Social Choice for Disambiguation of Pronominal Reference

Abstract

Disambiguation of pronoun reference has been an important issue for both theoretical and computational linguists. While linguistic theories on binding conditions eliminate impossible readings to a certain extent, many inter-sentential anaphora remain ambiguous between bound variable reading and coreference with salient discourse entities, as in *John said he broke his leg* where *he* can refer either to *John* or someone else previously mentioned (Reinhart 1983). This paper considers pronoun resolution as a social choice among discourse participants (Arrow 1963).

1 Binding Conditions, Proximity and Saliency

Pronouns are known for their referential ambiguities. In a dialogue taken from a speech corpus in (1), the italicized pronoun *she* has multiple candidates for its antecedent—Emma, Lisa and Lisa’s mom. Proximity and saliency of antecedents have been considered to be key factors to decide (Leass 1991). In (1), the most proximate antecedent *her (Lisa)'s mom* is identified to be the antecedent for *she*.

(1) Frances: ...Not while Emma’s not here. You know Emma
   Billy: Mm.
   Frances: she’s, she was walking with Lisa and I weren’t there and her Mum sh–
       jus– , like *she* muc– , *she* mucks about a lot and she told Leigh that if he don’t
       serve her he’s gonna die, she’s gonna punch him right! Cos she’s quite big, you
       know....

   (BNC KSW435-441)

However, proximity does not always resolve referential ambiguity of pronouns.

(2) a. John, likes him\(_{\neq i/j\neq i}\).

b. John, said he\(_{i/j}\) likes himself\(_{i/j}\).

*Him* in (2a) unambiguously means someone other than the closest *John*—some discourse-salient entity, as indicated by the indices. When *he* is embedded under the
matrix clause as in (2b), the pronoun becomes ambiguous between John and someone else salient in the discourse.

Linguistic binding theory (Chomsky 1981, Reinhart 1983) explains the phenomena well. Antecedents are called binders, which bind bindees that are anaphoric pronouns, such as him or himself.

- Condition B: pronouns must be free in their local domain.

Pronouns are free when they are not bound by the antecedent by means of coindexed and c-commanding relation. C-command is roughly equivalent to precedence, with some restrictions.

However, binding conditions are not by themselves sufficient. Consider another example:

(3) Anna: Bill, is a good goalkeeper.
Kim: John, said he, broke his leg recently.

(3) is ambiguous in four ways and can have either one of the following interpretations:

4 a. John broke John’s leg
b. John broke Bill’s leg.
c. Bill broke Bill’s leg.
d. Bill broke John’s leg.

He and his can be bound by either John or another salient discourse entity Bill. The binding theories have no way of disambiguating these pronouns since there is no way of knowing speaker intention. Proximity does not predict the readings in (4b-d) either.

2 Social Choice Theory

2.1 Collective Decision Making

When multiple interpretations are available for pronouns, speaker’s intention decides the reading. Although Social Choice Theory (Arrow 1963, Moulin 1988, Taylor 2005, Gaertner 2009) has not yet been studied from linguistic perspective, being only briefly mentioned in van Rooij (2011) in relation with interadjective comparison, it enables theoretical incorporation of speaker’s intention to pronoun resolution. This section will overview social choice theory which has been studied in economics and political science and explains decision making by a group of people. When each member has different views and preferences, aggregation of every individual’s view results in a single view. Computational social choice theory (Chevaleyre et al. 2007, Endriss 2011) is a fairly new theory developed from the classical social choice theory.

Typically, social choice theory explains collective decision making in case of voting and has solved the problems with majority decision. Elections may follow the majority
rule, which ranks one candidate x above another candidate y if and only if a majority of the individuals do. When people vote for their preferred candidate according to their own ranking among the candidates, aggregation of people’s preferences helps in selecting a certain candidate. Note that “\( \alpha \succ \beta \)” denotes \( \alpha \) is preferred to \( \beta \). For example, the dominating ranking in the situation in (5) would be “Obama (o) \succ Clinton (c)” and “Clinton \succ Mcain (m),” that is, “Obama \succ Clinton \succ Mcain,” which is a preferred ranking by the majority if the preference relation is transitive. However, this conflicts with “Mcain \succ Obama,” which is also preferred by three people (Condorcet Paradox (de Condorcet 1785)).

