

Enjoyable Conversation System

Tetsunori Kobayashi

Waseda University

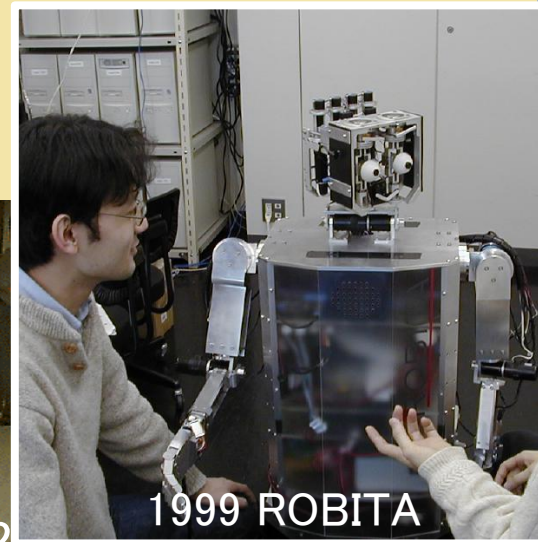
2016.7.15



1973 WABOT-1
The First Conv. RT



1995 Jijo2



1999 ROBITA

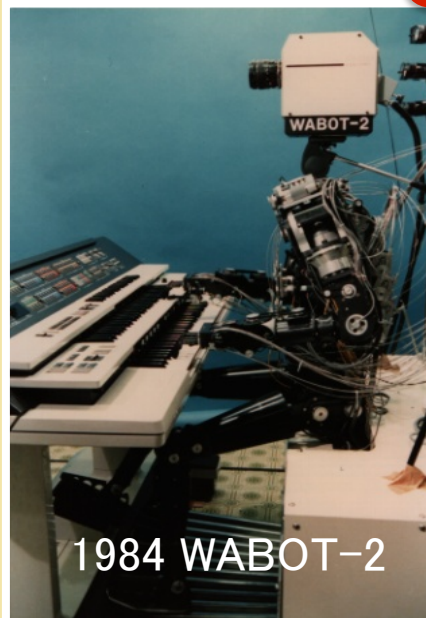
Group Convers

History of CONV. RT

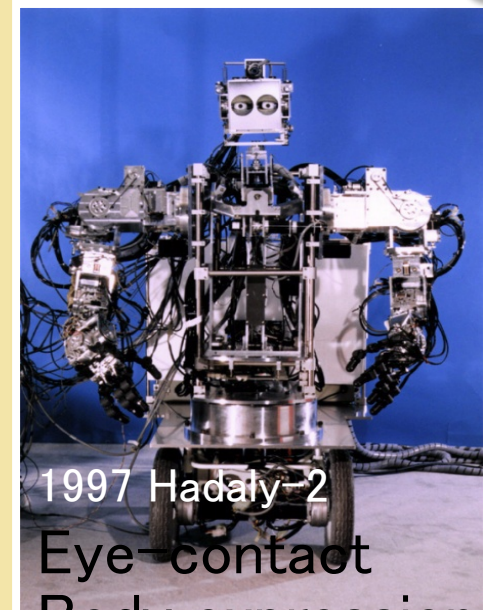
1970

1980

1990



1984 WABOT-2

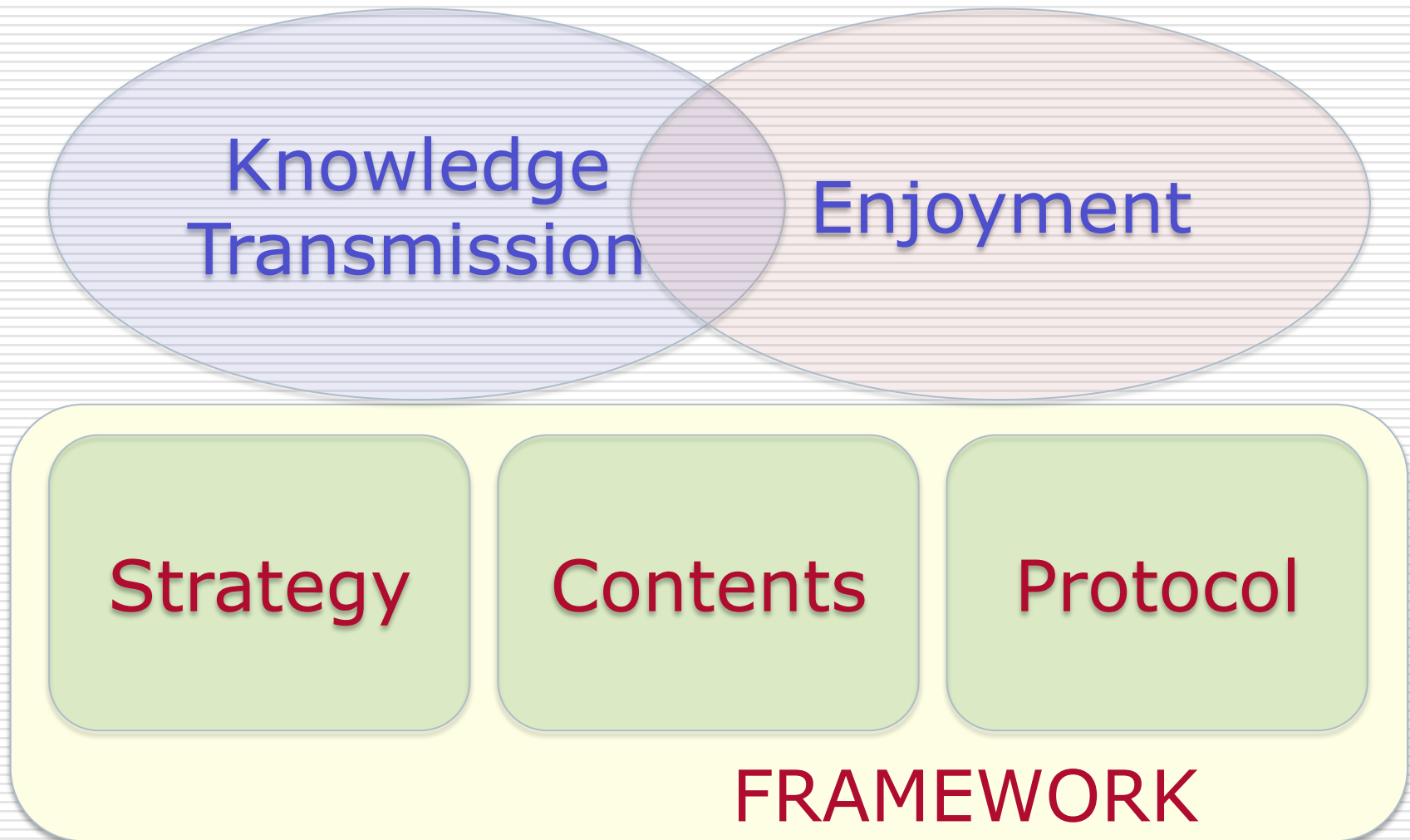


1997 Hadaly-2

Eye-contact
Body expression²

1998 R

My WORKs



from experience at
daycare for elderly people

- 2006-2009 at CareTown KODAIRA

I'd like to talk about :

1. Be Talkative:

- Framework to add one more comment to the usual answer.

2. Be Attentive:

- Framework to balance the engagement density of group communication.

難

「鰻背」

難読

Robot Vision

Subjects

EXTERNAL TRIGGER: "TSUGI"
BEHAVIOR: REACT TO LEVEL

2006-09 @Care Town KODAIRA

What elderly people enjoyed

1. Unexpected/surprising utterances

Oh! Robot speaks such things !!

2. Addressing to him/her

Oh! Robot talked to me !!

Add **ONE MORE** comment
Unexpected/surprising utterances

Enjoyable system should be
Talkative !!

Chaining of Episodes

Tedious guy

replies to only what he was asked.

Proposal: Add "+a" comment

Q: Which actor/actress do you like best, Robisuke?

A: I like Audrey Hepburn, best.

