



What are the limitations of Tabari?

The Conflict ITR Team
Rice University



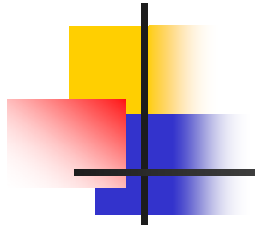
The experimental context

- 50 randomly selected stories on the Middle East from the Reuters corpus (August 1996 to August 1997).
- Only the lead sentence is coded.
- We let Tabari code these sentences with the Levant and Gulf coding dictionaries.
- We coded these sentences (5 per week) manually during a 10 week period in the summer of 2004.



The goal

- The question: should we modify Tabari (since we have the source code) or build a new event coder from scratch?
- Approach: determine how close Tabari comes to human coding, and identify systematic coding failures, if any. If the number and type of coding failures are small, consider “tweaking” Tabari. Otherwise, consider other approaches that allow robust automated coding of events from Reuters stories.



Summary of results

TABARI

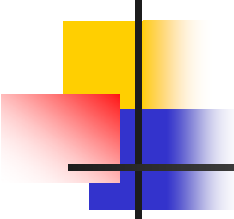
RiceTeam

		Coded	Not coded
Coded	Coded	24	7
	Not coded	9	10



Summary (in words)

- Tabari coded 9 sentences that we did not code, and we coded 7 sentences that Tabari did not code.
- There were 10 of the 50 total sentences that neither we nor Tabari coded.
- Both Tabari and we coded 24 of the 50 sentences.
- At this level, accuracy of Tabari is $34/50=68\%$.



A closer look at sentences coded by us and missed by Tabari

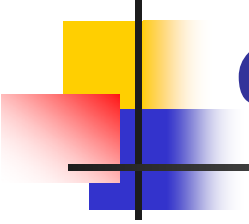
- 23355: Iran has denied a U.S. claim that Tehran was involved in an Iraqi scheme to sell oil in violation of United Nations sanctions.
 - [IRN USA 141](#)
- 39253: Syria firmly told Turkey on Saturday that it was against a Turkish security zone in northern Iraq, saying the move threatened Iraq's national integrity.
 - [SYR TUR 022](#)
- 44561: The United States urged Tehran and Baghdad on Monday to stop fueling inter-Kurdish violence in northern Iraq as the crisis began to grab center stage in the U.S. presidential campaign.
 - [USA IRN 102](#)
- 277991: The United States voiced regret on Friday at France's decision to withdraw from allied air patrols aimed at safeguarding Kurdish civilians in northern Iraq but said it saw no link to any other French-U.S. dispute.
 - [USA FRN 121](#)



A closer look(contd.)

- 468126: The U.N.'s Iraq Sanctions Committee on Monday approved procedures to deliver humanitarian aid to Iraqis on a prioritised need basis, pushing acceptance of food and medicine supply contracts ahead of less vital commodities.
 - UNO IRQ 042
- 580293: The air embargo on Iraq has caused the death of 1,918 people since it was imposed in 1990 for Baghdad's invasion of Kuwait, Health Minister Umeed Mubarak was quoted as saying on Tuesday.
 - IRQ IRQ 022
- 804197: A United Nations germ warfare team has ended a week-long inspection of Iraq's past biological weapons, a U.N. arms official said on Saturday.
 - UNO IRQ 020

Out of 7 sentences, 3 contain comment events only.



Sentences TABARI coded and we missed

- 23450: A U.N. Security Council committee on Wednesday put off Turkey's request to buy Iraqi oil until the impact of a limited oil sales deal with Iraq could be evaluated.
 - IRQ TUR 050
 - No verb.
- 24974: Iraq's U.N. ambassador Nizar Hamdoun said on Friday that minor problems still needed to be resolved before his country's oil-for-food deal with the United Nations could be implemented.
 - IRQ UNO 050
 - No target
- 36446: Dutch industrial inspectors Saybolt have withdrawn an advance team of three from Baghdad in the light of this week's Iraqi crisis and with the oil-for-food deal looking more distant, a spokesman said.
 - NTH IRQ 050
 - Lack of authority, no initiator



Tabari coded, we didn't

- 38368:Iraq's official media heaped scorn on President Bill Clinton on Friday, calling him an ignorant "womaniser".
 - IRQ USAGOV 111
 - No verb
- 39802:Foreign Ministers from the six Gulf Arab states meet in Saudi Arabia on Sunday to unify their response after last week's missile strikes against Iraq by the United States, their key Western ally.
 - ARB SAU 022 and SAU ARB 023
 - No initiator and no target
- 41477:President Bill Clinton said on Monday the United States is helping "those who have worked with us" to flee Iraq amid Kurdish strife that reportedly wrecked a CIA effort to topple Iraqi President Saddam Hussein.
 - USAGOV USA 010
 - No target