(5) a. Anna (a): Obama \succ Clinton \succ Mcain  
   b. Kim (k): Clinton \succ Mcain \succ Obama  
   c. Heather (h): Obama \succ Clinton \succ Mcain  
   d. George (g): Mcain \succ Obama \succ Clinton  
   e. Nathan (n): Mcain \succ Clinton \succ Obama  
   f. Social preference: oRc \land cRm \land mRo

2.2 Arrow’s Social Welfare Function

Arrow’s axiomatic methods of social choice theory have tackled the above-mentioned problems of aggregation (Arrow 1963). Preferences are ordering between alternatives and should satisfy the following two axioms. When \( R \) stands for a knowledge of all pairs and \( x, y \) and \( z \) for alternatives,

**Axiom 1.** For all \( x \) and \( y \), either \( xRy \) or \( yRx \).

**Axiom 2.** For all \( x, y, \) and \( z \), \( xRy \) and \( yRz \) imply \( xRz \).

Axiom 1 states that the relation \( R \) is connected—every candidate is related to each other. Relations that satisfy Axiom 2 are transitive. In (6), \( N \), a finite set of individuals or voters, consists of five individuals and \( \chi \), a nonempty set of alternatives or candidates, has three members. Let \( L(\chi) \) denote the set of all linear orders on \( \chi \). A profile \( R \) is a vector of linear orders, or preferences. \( R_i \) is a vector of preferences of an individual \( i \). \( N_{x>y}^{R} \) denotes the set of individuals that prefers the candidate \( x \) to \( y \). Supposing \( R \) the profile given in this model, \( N_{o>c}^{R} \) is a set of people who prefers Obama to Clinton, that are, Anna, Heather and George.

(6) a. \( N = \{a, k, h, g, n\} \)  
   b. \( \chi = \{o, c, m\} \)  
   c. \( R = (R_a, R_k, R_h, ...) \in L(\chi)^N \)  
   d. \( N_{o>c}^{R} = \{a, h, g\} \)
Individual’s preferences are aggregated and returns a single preference order, that results in collective decisions. A social welfare function (SWF) $F$ is a function which takes individual’s preferences and returns collective preference which is supposed to represent people.

$$\text{(7) SWF } F: \mathcal{L}(\chi)^{|N|} \rightarrow \mathcal{L}(\chi)$$

Arrow demonstrated that any SWF for three or more alternatives that satisfies the Pareto condition and IIA must be a dictatorship.

**Theorem 1 (General Possibility Theorem).** If there are at least three alternatives which the members of the society are free to order in any way, then every social welfare function satisfying Conditions 1 and 2 and yielding a social ordering satisfying Axioms 1 and 2 must be either imposed or dictatorial.

**Condition 1 (Pareto condition).** A SWF $F$ satisfies the Pareto condition if, whenever all individuals rank $x$ above $y$, then so does society: $N^{R}_{x>y} = N$ implies $xF(R)y$

Condition 2 states that the relative ranking of two candidates remains unchanged no matter how other candidates are ranked.

**Condition 2 (Independence of irrelevant alternatives (IIA)).** A SWF $F$ satisfies (IIA) if the relative social ranking of two alternatives only depends on their relative individual rankings: $N^{R}_{x>y} = N^{R'}_{x>y}$ implies $xF(R)y \iff xF(R')y$

When an individual’s preference dictates the collective preference, that individual is called a dictator. Dictatorship is a SWF that maps any member’s profile to a single individual profile.

3 Application of Social Choice Theory to Pronoun Resolution

3.1 Anaphora Resolution as Social Choice

This section will show that social welfare function for pronoun resolution satisfies Arrow’s General Possibility Theorem by satisfying Axioms 1, 2, Pareto Condition and IIA but demonstrating dictatorship.