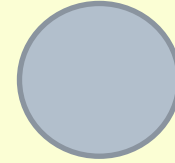
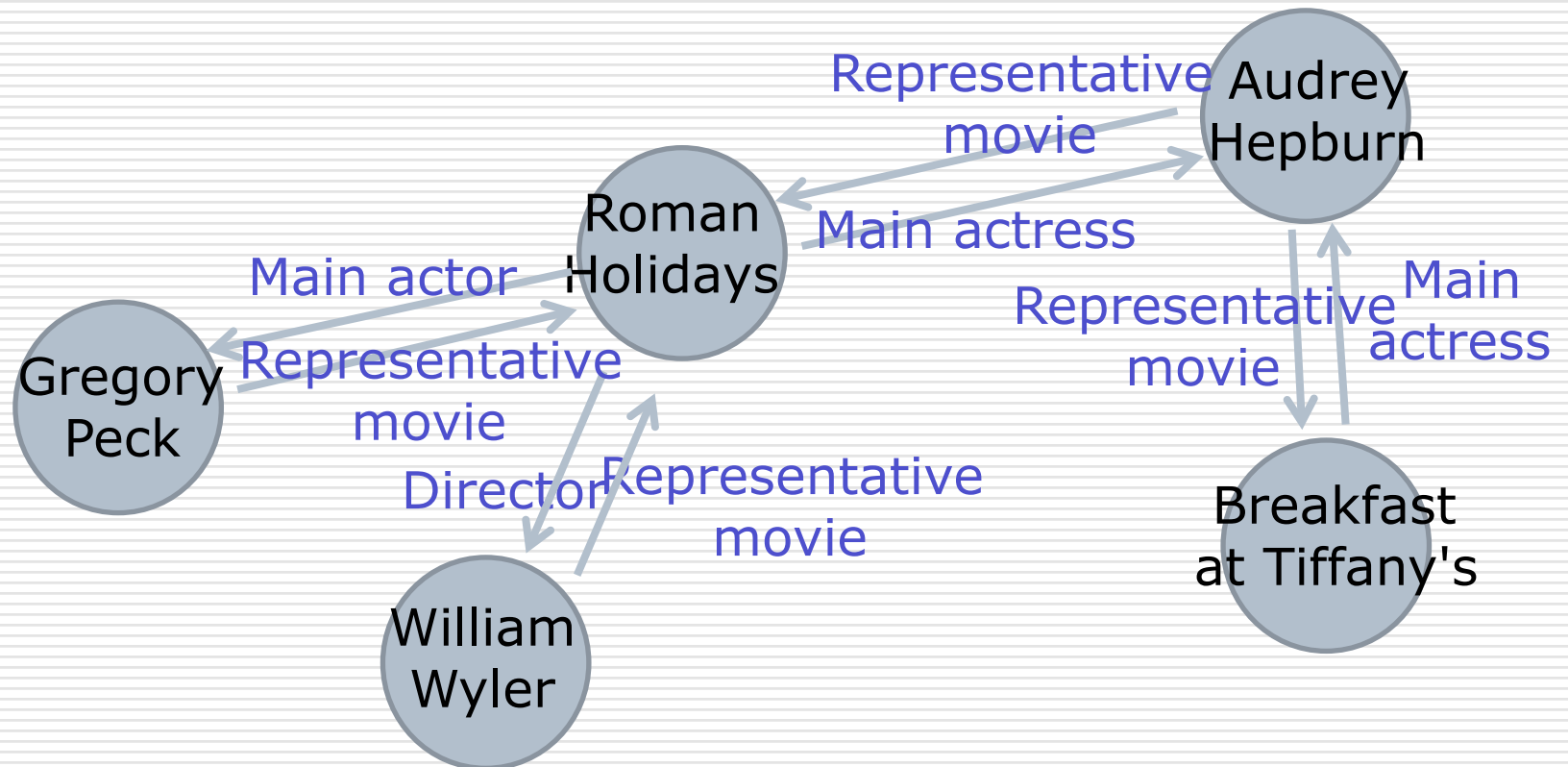
Roman Holiday is my favorite. (+a)

I love the scene of "Mouth of Truth" (+a)

Liar will lose his hand when he XXXXX (+a)

