Comma in Japanese is Conjunction

Sumiyo Nishiguchi
School of Management, Tokyo University of Science
500 Shimo-kiyoku, Kuki-city, Saitama, 346–8512, Japan
nishiguchi@rs.tus.ac.jp

June 14, 2012
JSAI2012@Yamaguchi, Japan
Abstract

- Logical conjunction and disjunction of entities are expressed in the form of “A, B and/or C” in English language.

- The word sense disambiguation of the comma “,” is simple—just copy the following connective and or or.

- In other words, “A, B and/or C” is translated into “A and/or B and/or C” where the meaning of the comma is identical to the following connective.

- Nevertheless, such method does not work for Japanese language in which, mere sequences of commas often denote conjunction or disjunction as in “A, B, C.”

- This paper claims that conjunctive commas between noun phrases always represent conjunction in Japanese; therefore, “,” can be translated to be “to (and),” and that disjunctive reading is obtained through the presence of distributive operators and world knowledge.
Roadmap

Word Sense Disambiguation of Commas in English

Commas in Japanese

Comma is Conjunction
  Logical Necessity: Double Negation
  Distributive Operator
    Distributive “Each”
    World Knowledge

Conclusion
In English, sequences of noun phrases (NP) are connected by commas and the connective *and* or *or*, as expressed in Regular Expression:

\[(1) \ (NP,)^+ \text{ and NP}\]

Even though commas between NPs are semantically ambiguous in two ways between *and* and *or* in English, the word sense disambiguation of “,” is an easy task because the meaning of the comma inherits the one of the following connective.

**Context-free Grammar:**

\[(2) \ a. \ NP, \ NP \text{ and NP} \rightarrow \ NP \text{ and NP and NP} \]

\[ b. \ NP, \ NP \text{ or NP} \rightarrow \ NP \text{ or NP or NP} \]
Copying And

While “,” is polysemous between and and or by itself, the meaning of “,” following Boroughs is disambiguated by copying and which follows Health Authorities. (3) is equivalent with (4).

(3) ACET News aims to—inform Boroughs, Health Authorities, and voluntary agencies of the work of ACET in the community.

(BNC A00 290)

(4) ACET News aims to—inform Boroughs and Health Authorities, and voluntary agencies of the work of ACET in the community.
Intersective Conjunction

Ignoring the modal expression *aims to*, the semantic contribution of the commas is to provide individual sum of *Boroughs, Health Authorities* and *voluntary agencies of the work of ACET*.

(5) \[ (4) \]
\[ = \text{inform}(b \oplus h \oplus v)(a) \]
\[ \iff \text{inform}(b)(a) \land \text{inform}(h)(a) \land \text{inform}(v)(a) \]
(a: ACET News, b: Boroughs, h: Health Authorities, v: voluntary agencies of the work of ACET in the community, \( \alpha \oplus \beta \): individual sum of \( \alpha \) and \( \beta \) à la Link (1993))
Intersective or Collective Conjunction

Hoeksema (1983) calls such noun phrase (NP) conjunction as intersective conjunction (cf. Schwarzschild 2010).

(6) \text{and}_i; \lambda \pi \lambda \phi \lambda P[\phi(P) \land \pi(P)]
- Every professor and every student ran. ⇔
- Every professor ran and every student ran.

(7) \text{and}_c; \lambda \pi \lambda \phi \lambda P[\phi(\lambda x. \pi(\lambda y. P(Grp[s, y])))]
- Some professors and\text{and}_c some students met. ⇔
- Some professors met and some students met.
Copying *Or*

On the other hand, commas before disjunctive *or* is interpreted as disjunctions. (7) is understood to be disjunction of five propositions, namely, *you live in Canterbury or you live in Brighton or you live in Farnman or you live in Southampton or you live in Portsmouth*.

(8) If you live in Canterbury, Brighton, Farnham, Southampton or Portsmouth, the exhibition will be coming your way.

\[(BNC\ KS5\ 31)\]

(9) \[
\begin{align*}
\llbracket(8)\rrbracket & = \text{live}(c \oplus b \oplus f \oplus s \oplus p)(h) \to \text{see}(e)(h) \\
\iff & \text{live}(c)(h) \lor \text{live}(b)(h) \lor \text{live}(f)(h) \lor \text{live}(s)(h) \lor \text{live}(p)(h) \\
\to & \text{see}(e)(h)
\end{align*}
\]

(h: hearer, c: Canterbury, b: Brighton, f: Farnham, s: Southampton, p: Portsmouth, e: exhibition)
Commas in Japanese

Commas representing conjunction

While disambiguation of internominal commas in English is fairly easy, the one in Japanese appears to be more complicated because mere sequences of commas represent conjunction or disjunction.

