can't we just push up reject rate until accuracy reaches 100%?

Remember, we can't change rules in ways that violate the law. VOTER INTENT is the definition we must always follow.

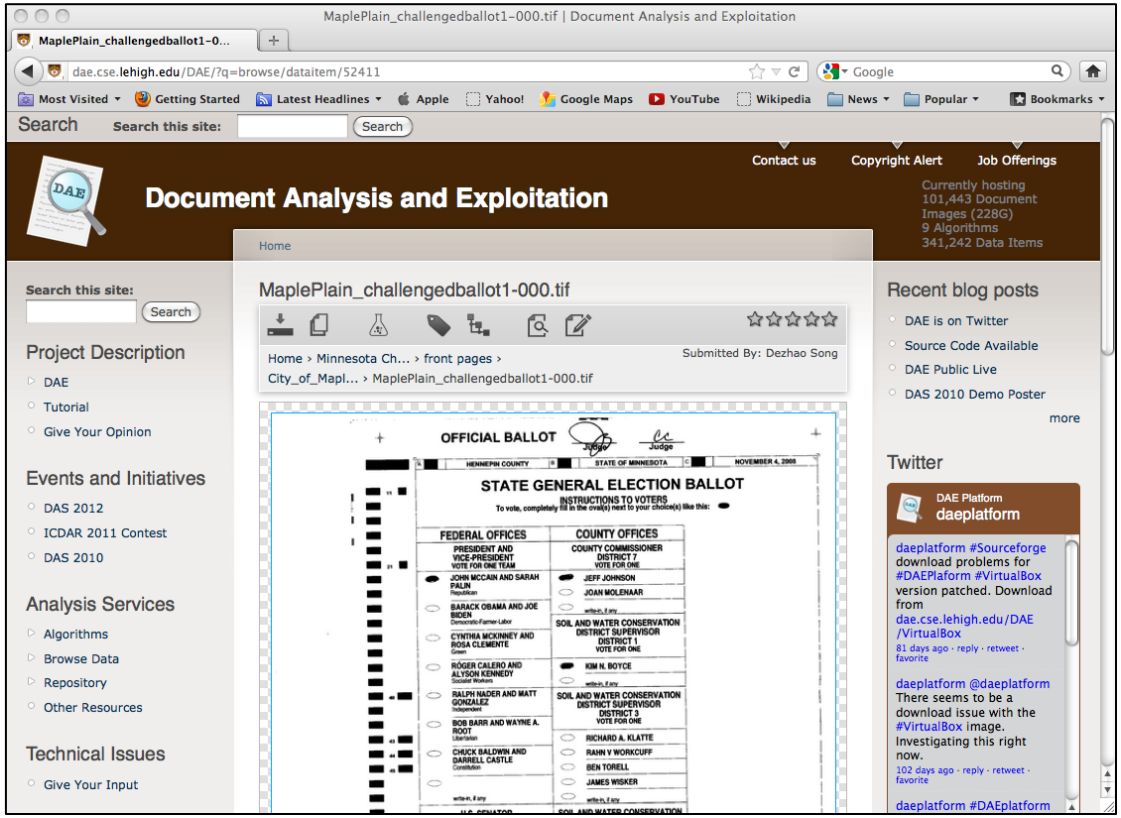
To do this right, we must be prepared to:

- Reject any ballot that may contain "identifying marks."
- Recognize intent when mark is atypical or far from target.
- Accurately identify when a vote has been cancelled.

Status

- Interpretations collected from 8 test subjects, 980 ballot sides.
- All 6,737 ballots now online on DAE server (see URL for more details on server and its capabilities: joint work with Bart).
- Approach is a bit traditional, so far …

http://dae.cse.lehigh.edu/DAE/



"An Open Architecture for End-to-End Document Analysis Benchmarking," B. Lamiroy and D. Lopresti, Proceedings of the Eleventh International Conference on Document Analysis and Recognition (ICDAR 2011), September 2011, Beijing, China, pp. 42-47.

Very Close Indeed

That's all well and good. But what really happened in Minnesota?