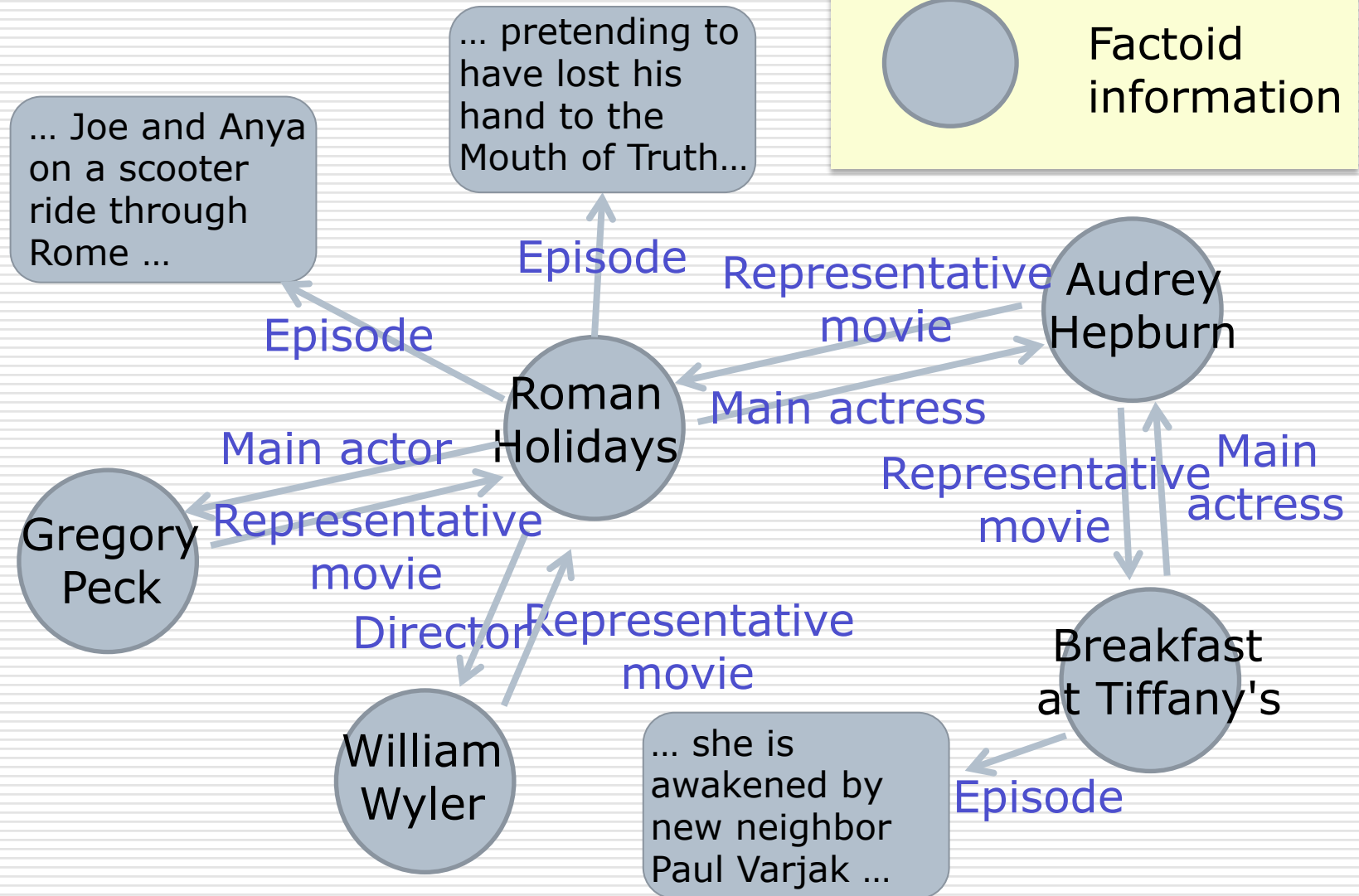
....