(10)

Indo, Pakisutan, Kitachosen-ga mishomei-deari
India Pakistan North Korea-NOM unsigned-be
"India, Pakistan and North Korea remain unsigned."

(BCCWJ2011, OW6X_00016)

(10) is synonymous with (11)

(11)

Indo-to Pakisutan-to Kitachosen-ga mishomei-deari
India-and Pakistan-and North Korea-NOM unsigned-be
"India, Pakistan and North Korean remain unsigned."

(12) \[(11)\] = \neg\text{sign}(x)(i \oplus p \oplus n) \iff \neg\text{sign}(x)(i) \land \neg\text{sign}(x)(p)
\land \neg\text{sign}(x)(n)
(13) Keikikaifukuki-ni-wa endaka, economic recovery period-DAT-TOP strong yen kabudaka, kinridaka-to nari strong stock prices high interest rate-COMP become “Strong yen, strong stock prices and high interest rate occur with economic recovery” (BCCWJ2011, PB53_00349)

(14) Keikikaifukuki-ni-wa endaka-to economic recovery period-DAT-TOP strong yen kabudaka-to kinridaka-to nari strong stock prices high interest rate-COMP become “Strong yen, strong stock prices and high interest rate occur with economic recovery”

(15) $[(14)] = k \rightarrow y \oplus s \oplus i$
$\Leftrightarrow [k \rightarrow y] \land [k \rightarrow s] \land [k \rightarrow i]$

(k: economic recovery, y: strong yen, s: strong stock prices, i: high interest rate)
(16) Denki, gasu, suido-ga tot-teiru.
electricity gas water-NON run-PROG
“There are electricity, gas and water supplies.”

(17) a. Denki-to gasu-to suido-ga tot-teiru.
electricity-and gas-and water-NOM run-PROG
“There are electricity, gas and water supplies.”

b. Denki-ka gasu-ka suido-ga tot-teiru.
electricity-or gas-or water-NOM run-PROG
“There is either electricity, gas or water supplies.”

(18) \[(16)\] = run(e \oplus g \oplus w)
\iff run(e) \land run(g) \land run(w)
(e: electricity, g: gas, w: water)

“Chiba University started admission of high-school students before graduation and Meiji University, Showa Women’s University, Seijo University, Elisabeth Music College, and Aizu University followed.”

(Asahi Newspaper, June 2, page 1, Tobinyugaku)
Commas representing disjunction

(20) is semantically equivalent with (21) in which commas are replaced with *ka* or *matawa* “or.”

(20) Kimuchi, tsukemono, sato, shio, Korean pickles Japanese pickles sugar salt
miso-nanka-o ireru-noni pittaride omiyage-ni-mo soybean paste-etc-ACC put-for fit souvenir-DAT-also
daikohyo popular
“(The pot) fits for keeping Korean and Japanese pickles, sugar, soybean paste and others and popular for souvenirs.”

(BCCWJ2011, LBq2_00068)
(21) Kimuchi-\textbf{ka}/matawa/*to tsukemono-\textbf{ka}/*to sato-\textbf{ka}
Korean pickles-or/or/and Japanese pickles-or/and sugar-or
shio-\textbf{ka} miso-nanka-o ireru-noni pittaride
salt-or soybean paste-etc-\text{ACC} put-for fit
omiyage-ni-mo daikohyo
souvenir-\text{DAT}-also popular
“(The pot) fits for keeping Korean or Japanese pickles, sugar, soybean paste or others and popular for souvenirs.”

(22) $[(21)] = \Diamond \text{put(p)(k } \oplus j \oplus s \oplus sa \oplus m)\quad \Leftrightarrow \Diamond [\text{put(p)(k } \lor \text{put(p)(j } \lor \text{put(p)(s } \lor \text{put(p)(sa } \lor \text{put(p)(m} \vert] \quad \Diamond$
(23) Okina rieki-o toritake-reba, Makao, Indo, Betonamu, large profit-ACC gain-if Macau India Vietnam Hawaioto-ya mikaitakuna basho-ga the Island of Hawaii-or unexploited location-NOM kakujitsuda-ga certain-but “If you want to gain large profit, Macau, India, Vietnam or the Island of Hawaii, or other unexploited locations would be surer.”

