Talk, Cheap Talk, and States of Knowledge

Rohit Parikh

City University of New York

COMSOC 2008
September 5, 2008

The peculiar character of the problem of a rational economic order is determined precisely by the fact that the knowledge of the circumstances of which we must make use never exists in concentrated or integrated form, but solely as the dispersed bits of incomplete and frequently contradictory knowledge which all the separate individuals possess.

The peculiar character of the problem of a rational economic order is determined precisely by the fact that the knowledge of the circumstances of which we must make use never exists in concentrated or integrated form, but solely as the dispersed bits of incomplete and frequently contradictory knowledge which all the separate individuals possess. The economic problem of society is thus not merely a problem of how to allocate "given" resources – if "given" is taken to mean given to a single mind which deliberately solves the problem set by these "data."

The peculiar character of the problem of a rational economic order is determined precisely by the fact that the knowledge of the circumstances of which we must make use never exists in concentrated or integrated form, but solely as the dispersed bits of incomplete and frequently contradictory knowledge which all the separate individuals possess. The economic problem of society is thus not merely a problem of how to allocate "given" resources – if "given" is taken to mean given to a single mind which deliberately solves the problem set by these "data." It is rather a problem of how to secure the best use of resources known to any of the members of society, for ends whose relative importance only these individuals know.

F. Hayek
Individualism and Economic Order

Not long ago, if you wanted to seize political power in a country, you had merely to control the army and the police. Today it is only in the most backward countries that fascist generals, in carrying out a coup d'etat, still use tanks. If a country has reached a high level of industrialization, the whole scene changes.

Not long ago, if you wanted to seize political power in a country, you had merely to control the army and the police. Today it is only in the most backward countries that fascist generals, in carrying out a coup d'etat, still use tanks. If a country has reached a high level of industrialization, the whole scene changes. The day after the fall of Khruschev, the editors of *Pravda, Izvestiia,* the heads of the radio and television were replaced; the army wasn't called out.

Not long ago, if you wanted to seize political power in a country, you had merely to control the army and the police. Today it is only in the most backward countries that fascist generals, in carrying out a coup d'etat, still use tanks. If a country has reached a high level of industrialization, the whole scene changes. The day after the fall of Khruschev, the editors of *Pravda, Izvestiia,* the heads of the radio and television were replaced; the army wasn't called out. Today, a country belongs to the person who controls communications.

<div align="right">
Umberto Eco

Towards a Semiological Guerrilla Warfare, 1967
</div>

Three people A, B, C walk into a coffee shop. One of them orders cappuccino, one orders tea, and one orders icecream. The waiter goes away and after ten minutes *another* waiter arrives with three cups. "Who has the cappuccino?" "I do," says A. "Who has the tea?" "I do," says C.

Three people A, B, C walk into a coffee shop. One of them orders cappuccino, one orders tea, and one orders icecream. The waiter goes away and after ten minutes *another* waiter arrives with three cups. "Who has the cappuccino?" "I do," says A. "Who has the tea?" "I do," says C.

Will the waiter ask a third question?"

Consider the possible situations for waiter 2. They are

1) CTI,   2) CIT,
3) TCI,   4) TIC,
5) ICT,   6) ITC

Consider the possible situations for waiter 2. They are

1) CTI,   2) CIT,
3) TCI,   4) TIC,
5) ICT,   6) ITC

When A says that he has the cappuccino, 3,4,5,6 are eliminated.

Consider the possible situations for waiter 2. They are

1) CTI,   2) CIT,
3) TCI,   4) TIC,
5) ICT,   6) ITC

When A says that he has the cappuccino, 3,4,5,6 are eliminated.
The waiter now has,

1) CTI,   2) CIT

Consider the possible situations for waiter 2. They are

1) CTI,   2) CIT,
3) TCI,   4) TIC,
5) ICT,   6) ITC

When A says that he has the cappuccino, 3,4,5,6 are eliminated.
The waiter now has,

1) CTI,   2) CIT

When C says that he has the tea, 1 is eliminated.

Now 2 alone is left and the waiter knows that B has the icecream.

A butler enters a hotel room to clean it and make the bed, but he encounters a woman guest, coming out of the bathtub and not even wearing a towel.

"Excuse me, sir," says the butler, and leaves the room.

A butler enters a hotel room to clean it and make the bed, but he encounters a woman guest, coming out of the bathtub and not even wearing a towel.

"Excuse me, sir," says the butler, and leaves the room.

Why did the butler say, "Excuse me, sir"?

A butler enters a hotel room to clean it and make the bed, but he encounters a woman guest, coming out of the bathtub and not even wearing a towel.

"Excuse me, sir," says the butler, and leaves the room.

Why did the butler say, "Excuse me, sir"?

In the woman's mind there were two possibilities.

A butler enters a hotel room to clean it and make the bed, but he encounters a woman guest, coming out of the bathtub and not even wearing a towel.
"Excuse me, sir," says the butler, and leaves the room.