Tabari coded, we didn't

- 121455: Iran is ready to mediate an end to the fighting between Iraq's Kurdish factions, Iranian Foreign Minister Ali Akbar Velayati was quoted as saying.
 - IRN IRQ 025
 - No target, no verb
- 250217: Iraq's Oil Minister Amir Muhammad Rasheed denied on Wednesday reports that the flow of Iraqi oil to Turkey was facing any political or technical problems.
 - IRQ TUR 126
 - No target
- 733642: Iraq's President Saddam Hussein will deliver a televised speech on Thursday to mark the 29th anniversary of the revolution which brought his Baath party to power, the Iraqi News Agency INA reported on Wednesday.
 - IRQ ARBBA 070
 - No verb and no target.

7 out of 9 sentences, coded incorrectly. 2 could have been coded by us with an extended actor dictionary.



Sentences coded by Tabari and by us

- Of the 24 sentences, that both Tabari and we coded
 - 1 exact match on (initiator, target, 3-digit Weis code).
 - 3 matches on (initiator, target, 1st two digits of Weis code).
 - 20 mismatches on either initiator, target or Weis code.
 - 4 initiator and target misidentifications each.
 - 20 misidentifications of Weis code.



Accuracy tally

- Tabari and we agree on $14/50 = 28\%$ of the sentences.



A closer look

- 31525: Iranian President Akbar Hashemi Rafsanjani criticised the United States on Wednesday for its cruise missile attacks against Iraq but said the Iraqis had provided the pretext.
 - Us: IRN USA 121, IRN IRQ 020
 - Tabari: IRN USA 111
 - 111 is turn down (-4.0) and 121 is criticize (-2.2)



A closer look

- 39845: Syria on Sunday warned Turkey against setting a security zone in northern Iraq, saying this would put Ankara in conflict with Arab states and threaten the stability of the Middle East.
 - Us: SYR TUR 160
 - Tabari: SYR TUR 130
 - 130 is protest (-1.9) and 160 is warn (-3)

Close enough



A closer look

- 50115: Iraq said on Thursday it was no longer necessary for Turkey to set up a proposed security zone in northern Iraq after a Baghdad-backed Kurdish militia took control of the region earlier this week.
 - Us: IRQ TUR 020
 - Tabari: IRQ TUR 010
 - 020 is comment (0.0) and 010 is yield (1.0)

Close enough.



A closer look

- 58481: The United States on Tuesday announced plans for a new diplomatic initiative with Iraqi Kurds, insisting it was not in the Kurds' long-term interest to remain allied with Iraqi President Saddam Hussein.
 - Us: USA KUR 025
 - Tabari: USA KUR 068, KUR USA 068
 - 025 is explain position (0.0) and 068 is ?



A closer look

- 58400: Kurdish leader Massoud Barzani is to meet a senior U.S. official in Turkey, just days after his fighters took control of northern Iraq with the backing of Baghdad's troops, a Kurdish group said on Tuesday.
 - Us: KUR USA 082
 - Tabari: USA KUR 100
 - Target and initiator interchanged!



A closer look

- 86317: United Nations Under-Secretary General Yasushi Akashi said on Monday he hoped the Iraq "oil-for-food" deal could go ahead in the near future.
 - Us: UNO IRQ 024
 - Tabari: UNO IRQ 013
 - 024 is opt. comment (0.4) and 013 is retract (2.0)



A closer look

- 131443: Chief U.N. weapons inspector Rolf Ekeus said on Monday that Iraq had not given any new information about unaccounted weapons but added the two agreed to focus on this issue.
 - Us: UNO IRQ 080, IRQ UNO 080, UNO IRQ 020, IRQ UNO 020
 - Tabari: UNO IRQ 090
 - 080 is agree (6.0) and 090 is request (3.0)



A closer look

- 173150: In a warning to the West that the world oil supply is vulnerable, Iraq's chief oil negotiator said on Thursday that the embargo on Iraqi oil was "dangerous" given rising crude prices and limited spare production capacity.
 - Us: IRQ IRQ 022
 - Tabari: West IRQ 012
 - 022 is pess. Comment (-0.4) and 012 is retreat (0.6)
 - Actors are different!