Since the referents of pronouns can be ambiguous as discussed in section 1, pronoun resolution can be compared with voting by multiple voters—in this case, discourse participants. The candidates or choices would be different interpretation of the sentence.

For example, the first sentence in the following dialogue in (8) is ambiguous between two interpretations: (i) the pronoun him refers to John, or (ii) him means someone else, Bob, who is salient in the discourse.

$$\text{(8) Chris: John likes him. I mean, John likes Bob. Naomi: I thought you meant John liked himself.}$$
Such ambiguity actually does not exist at all since the first reading is impossible, as Binding Condition B properly eliminates the first reading. \textit{Him} should be free in the locality.

In the following dialogues (9) and (10), the referent of the pronouns \textit{he} in both (9) and (10) is ambiguous between \textit{John}, the binder, and some other discourse referent. Suppose that the speaker meant the referent of \textit{he} to be \textit{Bob} who appeared in their previous discourse, while the hearer interpreted \textit{him} to be \textit{John}.

(9) Chris: John said he broke his leg.
Naomi: Did he? John looked fine when I saw him this morning.
Chris: It is Bob who broke his leg.
Naomi: I thought you were talking about John.

(10) Naomi: I saw Bob’s car dented.
Chris: John said he met with an accident.
Naomi: John, too?
Chris: I mean, it was Bob who met an accident.
Naomi: I thought you said John was hit.

\textit{He} in the embedded clause has been considered as a bound variable bound by \textit{John} (Heim & Kratzer 1998). Since both variable binding and coreference with salient discourse entities are available, ambiguities remain. Binding Conditions alone do not eliminate such ambiguity.

Therefore, I argue that identifying the antecedent of pronouns is a social choice and SWF decides the antecedent.

(11) a. Individuals I = \{c, n\}
b. Candidates \chi = \{j, b\}
c. Ordering jRb \land bRj

d. Denote the set of linear orders on \chi by L(\chi). Preferences (or ballots) are taken to be elements of L(\chi).

e. A profile R = (R_c, R_n) \in L(\chi)^I\{I\} is a vector of preferences

f. A social choice function (SCF) or voting rule is a function F : L(\chi)^I\{I\} \rightarrow 2^x \setminus \{\emptyset\} mapping any given profile to a nonempty set of winners.

g. A social welfare function (SWF) is a function F : L(\chi)^I\{I\} \rightarrow L(\chi) mapping any given profile to a (single) collective preference order.

h. R_c = F(R_c) = F(R_n) = R

In dialogues (9) and (10), there are two voters, Chris and Naomi who are discourse participants. The decision is made regarding the referent of the pronoun \textit{he}. The preferences between the candidates vary between the individuals. The social welfare function returns a single preference order for the references since ambiguities are resolved during the conversation.
3.2 Axioms 1 and 2

There are three possible antecedents for *she* in (1)—Emma, Lisa and Lisa’s mother. Let us say that Billy (b) prefers Emma (e) to Lisa (l), and also Lisa to Lisa’s mother (m) to be the antecedent. On the other hand, the speaker Francis (f) prefers Lisa’s mother to the other two, Lisa’s mother to Lisa, and Lisa to Emma according to the proximity. All three candidates are ordered in accordance with Axiom 1.

\[
\begin{align*}
(12) & \quad a. \ eR_b l \land lR_b m \\
& \quad b. \ mR_f l \land lR_f e
\end{align*}
\]

Transitivity also holds for pronoun antecedent preferences. (12a) and (12b) each implies (13a) and (13b).

\[
\begin{align*}
(13) & \quad a. \ eR_b lR_b m \\
& \quad b. \ mR_f lR_f e
\end{align*}
\]

3.3 Pareto Condition

Pareto condition states if, whenever all individuals rank x above y, then so does the society. Suppose the situation when Billy, Francis and another discourse participant Heather, who stayed quiet during the conversation, all prefer the same ranking in (14).

\[
\begin{align*}
(14) & \quad mR_b l \land lR_b e \land mR_f l \land lR_f e \land mR_h l \land lR_h e \rightarrow mR_l \land lR_e
\end{align*}
\]

Therefore, \( N_{m,l,e} \times N \) implies \( mF(R)lF(R)e \)

When everyone understands that Francis referred to Lisa’s mom with *her*, the society which consists of discourse participants also refers Lisa’s mom to be the antecedent. Therefore, Pareto condition is met.