Why did the butler say, "Excuse me, sir"?

In the woman's mind there were two possibilities.
S1 = "The butler saw her clearly"

A butler enters a hotel room to clean it and make the bed, but he encounters a woman guest, coming out of the bathtub and not even wearing a towel.

"Excuse me, sir," says the butler, and leaves the room.

Why did the butler say, "Excuse me, sir"?

In the woman's mind there were two possibilities.
S1 = "The butler saw her clearly"
S2 = "The butler did not see her clearly"
The butler's remark eliminated S1 and saved her from embarrassment.

Numerical Foreheads

Two players Ann and Bob are told that the following will happen. Some positive integer n will be chosen and *one* of n, $n+1$ will be written on Ann's forehead, the other on Bob's. Each will be able to see the other's forehead, but not his/her own.

Note that each can see the other's number, but not their own. Thus if Ann has 5 and Bob has 6, then Ann knows that her number is either 5 or 7 and Bob knows that his number is either 6 or 4.

After this is done, they are asked repeatedly, beginning with Ann, if they know what their own number is.

Theorem 1: In those cases where Ann has the even number, the reponse at the nth stage will be, "my number is $n + 1$", and in the other cases, the response at the $(n + 1)$st stage will be "my number is $n + 1$". In either case, it will be the person who *sees* the smaller number, who will respond first.

Definition 1: A *Kripke model M* for a (two person) knowledge situation consists of a state space W and two equivalence relations $\equiv_1$ and $\equiv_2$. Intuitively $s \equiv_1 t$ means that states s and t are indistinguishable to player 1 (Ann) and $s \equiv_2 t$ means that they are indistinguishable to player 2 (Bob). We shall assume in this talk that W is finite or countable.

In the example we are looking at,
$W = \{(m, n) | m, n \epsilon N^+ \text{ and } |m - n| = 1\}$.
If $s, t \epsilon W$ and $i \epsilon \{1, 2\}$, then $s \equiv_i t$ iff $(s)_j = (t)_j$, where $j = 3 - i$, and $(s)_j$ is the j-the component of s.

Intuitively, $s \equiv_i t$ means that when the dialogue begins, player i cannot distinguish between s and t, where Ann is player 1 and Bob is player 2.

Definition 2: A subset X of W is *i-closed* if $s \epsilon X$ and $s \equiv_i t$ imply that $t \epsilon X$. X is *closed* if it is both 1-closed and 2-closed.

The subset where Ann has the odd number is closed, as is the subset where Bob has the odd number.

Definition 3: Given Kripke model M, $X \subseteq W$, and $s \epsilon X$, then i *knows X at s* iff for all t, $s \equiv_i t$ implies that $t \epsilon X$. In other words, there is an i-closed subset Y such that $s \in Y$ and $Y \subseteq X$. X is *common knowledge* at s iff there is a closed set Y such that $s \epsilon Y \subseteq X$.

Thus if Ann has an odd number then that fact is common knowledge.

Observation: If an announcement of a formula ϕ is made, then the new Kripke structure is obtained by deleting all states $s \in W$ where ϕ is false. (Note: there *are* some caveats.)

We now look at what is happening when Ann has 5 and Bob has 6.

Start situation

.
.

(5,6)

(5,4)

(3,4)

(3,2)

(1,2)
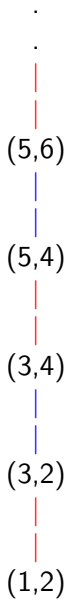
Bob has just said, I don't know my number

$\vdots$

(5,6)

(5,4)

(3,4)

(3,2)

(1,2)

Ann said no also

.
.

(5,6)

(5,4)

(3,4)

(3̶,2̶)

(1̶,2̶)

Bob said a second "no"

.

.

(5,6)

(5,4)

(~~3,4~~)

(~~3,2~~)

(~~1,2~~)
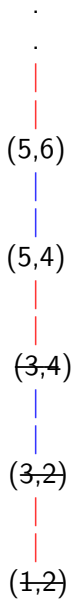
Ann said a second "no"

.

.

(5,6)

(5̶,4̶)

(3̶,4̶)

(3̶,2̶)

(1̶,2̶)
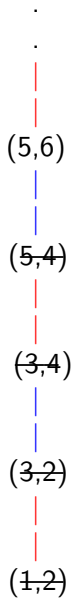
Bob knows his number is 6

A social structure with certain logical properties is a queue, like at a bus stop or in a bank.

- ▶ Someone who came earlier gets service earlier.
- ▶ Violations are easily detectable.

The problem of parking is a similar problem. A scarce resource needs to be allocated on the basis of some sort of priority, which, however, is difficult to determine.

When people are looking for parking in a busy area, they tend to cruise around until they find a space. There is no queue as such, but in general we do want that someone who arrives first should find a parking space and someone who arrives later may not.