Date	Description	Votes for Coleman	Votes for Franken
11/18/08	Initial State Canvassing Board meeting.	1,211,590	1,211,375
12/5/08	After hand recount, not including challenged ballots.	1,209,240	1,209,228
12/20/08	After review of challenged ballots by State Canvassing Board.	1,211,901	1,211,950
1/5/09	After counting of improperly rejected absentee ballots by order of Minnesota Supreme Court. This total was certified by State Canvassing Board.	1,212,206	1,212,431
4/13/09	After counting of improperly rejected absentee ballots by order of three-judge panel during the election contest.	1,212,317	1,212,629

http://www.sos.state.mn.us/elections-voting/2008-general-election-results/2008-state-recounts/

FAQ for Official Recount

WHEN DID THE RECOUNT BEGIN?

WHO SITS ON THE STATE CANVASSING BOARD?

HOW WILL THE RECOUNT WORK?

Every single vote cast for the U.S. Senate candidates will be recounted by hand.

The official recount is being conducted in approximately 110 locations throughout the state, generally in every county courthouse and in the city halls of major cities. In some locations more than one recount "station" will be used depending on the size of the jurisdiction.

The people doing the recounting are county election officials and election judges. Teams of recounters will examine each ballot and record the vote.

As many as four, perhaps even more, observers have been present as each ballot is recounted -- the election judge doing the recounting, representatives from each candidate's campaign, and any other interested parties. The recounts and canvassing board meetings are all open to the public.

https://www.mprnews.org/story/2008/11/06/recount_faq

WHAT ARE THE RECOUNT OFFICIALS LOOKING FOR?

The recounters are trying to determine the intent of the voter when they encounter problem ballots.

Most voters fill in the circle next to the candidate they choose. But sometimes an individual will put a check mark or an X next to a name. Others will circle a name. Ballots marked in that way cannot be scanned by the voting machines, so they wouldn't have been counted the first time around.

If a voter's intentions aren't clear by looking at a ballot, or if there is any objection to the decision being made by the election official by either one or both of the candidates' representatives, the ballots in dispute become "challenged" ballots that will go to the State Canvassing Board for review.

HOW LONG WILL THIS PROCESS TAKE?

HOW MUCH WILL THIS COST?

COULD THIS END UP BEING TAKEN TO COURT?

HOW WOULD THE CASE PROCEED?

WHY WOULD A CANDIDATE CHOOSE THIS PROCESS INSTEAD OF GOING WITH THE SECRETARY OF STATE'S RECOUNT?

More Details on Official Recount

POLITICS & POLICY

Judges' 'three votes' give Al Franken convincing win in Senate recount trial

By Jay Weiner | 04/13/09 Email Share Tweet Print

When this 2008 U.S. Senate race is finally over, when all the appeals are exhausted, when its history is written – based on what we know today -- it will be said that Al Franken won the election and the seat of the junior senator from Minnesota by 312 votes.

But that will not be exactly correct.

After Monday's long-awaited final legal ruling (PDF), add three more votes to Franken's tally: those of Judges Elizabeth Hayden, Kurt Marben and Denise Reilly.

Technically, the three-judge panel that oversaw a seven-week-long trial that generated 19,181 pages of legal filings "voted" against Norm Coleman in their unanimous 56-page opinion, with another 12 pages of exhibits.

It was Coleman's case to prove, and now he'll get another chance when he appeals today's ruling to the Minnesota Supreme Court within 10 days.

https://www.minnpost.com/politics-policy/2009/04/judges%E2%80%99-%E2%80%98three-votes%E2%80%99-give-al-franken-convincing-win-senate-recount-trial