(24) Okina rieki-o toritake-reba, Makao-ya/matawa/*to large profit-ACC gain-if Macau-or/or/and Indo-matawa Betonamu-ya Hawaioto-ya India-or Vietnam the Island of Hawaii-or mikaitakuna basho-ga kakujitsuda-ga “If you want to unexploited location-NOM certain-but gain large profit, Macau, India, Vietnam or the Island of Hawaii, or other unexploited locations would be surer.”
(25) Ato 2, 3-nichi-wa koko-ni todomat-teor-anebanarumai.  
rest 2 3-day-TOP here-LOC stay-PROG-should  
“Should stay here at least a few days longer.”  

(BCCWJ 2011 OB2X_00056)
I have abstracted 200 examples out of 58689 instances in the form “NP, NP,” in BCCWJ2011 by using Chunagon. The commas between NPs were annotated with either *and*, *or*, reduplication or punctuation.

<table>
<thead>
<tr>
<th>translation of commas</th>
<th>instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>and</td>
<td>174</td>
</tr>
<tr>
<td>or</td>
<td>18</td>
</tr>
<tr>
<td>reduplication</td>
<td>6</td>
</tr>
<tr>
<td>punctuation</td>
<td>2</td>
</tr>
<tr>
<td>total</td>
<td>200</td>
</tr>
</tbody>
</table>
Reduplication

Reduplication is the repetition of the same words. (26) is an example of the reduplication of onomatopoeic words.

(26) Aburazemi-de-wa jij, jij, ji, nigrofuscata cicada-LOC-TOP ONOM ONOM ONOM
ezozemi-de-wa geeg, gee, to izuremo Lyristes japonicus-LOC-TOP ONOM ONOM COMP both hakkiri kugiri-masu.
clearly punctuate-HON

“Both Nigrofuscata and Lyristes japonicus cicadas clearly punctuate while singing—“jij, jij, ji,” and “geeg, geeg.”

(BCCWJ2011, LBf4_00023)

Excluding commas used for reduplication and punctuation, this paper claims that Japanese comma should be translated as conjunction from logical perspective. Disjunctive reading arises only with the presence of quantifiers and world knowledge.
Comma is Conjunction

This section argues that internominal commas monosemously represent conjunction in view of (i) its logical contribution, and (ii) the presence of distributive keys.
Logical Necessity: Double Negation

The examples with double negation support the claim that Japanese commas represent conjunction. A conjunct of two nominals is embedded under two negations. The direction given by the sentence (27) is that everyone should both write and mark the form. Therefore, the sentence should not be translated into (28b) but (28a)

(27)

Kinju, makushitei-nai-hito-wa ima-sen-ne.
write mark-NEG-person-TOP be-NEG-PAR
“There is not anyone who has not written or marked, is there?”

(28) a. \( \neg \exists x [ \neg (\neg \text{write}(x) \land \text{mark}(x))] \)
\( \Leftrightarrow \neg \exists x [\neg \text{write}(x) \lor \neg \text{mark}(x)] \)
\( \Leftrightarrow \forall x [\neg \text{write}(x) \lor \neg \text{mark}(x)] \)
\( \Leftrightarrow \forall x [\text{write}(x) \land \text{mark}(x)] \)

b. \( \neg \exists x [\neg (\neg \text{write}(x) \lor \text{mark}(x))] \)
\( \Leftrightarrow \neg \exists x [\neg \text{write}(x) \land \neg \text{mark}(x)] \)
\( \Leftrightarrow \forall x [\neg \text{write}(x) \land \neg \text{mark}(x)] \)
\( \Leftrightarrow \forall x [\text{write}(x) \lor \text{mark}(x)] \)
The presence of distributive keys forces disjunctive reading of commas. Otherwise, commas represent conjunction. Distributive keys can be either overt quantifiers or world knowledge.
Distributive “Each”

Some examples contain a quantifier _dono_mo_ “each/any” which associates with a distributive operator (Dowty & Brody 1984). A quantifier forces distributive reading. In (29), the events denoted by _taikan_ “retirement,” _shutsuba_ “running for election,” and _kiso_ “indictment” are temporally distributed into separate events by the distributive operator so that the commas connecting these event nouns substitute for the disjunction.