When my students and I studied cruising for parking in a 15-block business district in Los Angeles, we found the average cruising time was 3.3 minutes, and the average cruising distance half a mile (about 2.5 times around the block). This may not sound like much, but with 470 parking meters in the district, and a turnover rate for curb parking of 17 cars per space per day, 8,000 cars park at the curb each weekday. Even a small amount of cruising time for each car adds up to a lot of traffic.

Over the course of a year, the search for curb parking in this 15-block district created about 950,000 excess vehicle miles of travel, equivalent to 38 trips around the earth, or four trips to the moon. And here's another inconvenient truth about underpriced curb parking: cruising those 950,000 miles wastes 47,000 gallons of gas and produces 730 tons of the greenhouse gas carbon dioxide. If all this happens in one small business district, imagine the cumulative effect of all cruising in the United States.   Donald Shoup

Shoup regards this problem as one of incentive and suggests that parking fees be raised so that occupancy of street parking spaces is only 85%.

But perhaps this is really a knowledge problem?

Find a Place to Park on Your GPS – Spark Parking Makes it Possible

Navigation Developers Can Access Spark Parking Points of Interest Through New Tele Atlas ContentLink Program

San Francisco, CA, March 21, 2007

Running late for a meeting and worried about finding a place to park? Unhappy about paying outrageous valet parking fees at your favorite restaurant? These headaches will soon be a thing of the past. Spark Parking's detailed parking location information data is now available through the newly released Tele Atlas ContentLinkSM portal for application developers to incorporate into a range of GPS devices and location-based services and applications.

Spark Parking's detailed parking information provides the locations of every paid parking facility in each covered city – from the enormous multi-level garages to the tiny surface lots hidden in alleys. In addition, Spark Parking includes facility size, operating hours, parking rates, available validations, and many more details not previously available from any source. As a result, drivers will easily be able to find parking that meets their needs and budgets. http://www.pr.com/press-release/33381

Where's the bus? NextMuni can tell you.
System uses GPS to let riders know when streetcar will arrive

Rachel Gordon, Chronicle Staff Writer
Thursday, March 29, 2007

San Francisco's Municipal Railway may have a hard time running on time, but at least the transit agency is doing more to let riders know when their next bus or streetcar is due to arrive.

The "NextMuni" system, which tracks the location of vehicles via satellite, is now up and running on all the city's electrified trolley bus lines. It had been available only on the Metro streetcar lines and the 22-Fillmore, a trolley bus line that served as an early test. The whereabouts of the Global Positioning System-equipped vehicles are fed into a centralized computer system that translates the data into user-friendly updates available on the Internet and on cell phones and personal digital assistants.

http://www.sfgate.com/

Common Knowledge

Defined independently by Lewis and Schiffer. Used first in Game theory by Aumann.

Aumann showed that common knowledge implies same opinion.

Common Knowledge

Defined independently by Lewis and Schiffer. Used first in Game theory by Aumann.

Aumann showed that common knowledge implies same opinion.

Geanakoplos and Polemarchakis showed that communication between two agents leads to common knowledge and same opinion.

Common Knowledge

Defined independently by Lewis and Schiffer. Used first in Game theory by Aumann.

Aumann showed that common knowledge implies same opinion.

Geanakoplos and Polemarchakis showed that communication between two agents leads to common knowledge and same opinion.

Parikh and Krasucki showed that among n agents communicating in pairs, common opinion about some quantity can come about without most agents communicating with others.

Aumann's argument

Column

Now Row's value is
$$v = (1/4)[v_{1,1,} + v_{1,2} + v + 1, 3 + v_{1,4}]$$
And Column's value is
$$w = (1/4)[(v_{1,1,} + v_{2,1} + v_{3,1} + v_{4,1}]$$

Now Row's value is
$$v = (1/4)[v_{1,1,} + v_{1,2} + v + 1, 3 + v_{1,4}]$$
And Column's value is
$$w = (1/4)[(v_{1,1,} + v_{2,1} + v_{3,1} + v_{4,1}]$$

Since these values are common knowledge, they must both equal
$$(1/16)[\Sigma v_{i,j} : i \leq 4, j \leq 4]$$

Now Row's value is
$$v = (1/4)[v_{1,1,} + v_{1,2} + v + 1, 3 + v_{1,4}]$$
And Column's value is
$$w = (1/4)[(v_{1,1,} + v_{2,1} + v_{3,1} + v_{4,1}]$$

Since these values are common knowledge, they must both equal
$$(1/16)[\Sigma v_{i,j} : i \leq 4, j \leq 4]$$

Thus $v = w$.

"I'd never join any club that would have me for a member"

Groucho Marx

**Using Aumann's reasoning,
Milgrom and Stokey proved a famous**
No Trade **theorem!**

If A is selling a stock to B, and B is buying it, then obviously A thinks the stock will go down and B thinks it will go up. But this fact is common knowledge! By a proof based on Aumann, it cannot be common knowledge that they have different views of the stock and the sale cannot take place.