By means of reduction, I will show that the assumption that, all individuals rank x above y but society does not rank x above y, leads to contradiction.

**Proof.** Let us prove the contrary does not hold.

Assume, \( N_{m,l,e} \times N \) implies \( yF(R)x \) —(A)

Suppose,

\( jR_e b \land jR_ab \rightarrow jRb \)

Therefore, (A) does not hold.

QED.
3.4 IIA

A SWF $F$ satisfies IIA if the relative social ranking of two alternatives only depends on their relative individual rankings. Suppose the dialogue in (9) is modified into (15) and (16). The preference relations are denoted by $R$ for (15) and $R'$ for (16). As shown in (17), the relative rankings between Bob and John remain unaffected by irrelevant candidate Victor’s ranking; therefore IIA is satisfied.

(15) Chris: Bob is a good skier. But John said he broke his leg.
    Naomi: Did he? Poor Bob!

(16) Chris: Victor is a good skier and so is Bob. But John said he broke his leg.
    Naomi: Did he? Poor Bob!

(17) $bR'c \land vR'j \land bR'nj \land jR'nv$
    Then, $N_{b>j}^R = N_{b>j}^{R'}$ implies $bF(R)j \iff bF(R')j$

Below, by means of reduction, I will show that it is contradictory to say that $N_{x>y}^R = N_{x>y}^{R'}$ implies $xF(R)y \iff xF(R')y$.

Proof. Let us assume the following:
$bRcj \land bRnj \land bR'cv \land vR'ej \land bR'nj \land jR'nv$
Then, $N_{b>j}^R = N_{b>j}^{R'}$
Suppose,
$bF(R)j \iff bF(R')j$ —(B)
However,
$bF(R)j \iff bF(R')j$
Therefore, (B) is contradictory.
QED.

3.5 Dictatorship

Corrections made by the speaker in dialogues suggest that the speaker dictates the decision. The different preference rankings are aggregated to a single choice made by the speaker. The hearers are forced to interpret the speaker utterance as he meant.

What is called the plurality rule does not apply to pronoun resolution. The plurality rule is one to elect the candidate ranked first most often. Although voting usually satisfies the plurality rule, the selection of possible interpretation of pronouns does not depend on the sum of the discourse participants who share the same interpretation. Rather, pronoun resolution is dominated, or dictated, by the speaker’s meaning.

(18) Dictator: speaker $c \in I$

The speaker dominates the choice of referents and reserves right to correct the hearer’s interpretation.
$F(R) = R_c$ for any profile $R$, that is, the outcome is always identical to the preference supplied by the dictator.
Proof. Suppose
\[ F(R) \neq R_c \text{ where } c \text{ is a speaker} \]
Suppose a following dialogue:

(19) 1) Chris: Bob is a good skier. But John said he broke his leg.
    2) Naomi: Did he? Poor Bob!
    3) Chris: No. I mean John broke his leg.

\[ F(R) = R_n \text{ until the line 2. However, } F(R) = R_c \text{ after all at line 3. Therefore, the} \]
premise leads to contradiction.
QED.

Lemma 1. The choice of antecedent for anaphora is a social welfare function which
satisfies conditions 1 and 2.

4 Conclusion

This paper is one of the few attempts to apply social choice theory to linguistic problems. While
anaphora resolution has been a problem for natural language processing and theoretical linguistics, it is considered to be a social choice function dictated by a dictator, the speaker.

References

Haven, 2 edn.


N. Chomsky (1981). Lectures on Government and Binding. Foris Publications, Dor-
drecht.


tives in Economic Analysis. Oxford University Press.