A closer look

- 211961: Baghdad has called on the United Nations to investigate alleged Turkish bombings and incursions into northern Iraq, official newspapers reported on Friday.
 - Us: IRQ UNO 094
 - Tabari: IRQ UNO 051
 - 094 is call for (-0.1) and 051 is promise policy support (4.5)



A closer look

- 211965: The United Nations official in charge of destroying Iraq's deadliest weapons said in remarks published on Friday that Baghdad has obstructed U.N. efforts to study parts of Scud missiles.
 - Us: UNO UNO 020, IRQ UNO 190
 - Tabari: UNO IRQ 112
 - Initiator and target interchanged!



A closer look

- 459323: Iraq and Russia signed on Friday a contract to develop seven to eight billion barrels of reserves in southern Iraq's Qurana oilfield, Iraqi and Russian officials said.
 - Us: IRQ RUS 081, RUS IRQ 081, IRQ RUS 020, RUS IRQ 020
 - Tabari: IRQ RUS 061, RUS IRQ 061
 - 081 is make agreement (6.5) and 061 is apologize (1.8)



A closer look

- 464168: A senior Russian diplomat on Tuesday will visit Kuwait, which has expressed concern over a new Moscow-Baghdad oil deal, diplomats said on Monday.
 - Us: RUS KUW 032
 - Tabari: RUS KUW 068
 - 068 is ? and 032 is visit (1.9)



A closer look

- 489889:Iraq has given the U.N.'s chief arms inspector a detailed report on its chemical weapons, newspapers said on Sunday.
 - Us: IRQ UNO 066
 - Tabari: IRQ UNO 090
 - 066 is release (1.9) and 090 is request (3.0)



A closer look

- 502824: U.N. overseers have approved a 48th oil contract for Iraq under terms of the oil-for-food accord, allowing a spot deal for 100,000 barrels, a U.N. official said Friday.
 - Us: UNO IRQ 040, UNO IRQ 020
 - Tabari: IRQ UNO 020
 - Initiator and target interchanged!



A closer look

- 633907: The United States will support renewal of the United Nations' oil-for-food programme for Iraq, the State Department said on Tuesday.
 - Us: USA IRQ 042, USA IRQ 025
 - Tabari: USA IRQ 031
 - 042 is endorse (3.6) and 025 is explain position (0.0) and 031 is meet (1.0)



A closer look

- 645087: Iran on Sunday called on Turkey to immediately pull its troops out of northern Iraq, the official Iranian news agency IRNA said.
 - Us: IRN TUR 102, IRN TUR 020
 - Tabari: IRN TUR 050
 - 102 is urge (-0.1) and 050 is promise (4.0)



A closer look

- 662350: Iraqi leader Saddam Hussein assailed the United States over a vote by its House of Representatives reaffirming an earlier commitment to move the U.S. embassy in Israel to Jerusalem by 1999, Baghdad radio reported on Sunday.
 - Us: IRQ USA 121, IRQ USA 020
 - Tabari: IRQ USA 111
 - 121 is criticize (-2.2) and 111 is turn down (-4.0)



A closer look

- 727092: Iraq has postponed a visit by a business delegation to Lebanon after Beirut refused to allow Iraqi athletes to participate in the Pan Arab Games, the Iraqi News Agency INA reported on Monday.
 - Us: IRQ LEB 191, IRQ LEB 020
 - Tabari: IRQ ARB 120
 - Target is misidentified, and the event codes are also incorrect!



A closer look

- 744443: Richard Butler, the U.N. official in charge of scrapping Iraq's weapons of mass destruction, headed for Baghdad on Monday for talks with Iraqi officials, a United Nations spokesman said.
 - Us: UNO IRQ 032, UNO IRQ 020
 - Tabari: UNO IRQ 112
 - 032 is visit (1.9) and 112 is refuse (-4.0)



A closer look

- 237036: Turkey's Energy Ministry said Iraqi oil began to be pumped into the Iraq-Turkey pipeline around midday on Thursday as a test under United Nations supervision.
 - Us: TUR IRQ 020
 - Tabari: TUR IRQ 010
 - 020 is comment (0.0) and 010 is yield (1.0)



Summary

- 3 out of 24 sentences have target and initiator interchanged.
- 2 out of 24 sentences have a target or initiator misidentified.
- On 1 out of 24 sentences, do the event codes match to 3 digits.
- 20 out of 24 event codes do not match at the 2 digit level.



Need for a new coder?

- The verb dictionary is Tabari's greatest vulnerability.
- Tabari misidentifies the main verb in the sentence and this causes misidentification at the 2 digit Weis code level.
 - A full parse of the sentence rather than a shallow parse will help separate the components of the sentence better.
- Lack of authority rules need to be coded.
- Targets need to be identified with greater accuracy.