But what if the value is not common knowledge?

Will communication help?

GP argument

Column

②2	3	5	4
7	8	9	10
3	2	5	4
5	4	3	2

Row

At this point Row announces that her expected value is 3.5, and column eliminates row 2

Column

(2)	3	5	4
7	8	9	10
3	2	5	4
5	4	3	2

Row

Now column announces that his value is 3.33, and row eliminates columns 2,3

Column

(2)	3	5	4
7	8	9	10
3	2	5	4
5	4	3	2

Row

Now Row announces his value as 3 = (2+4)/2 and Column eliminates row 3, 4, announcing his value as 2.

Column

(2)	3	5	4
7	8	9	10
3	2	5	4
5	4	3	2

Row

At this point Row eliminates column 4, also announces his value at 2, and they have consensus.

Column

|			
②2	3	5	4
7	8	9	10
3	2	5	4
5	4	3	2

Row

A brief overview of the [PK] result:

Suppose we have n agents connected in a strongly connected graph. They all share initial probability distribution, but have now received, each of them, a finite amount of private information. Thus their estimate of the probability of some event or the expected value of some random variable v may now be different.

Let g be a function which, at stage n picks out a sender $s(n)$ and a recipient $r(n)$. $s(n)$ sends his latest value of v to $r(n)$ who then revises *her* valuation of v.

If the graph G is strongly connected, and for each pair of connected agents i, j, i repeatedly sends his value of v to j, then eventually all estimates of the value of v become equal.

Parikh-Krasucki result
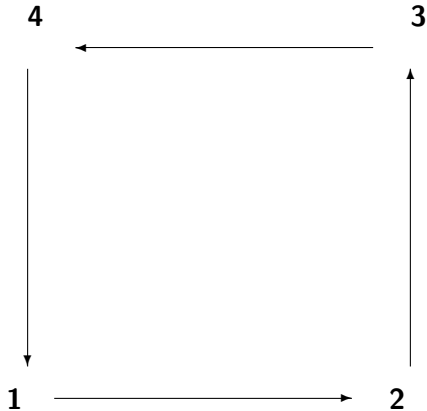
History Based Knowledge

On Monday Jack writes to Ann that he got a dog (**D**) E_1

History Based Knowledge

On Monday Jack writes to Ann that he got a dog (**D**) E_1
On Wednesday Ann receives his letter, E_2

History Based Knowledge

On Monday Jack writes to Ann that he got a dog (**D**) E_1
On Wednesday Ann receives his letter, E_2
On Thursday, Jack looks at the calendar and sees that three days have passed since he wrote, E_3

History Based Knowledge

On Monday Jack writes to Ann that he got a dog (**D**) E_1
On Wednesday Ann receives his letter, E_2
On Thursday, Jack looks at the calendar and sees that three days have passed since he wrote, E_3

$E_1 \longrightarrow E_2 \longrightarrow E_3$

History Based Knowledge

On Monday Jack writes to Ann that he got a dog (**D**) E_1
On Wednesday Ann receives his letter, E_2
On Thursday, Jack looks at the calendar and sees that three days have passed since he wrote, E_3

$E_1 \longrightarrow E_2 \longrightarrow E_3$

Suppose that a letter takes at most three days to arrive. Then on Wednesday, Ann knows D, but Jack does not know that Ann knows D.
On Thursday, Jack knows that Ann knows that D.

See no Evil, Hear no Evil

A pretty woman (Eve) shoots a man dead.

See no Evil, Hear no Evil

A pretty woman (Eve) shoots a man dead.

Wally, who is blind, hears a shot.

See no Evil, Hear no Evil

A pretty woman (Eve) shoots a man dead.

Wally, who is blind, hears a shot.

Dave, who is deaf, sees a woman leave in a hurry (his back was turned when she fired)

See no Evil, Hear no Evil

A pretty woman (Eve) shoots a man dead.

Wally, who is blind, hears a shot.

Dave, who is deaf, sees a woman leave in a hurry (his back was turned when she fired)

Together they know who committed the murder. But neither of them knows it by himself.

A global history is the sequence of all events which happen.
The corresponding local history for an agent i, is all the events (or aspects of them) which i 'sees'.
The protocol is the set of all possible global histories.
Suppose an agent sees local history h, and X is the set of all global histories which are compatible with h.
If some property P is true of all histories in X, then the agent knows P.

Example 1: Uma is a physician whose neighbour is ill. Uma does not know and has not been informed. Uma has no obligation (as yet) to treat the neighbour.

Example 1:   Uma is a physician whose neighbour is ill. Uma does not know and has not been informed. Uma has no obligation (as yet) to treat the neighbour.

Example 2:   Uma is a physician whose neighbour Sam is ill. The neighbour's daughter Ann comes to Uma's house and tells her. Now Uma does have an obligation to treat Sam, or perhaps call in an ambulance or a specialist.