Deciphering Official Recount Results

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	County Name	County Number	Precinct Number	Precinct Name	**Nov. 4, 2008**	**Nov. 4, 2008**	**RECOUNT**					Ballot Disposition for COLEMAN		Ballot Disposition for FRANKEN		Ballot Disposition for Other		Change in Ballots Counted for COLEMAN	Change in Ballots Counted for FRANKEN	**Final Recount Totals for COLEMAN**	**Final Recount Totals for FRANKEN**
2					Votes Counted for COLEMAN	Votes Counted for FRANKEN	Number of Ballots for COLEMAN (as recounted)	Number of Ballots for FRANKEN (as recounted)	Number of All Other Ballots (as recounted)	COLEMAN and Other Ballots Challenged by FRANKEN	FRANKEN and Other Ballots Challenged by COLEMAN	W	CB	W	CB	W	CB				
144	ANOKA	2	4810	FRIDLEY W	407	547	407	547	206	0	0	0	0	0	0	0	0	0	0	407	547
145	ANOKA	2	4820	FRIDLEY W	575	710	574	709	292	2	1	1	0	1	0	1	0	0	0	575	710
146	ANOKA	2	4830	FRIDLEY W	231	416	231	417	153	1	0	0	0	0	0	1	0	0	1	231	417
147	ANOKA	2	4840	FRIDLEY W	297	498	297	497	186	0	1	0	0	1	0	0	0	0	0	297	498
148	ANOKA	2	5010	HAM LAKE	541	344	539	344	187	3	0	0	2	0	0	0	1	0	0	541	344
149	ANOKA	2	5020	HAM LAKE	850	396	849	396	252	1	0	1	0	0	0	0	0	0	0	850	396
150	ANOKA	2	5030	HAM LAKE	606	360	606	360	242	0	0	0	0	0	0	0	0	0	0	606	360
151	ANOKA	2	5040	HAM LAKE	904	488	902	488	283	2	1	2	0	0	1	0	0	0	1	904	489
152	ANOKA	2	5050	HAM LAKE	1029	531	1025	531	320	4	1	4	0	0	0	0	1	0	0	1029	531
153	ANOKA	2	5060	HAM LAKE	895	436	895	436	296	0	1	0	0	1	0	0	0	0	1	895	437
154	ANOKA	2	5210	HILLTOP P	67	189	67	188	69	0	1	0	0	1	0	0	0	0	0	67	189
155	ANOKA	2	5410	LEXINGTO	350	401	348	401	215	3	0	1	1	0	0	1	0	0	0	350	401
156	ANOKA	2	5610	LINO LAKE	511	398	511	396	249	0	2	0	0	2	0	0	0	0	0	511	398
157	ANOKA	2	5620	LINO LAKE	821	588	822	588	331	1	0	1	0	0	0	0	0	2	0	823	588
158	ANOKA	2	5630	LINO LAKE	647	504	647	505	299	1	0	1	0	0	0	0	0	1	1	648	505
159	ANOKA	2	5640	LINO LAKE	1177	548	1174	548	310	3	0	3	0	0	0	0	0	0	0	1177	548
160	ANOKA	2	5650	LINO LAKE	487	383	485	383	204	3	0	1	1	0	0	1	0	0	0	487	383
161	ANOKA	2	5660	LINO LAKE	1025	444	1023	442	275	2	2	2	0	1	1	0	0	0	0	1025	444
162	ANOKA	2	5670	LINO LAKE	1007	545	1005	544	297	1	1	1	0	0	0	0	1	-1	-1	1006	544
163	ANOKA	2	5810	LINWOOD	1460	983	1455	981	606	4	5	3	2	2	0	1	1	0	0	1460	983
164	ANOKA	2	5910	NOWTHEN	1452	677	1448	672	479	4	5	4	0	2	2	0	1	0	-1	1452	676

Freely available as MS Excel file. But note ambiguity: work is needed to translate this into decisions on a ballot-by-ballot basis.

https://www.sos.state.mn.us/media/1979/2008-final-recount-summary-by-precinct.xls