(29) Tatematsu-wa Kishimoto-no taikan, shutsuba,
    Tatematsu-TOP Kishimoto-GEN retire running for election
    kiso-no _dono_ fushime-de-mo, watashi-ni
    indictment-GEN _any_ turning point-at-even me-DAT
    hitokoto-no kanso-mo morasa-nakat-ta.
    one word-GEN opinion-even share-GEN-PAST

“Tatematashu never shared me his opinions at any turning point of Kishimoto—his retirement, running for election and indictment.”
Eliminating Negation

(30) Tatematsu-wa Kishimoto-no taikan, shutsuba, Tatematsu-TOP Kishimoto-GEN retire running for election kiso-no dono fushime-mo, shijishi-ta. indictment-GEN any turning point-even support-PAST

“Tatematsu supported any turning point of Kishimoto—his retirement, running for election and indictment.”

(31) \[ (30) \equiv \text{supported}(r \oplus e \oplus i)(t) \equiv \text{supported}(r)(t) \land \text{supported}(e)(t) \land \text{supported}(i)(t) \]
World Knowledge

Eliminating the quantifier *dono-mo* “each/any” as in (32) allows cooccurrence of the three events—retirement, running for election and indictment, which is hard in this example due to our world knowledge that these three events usually do not occur at one time.

(32) Tatematsu-wa Kishimoto-no taikan, shutsuba, Tatematsu-TOP Kishimoto-GEN retire running for election kiso-no fushime-de, watashi-ni hitokoto-no indictment-GEN turning point-at me-DAT one word-GEN kanso-mo morasa-nakat-ta.

“Tatematasu never shared me his opinions at any turning point of Kishimoto—his retirement, running for election and indictment.”
World Knowledge

(33) (Kame-wa) Kimuchi, tsukemono, sato, shio, jar-TOP Korean pickles Japanese pickles sugar salt miso-nanka-o ireru-noni pittaride omiyage-ni-mo soybean paste-etc-ACC put-for fit souvenir-DAT-also daikohyo popular
“(The jar) fits for keeping Korean and Japanese pickles, sugar, soybean paste and others and popular for souvenirs.”

(BCCWJ2011, LBq2_00068)

World knowledge that we do not usually mix pickles, sugar, salt and soybean paste into one jar forces disjunctive reading. There cannot be only jar for everything.

(34) [(33)] = ◊put(p)(k ⃝ j ⃝ s ⃝ sa ⃝ m)
    ⇔ ◊[put(p)(k) ∨ put(p)(j) ∨ put(p)(s) ∨ put(p)(sa) ∨ put(p)(m)]
(35) Shakaiteki-mibun, monchi, jinshu, minzoku, shinjo, social status, family origin, race, tribe, religion, seibetsu, shogai-to-nyoroo, hutona, sabetsu, nado, gender, disabilities-etc.-by unjust, discrimination etc.

samazamana, katachide-no, jinkenshingaikoi-ga, aru. various, form-{GEN}, violation of human rights, exist

“There exist various forms of violation of human rights such as discrimination based on social and family status, race, tribe, religion, gender, and disabilities.”

(BCCWJ2011, LBs6_00031)
World Knowledge Forces Disjunctive Reading

(36) Shakaiteki-mibun-matawa/*to monchi-matawa/*to
social status-or/and family origin
jinshu-matawa/*to minzoku, shinjo, seibetsu,
race tribe religion gender
shogai-to-niyoru hutona sabetsu nado, samazamana
disabilities/etc.-by unjust discrimination etc. various
katachide-no jinkenshingaikoi-ga aru.
form-GEN violation of human rights exist

World knowledge that a single person would not usually undergo
discrimination, based on social status, family origin, race, tribe,
religion, gender and disabilities at the same time, forces disjunctive
reading.

(37) \( \forall x \exists y [x \in X \land [X = \text{social status} \oplus \text{family origin} \oplus \text{race} \oplus \text{tribe} \oplus \text{religion} \oplus \text{gender} \oplus \text{disabilities}] \rightarrow \text{base}(x)(y) \land \text{discrimination}(y)] \)
Conclusion

- This paper surveyed occurrences of inter-nominal commas and represented their meaning as conjunction.
- The disjunctive reading is forced by the presence of distributive keys or world knowledge.
- The conjunctive interpretation is also required from logical standpoint.
References

British National Corpus (BNCweb), CQP edition.
Chunagon, https://chunagon.ninjal.ac.jp/search