Example 1:   Uma is a physician whose neighbour is ill. Uma does not know and has not been informed. Uma has no obligation (as yet) to treat the neighbour.

Example 2:   Uma is a physician whose neighbour Sam is ill. The neighbour's daughter Ann comes to Uma's house and tells her. Now Uma does have an obligation to treat Sam, or perhaps call in an ambulance or a specialist.

The global history contained the event E of Sam being sick, but until Uma was told, she did not know it and did not know that she needed to act.

The Kitty Genovese Murder

"Along a serene, tree-lined street in the Kew Gardens section of Queens, New York City, Catherine Genovese began the last walk of her life in the early morning hours of March 13, 1964.....As she locked her car door, she took notice of a figure in the darkness walking towards her. She became immediately concerned as soon as the stranger began to follow her.

'As she got of the car she saw me and ran,' the man told the court later, 'I ran after her and I had a knife in my hand.... I could run much faster than she could, and I jumped on her back and stabbed her several times,' the man later told the cops."

Many neighbours saw what was happening, but no one called the police.

"Mr. Koshkin wanted to call the police but Mrs. Koshkin thought otherwise. 'I didn't let him,' she later said to the press, 'I told him there must have been 30 calls already.' "

"When the cops finished polling the immediate neighbourhood, they discovered at least 38 people who had heard or observed some part of the fatal assault on Kitty Genovese."[1]

Some 35 minutes passed between Kitty Genovese being attacked and someone calling the police, why?

---

[1]This quote is from the article 'A cry in the night: the Kitty Genovese murder', by a police detective, Mark Gado, and appears on the web in *Court TV's Crime Library*.

Gricean Implicature

Gricean Implicature

A: My car is out of gasoline.

Gricean Implicature

A: My car is out of gasoline.

B: There is a gas station around the corner

Gricean Implicature

A: My car is out of gasoline.

B: There is a gas station around the corner

The assumption is that B is co-operating with A and would not say what he said unless he knew that the gas station was (likely to be) open.

But, can we always believe what others tell us?

But, can we always believe what others tell us?

Sally is applying to Rayco for a job and Rayco asks if her ability is high or low.

.

Sally has nothing to gain by lying about her qualifications and Rayco can trust her.

	Rayco	
	High	**Low**
High	(3,3)	(0,0)
Low	(3,0)	(2,2)

Sally

.

Sally has nothing to lose by lying about her qualifications and Rayco cannot trust her.

The extent to which one agent (the listener) can believe
another agent (the speaker)
depends on how much they have in common.

Something interesting has happened recently in the kerfuffle between Barack Obama and his putative pastor, Jeremiah Wright.

Obama denounced comments made by Wright at the NAACP and at the Press Club.

Something interesting has happened recently in the kerfuffle between Barack Obama and his putative pastor, Jeremiah Wright.

Obama denounced comments made by Wright at the NAACP and at the Press Club.

Wright responded, "It went down very simply. He's a politician. I'm a pastor. We speak to two different audiences. And he says what he has to say as a politician. I say what I have to say as a pastor. Those are two different worlds. I do what I do, he does what politicians do. So what happened in Philadelphia where he had to respond to the sound bites, he responded as a politician."

Similar strategizing takes place when candidates speak. Suppose that (certain positions on) the issues in question are represented by propositional variables p, q, r, s. Voter v would like p, q to be true and r, s, to be false. The candidate has only said $p \to q$ and $r \to s$. Many truth assignments are compatible with the candidate's theory T_c which is the logical closure of $\{p \to q, r \to s\}$. What should the voter think?

Similar strategizing takes place when candidates speak. Suppose that (certain positions on) the issues in question are represented by propositional variables p, q, r, s. Voter v would like p, q to be true and r, s, to be false. The candidate has only said $p \to q$ and $r \to s$. Many truth assignments are compatible with the candidate's theory T_c which is the logical closure of $\{p \to q, r \to s\}$. What should the voter think?

If v is optimistic, he will like the candidate, and if he is pessimistic, he will take the worst option. Or, he may take some sort of average.

Similar strategizing takes place when candidates speak. Suppose that (certain positions on) the issues in question are represented by propositional variables p, q, r, s. Voter v would like p, q to be true and r, s, to be false. The candidate has only said $p \rightarrow q$ and $r \rightarrow s$. Many truth assignments are compatible with the candidate's theory T_c which is the logical closure of $\{p \rightarrow q, r \rightarrow s\}$. What should the voter think?

If v is optimistic, he will like the candidate, and if he is pessimistic, he will take the worst option. Or, he may take some sort of average.

What should the candidate say?

Similar strategizing takes place when candidates speak. Suppose that (certain positions on) the issues in question are represented by propositional variables p, q, r, s. Voter v would like p, q to be true and r, s, to be false. The candidate has only said $p \to q$ and $r \to s$. Many truth assignments are compatible with the candidate's theory T_c which is the logical closure of $\{p \to q, r \to s\}$. What should the voter think?