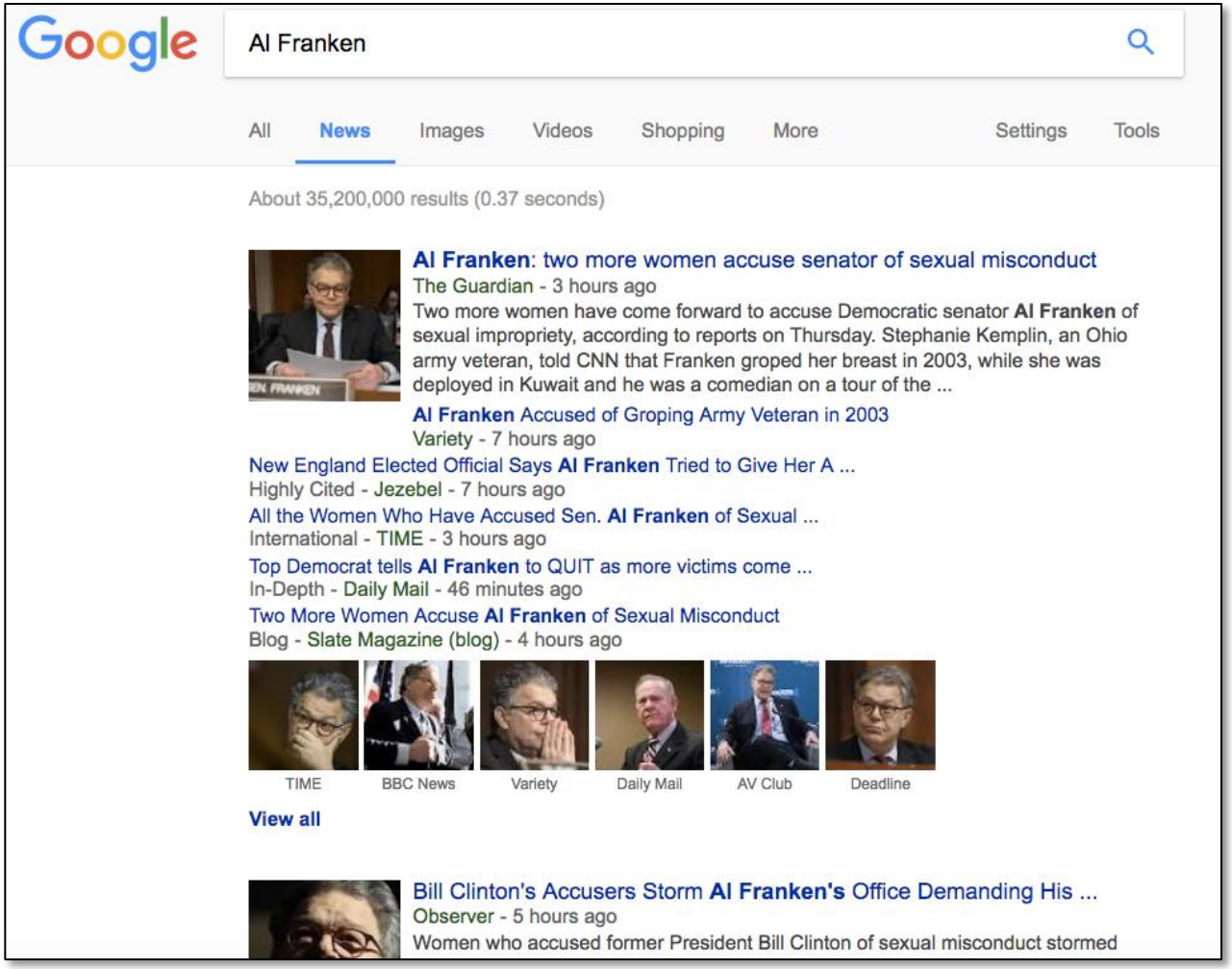
What Can Be Learned Here?

Counting ballots not just an abstract pattern recognition problem:

- A real task defined by pre-determined laws and processes.
- Important to society (not just labeling "cute cat" photos).
- Inherently political, but designed to be as fair as possible.
- Expressed in terms of human interpretation.
- Ambiguity is utterly inherent (real world is messy).
- "Noisy labeling" is utterly inherent.
- AI (pattern recognition) can and must do better.

In other words, this is a perfect problem to study for those of us who want our research to have an impact in the real world.

A Sad Epilogue …



Hopefully this has given you some points to think about … if you're interested in collaborating to turn MN ballots into a community resource for exploring interesting and important "noisy labeling" problem, let me know!

Thank you!
Merci!!