If v is optimistic, he will like the candidate, and if he is pessimistic, he will take the worst option. Or, he may take some sort of average.

What should the candidate say? The candidate's problem is to make statements which she believes (or at least does not disbelieve) which will improve her image in the eyes of the (different groups of) voters.

NEW YORK After Sen. Barack Obama's comments last week about what he typically eats for dinner were criticized by Sen. Hillary Clinton as being offensive to both herself and the American voters, the number of acceptable phrases presidential candidates can now say is officially down to four. "At the beginning of 2007 there were 38 things candidates could mention in public that wouldn't be considered damaging to their campaigns, but now they are mostly limited to 'Thank you all for coming,' and 'God bless America,'" ABC News chief Washington correspondent George Stephanopoulos said on Sunday's episode of This Week.

The Onion,[2] *May 8, 2008*

---

[2] *The Onion* is a tongue-in-cheek weekly newsmagazine.

When a candidate utters a sentence A, she is evaluating its effect on several groups of voters, $G_1, ..., G_n$ with one group, say G_1 being her primary target *at the moment*.

Thus when Clinton speaks in Indiana, the Indiana voters are her primary target but she is well aware that other voters, perhaps in North Carolina, are eavesdropping. Her goal is to increase the likelihood that a particular group of voters will vote for her, but without undermining the support she enjoys or hopes to enjoy from other groups. If she can increase *their support* at the same time as wooing group G_1, so much the better, but at the very least, she does not want to undermine her support in G_2 while appealing to G_1. Nor does she want to be caught in a blatant contradiction. She may not always succeed, as we all know, but remaining consistent, or even truthful, is surely part of her strategy. Lies are expensive.

We will represent a particular group of voters as *one* formal voter, but since the groups are of different sizes, these formal voters will not all have the same influence. A formal voter who represents a larger group of actual voters will have a larger *size.* We will assume that each voter has a preferred ideal world – how that voter would like the world to be as a result of the candidate's policies, should she happen to be elected.

Thus suppose the main issues are represented by $\{p, q, r\}$, representing perhaps, policies on the Iraq war, abortion, and taxes. If the agent's ideal world is $\{p, q, \neg r\}$, then that means that the voter wants p, q to be true, and r to be false. But it may be that p is more important to the voter than q. Then the world $\{\neg p, q, \neg r\}$ which differs from the ideal world in just p will be worse for the voter than the one, $\{p, \neg q, \neg r\}$, which differs in just q.

We represent this situation by assigning a utility of 1 to the ideal world, and assigning weights to the various issues, adding up to at most 1. If the weights of p, q, r are .4, .2 and .4 respectively and the ideal world is $p, q, \neg r$, then a world in which p, q, r are all true will differ from the ideal world in just r. It will thus have a utility of (1 - .4), or .6.

Each voter also has a theory T_c of the candidate, and in the first pass we will assume that the theory is simply generated by things which the candidate has said in the past. If the candidate has uttered (presumably consistent) assertions $A_1, ..., A_5$, then T_c will be just the logical closure of $A_1, ..., A_5$. If the candidate is truthful, then T_c will be a subtheory of T_a which is the candidate's own theory of the world.

The voter will assume that if the candidate is elected, then one of the worlds which model T_c will come to pass. The voter's utility for the candidate will be obtained from the utilities of these worlds, perhaps by calculating the expected utility over the (finitely many) models of T_c.

(Note that we are implicitly assuming that all the worlds are equally likely, something which is not always true, but even such a simple setting turns out to be rich enough for some insights.)

Suppose now that the candidate (who knows all this) is wondering what to say next to some group of voters. She may utter some formula A, and the perceived theory T_c will change to $T'_c = T_c + A$ (the logical closure of T_c and A) if A is consistent with T_c, and $T_c * A$ if not. Here the $*$ represents an AGM like revision operator. (Note: The AGM operator $*$ accommodates the revision of a theory T by a formula A which is inconsistent with T. For the moment we will assume that A is in fact something which the candidate believes and is consistent with T_c which is a subtheory of T_a.)

Thus the candidate's utterance of A will change her perceived utility in the minds of the voters and her goal is to choose *that A* which will maximize her utility summed over all groups of voters. We can now calculate the utility to her of the utterance of a particular formula A.

Each group of voters will revise their theory of the candidate by including the formula A, and revising their utility evaluation of the candidate.

Let the old utility to group G_i calculated on the basis of T_c be U_i and the new utility calculated on the basis of $T_c * A$ be U'_i. Let w_i be the weight of the group G_i calculated on the basis of size, likelihood of listening to A which is greater for the current target group, and the propensity to actually vote. Then the change in utility on the basis of uttering A, or the *value* of A, will be

$$val(A) = val(A, T_c) = \Sigma w_i(U'_i - U_i)$$

Let the old utility to group G_i calculated on the basis of T_c be U_i and the new utility calculated on the basis of $T_c * A$ be U_i'. Let w_i be the weight of the group G_i calculated on the basis of size, likelihood of listening to A which is greater for the current target group, and the propensity to actually vote. Then the change in utility on the basis of uttering A, or the *value* of A, will be

$$val(A) = val(A, T_c) = \Sigma w_i(U_i' - U_i)$$

The rational candidate should utter *that* A which will have the largest value for $val(A)$.

Example 1: Quite recently, Hillary Clinton announced that she had shot a duck as a child. Now ducks do not vote, so we know she was not appealing to them. Who was she appealing to? Clearly those voters who oppose gun control. Other things being equal, a world in which there is gun control is worse for them than a world in which there isn't, and Hillary's remark will clearly decrease the set of worlds (in the voters' perception) in which Hillary is president and there *is* gun control. Presumably this will increase her utility in the minds of these voters.

But what about other voters who *do* prefer gun control? Now first of all, note that the fact that she shot a duck as a child does not *eliminate* worlds in which she is president and there is gun control – it merely decreases their number. Moreover, when she is campaigning in Pennsylvania or Indiana, these voters are not her primary voters. The likelihood that Massachusetts voters will be affected by the duck story will be (hopefully) less than the likelihood of a Pennsylvania voter being so affected. There is even the likelihood that voters who disfavor gun control – perhaps because they own a gun, will be more passionate about gun control (against it), than voters who favor gun control for more abstract reasons.

C will denote the candidate under consideration. V (or v as a subscript) will denote the group of voter (in the single block case). Otherwise $\mathcal{B} = \{B_1, \ldots, B_k\}$ will denote blocks of voters (this case will be considered later).

$T_c =$ voters' theory of candidate C

$T_a =$ candidate C's actual theory

$At = \{P_1, \ldots, P_n\}$ atomic propositions corresponding to issues (which we may identify with the integers $\{1, ..., n\}$).

$\mathcal{W}$ a finite set of worlds. Worlds will be seen as truth assignments, i.e., as functions $w : At \rightarrow \{1, -1\}$ such that $w(i) = 1$ if $w \models P_i$ and $w(i) = -1$ if $w \not\models P_i$ and we write $w(i)$ to denote the ith component of w. It may well happen that there is a non-trivial theory T_0 which is shared by both voters and candidates, and then of course the worlds to consider (even initially) will be those which model T_0.

$\mathcal{L} =$ the propositional language over At, which we may occasionally identify with the corresponding *propositions*, or subsets of $\mathcal{W}$.

$p_v : At \rightarrow \{1, 0, -1\} = V$'s preferred world, represented as follows

$$p_v(i) = \begin{cases} 1 & \text{if } V \text{ would prefer } P_i \text{ to be true} \\ 0 & V \text{ is neutral about } P_i \\ -1 & V \text{ would prefer that } P_i \text{ be false} \end{cases}$$

$w_v : At \rightarrow [0, 1]$ V assigns weight $w_v(i)$ to proposition i. To simplify thought, we assume $\sum_{1 \leq i \leq n} w_v(i) \leq 1$.
$u_v(w) = $ the utility of world w for V

$$u_v(w) = \sum_{1 \leq i \leq n} p_v(i) \cdot w_v(i) \cdot w(i)$$

Voter types: [o]ptimistic, [p]essimistic, [e]xpected value.
Given a possible set of worlds, according to the candidate's
position T_c so far, the optimistic voters will assume that the
candidate will implement the best one which is compatible with
T_c. The pessimistic voters will assume the worst, and the expected
value voters will average over the possible worlds.
$ut_v^t(T) =$ the utility of the theory T for V of type t (we leave out
the subscript v below).

- $ut^o(T) = \max\{u(w) : w \models T\}$
- $ut^p(T) = \min\{u(w) : w \models T\}$
- $ut^e(T) = \frac{\sum_{w \models T} u(w)}{|\{w : w \models T\}|}$

We could think, with slight abuse of language, of the ut functions as applying to sets of worlds rather than to theories, and if X, Y are sets of worlds, we will have,

- $ut^o(X \cup Y) = max(ut^o(X), ut^o(Y))$
- $ut^p(X \cup Y) = min(ut^p(X), ut^p(Y))$
- $ut^e(X \cup Y) \in$ the closed interval of $ut^e(X), ut^e(Y)$

The last claim about ut^e requires that X, Y be disjoint.

Note: There seems to be a plausible argument that typical voters would be of type p. Although such a voter always "assumes the worst" he is also such that hearing additional messages can never *decrease* the utility he assigns to his current theory of the candidate. As such, such a voter will always *prefer* to hear more information on which to base his vote. This seems like a rational strategy. Of course, a pessimistic voter can also be regarded as a 'play it safe', or 'worst outcome' voter.

Let $val(A, T)$ = the value of announcement $A \in \mathcal{L}$ be what a particular announcement A is worth to the candidate.

$$val(A, T) = ut(T \dotplus A) - ut(T)$$

What are the sets of formulas from which the candidate might choose? Possible sets $\mathcal{X} \subseteq \mathcal{L}$ of statements from which C might select to choose the message A she will utter:

- $\mathcal{X} = \mathcal{L}$ (this would allow for contradicting a statement already in T_c or lying about statements in T_a)
- $\mathcal{X} = T_a$ (only select a message from her actual theory)
- $\mathcal{X} = \mathcal{L} - \{\neg A : A \in T_c\}$ (allow for any message which is consistent with T_c)
- $\mathcal{X} = \mathcal{L} - \{\neg A : A \in T_a\}$ (allow for any message which is consistent with T_a)

An honest candidate will only choose a message which she actually believes, but a Machiavellian candidate may well choose a message which she does not believe, even disbelieves, but which is compatible with what she has said so far. But as we see, even an honest candidate has options.

$best(T, \mathcal{X}) =$ the most advantageous message for C which is an element of $\mathcal{X}$.

$$best(T, \mathcal{X}) = argmax_A val(A, T) : A \in \mathcal{X}$$

Proposition

Assume e-voters. For all $A \in \mathcal{X}$, there exist positive $a, \ldots, f$ such that

1. $a.val(A, T) + b.val(\neg A, T) = 0$
2. $val(A \wedge B)$
   $= val(A, T) + val(B, T \dotplus A)$
   $= val(B, T) + val(A, T \dotplus B)$
3. $c.val(A \vee B) + d.val(A \wedge B, T) = e.val(A, T) + f.val(B, T)$

Here the numbers $a, ..., f$ represent the number of worlds satisfying particular (relevant) theories. we immediately get,

Proposition

Assume e-voters. For all $A \in \mathcal{X}$, either exactly one of $val(A, T), val(\neg A, T)$ is positive and the other negative, or they are both zero.

Example: While it is clear that statements A and $\neg A$ cannot both benefit (or both hurt) a candidate, we *could* have a situation where $A \wedge B$ and $(\neg A) \wedge (\neg B)$ are both beneficial or both hurtful. Worlds which satisfy T fall into four groups:

Example: While it is clear that statements A and $\neg A$ cannot both benefit (or both hurt) a candidate, we *could* have a situation where $A \wedge B$ and $(\neg A) \wedge (\neg B)$ are both beneficial or both hurtful. Worlds which satisfy T fall into four groups:

1. **X:** worlds which satisfy both A, B.
2. **Y:** worlds which satisfy $A, \neg B$
3. **Z:** worlds which satisfy $\neg A, B$
4. **U:** worlds which satisfy $\neg A, \neg B$

Each of the sets **X,Y,Z,U** could have on average, better of worse worlds than the full set of wolrds which satisfy T_c. It can be that any of them are good and any of them are bad, provided that if one is good, at least one is bad, and if one is bad then at least one is good.

For instance, voters may feel that it is good to have a military buildup *and* go to war with Iran, or *not* have a buildup and not go to war, while thinking it foolish to have *just* one and *not* the other.

The utilities over the models of $T + A \wedge B$, $T + A \wedge \neg B$, $T + \neg A \wedge B$, $T + \neg A \wedge \neg B$, must average out to utilities over models of T. But any of them can be higher or lower, as long as they average out.

For optimistic voters, the best world could be in any of sets **X,Y,Z,U**, and the next best in any of them also.

If the voters are optimistic, the candidate is best off saying nothing. In other words, the best thing to say is something like, "These pancakes are nice."

If $T_c \subseteq T_d$, where c, d, are two candidates, then for all types of voters, the best possible extension of T_c is at least as good as the best possible extension of T_d. But for optimistic voters and e-voters, the best extension of T_c is likely to be strictly better.

The good world might stay when more information is received, or it can be deleted. So for optimistic voters, more information can only make things worse. For pessimistic voters, it is the other way.

Independent topics:

It is shown in [Parikh99] that given a propositional theory T in language L, L can be split into disjoint sublanguages $L_1, .., L_k$ such that there exist theories T_i in languages L_i with T being the logical closure of $\bigcup T_i$. Thus the theory T can seen as being obtained from unrelated theories T_i in languages L_i. Moreover, there is a finest splitting of T which refines every other splitting.

Proposition

Let the voters' perception of the candidate be expressed by theory T_c. Suppose that the language L of T_c can be split into two disjont sublanguages L_1, L_2 with theories T_1, T_2 in L_1, L_2 respectively which jointly imply T_c. Let $A \in L_1$ and $B \in L_2$ be possible statements that the candidate could make. Then $val(A, T) = val(A, T + B)$ and $val(B, T) = val(B, T + A)$. The statement A has the same effect whether it is made before or after statement B is made and similarly for B.

In other words, if L_1, L_2 are seen as unrelated then the candidate's statements in one language has no bearing on the statements in the other.