Chaining Episodes



Factoid
information

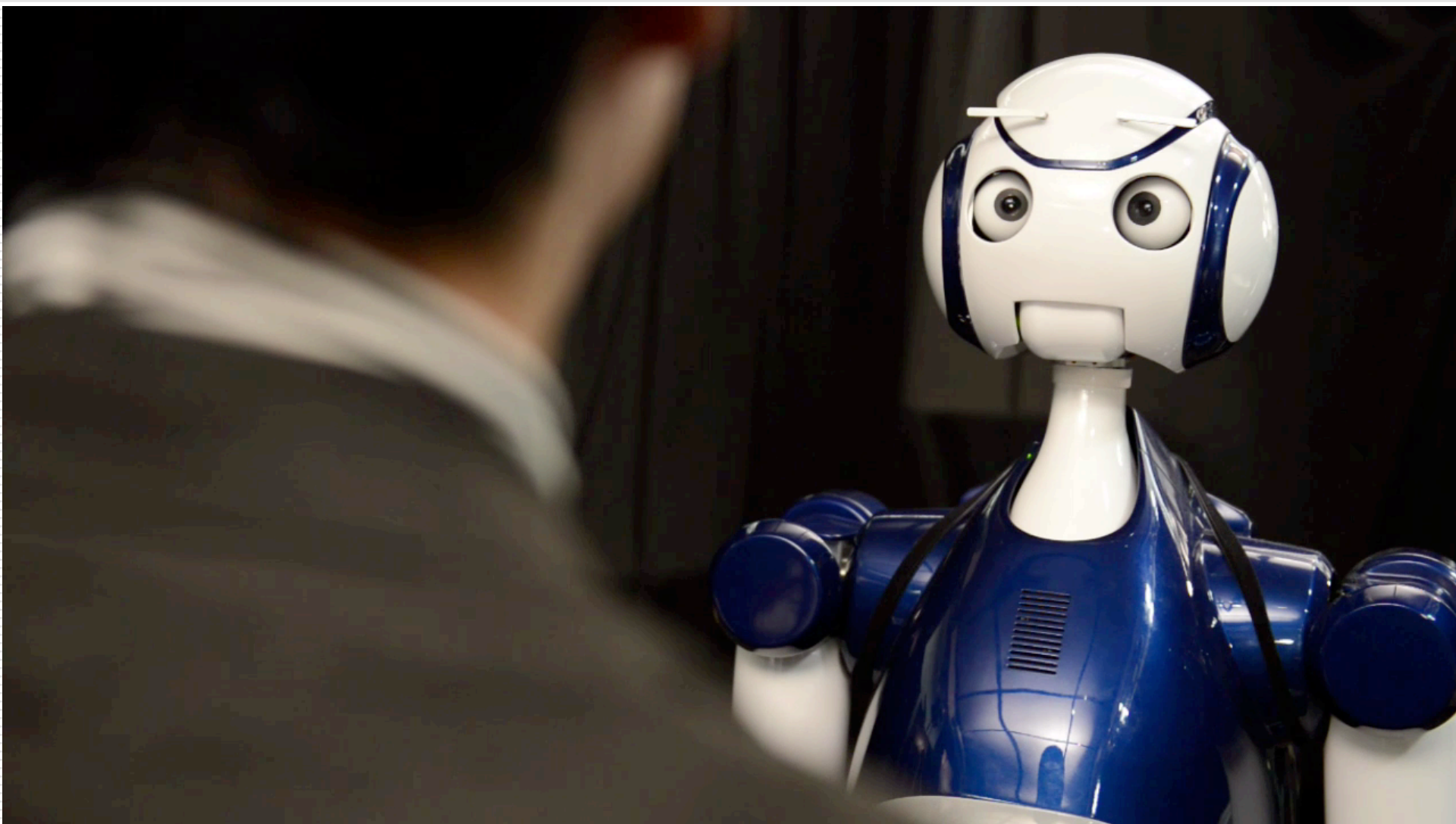
Chaining Episodes



Automatic Sentence Generation

1. Collect sentences from review sites.
2. Extract the subjectively informative parts
(using Conditional Random Fields).
3. Change style from written to spoken
4. Ranking & select
 - Sentence length (not so long = easy to speak)
 - TF-IDF of nouns (high = topic related)
 - Number of adjectives
(many = quantitatively more informative)
 - Frequency of adjectives
(low = qualitatively more informative)

Demo :



Facilitation :

balance maintenance of

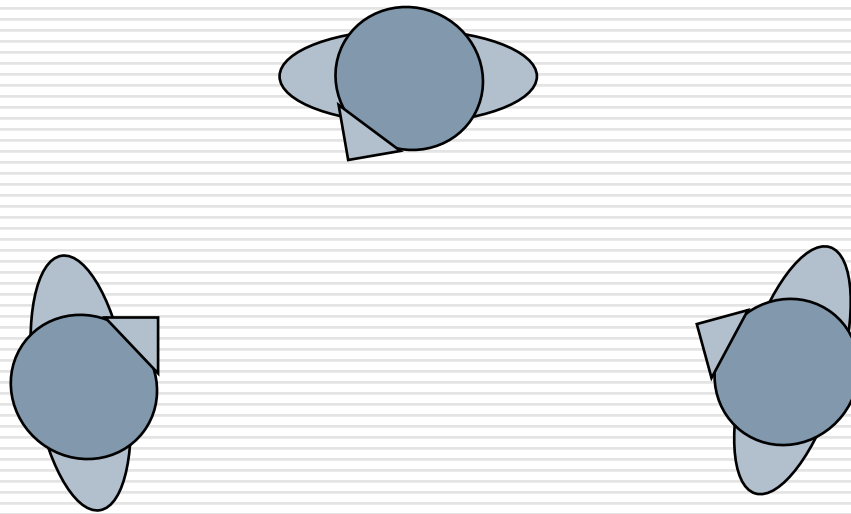
ENGAGEMENT DENSITY

Addressing to everyone

Enjoyable system should be

Attentive/thoughtful !!

What is engagement density

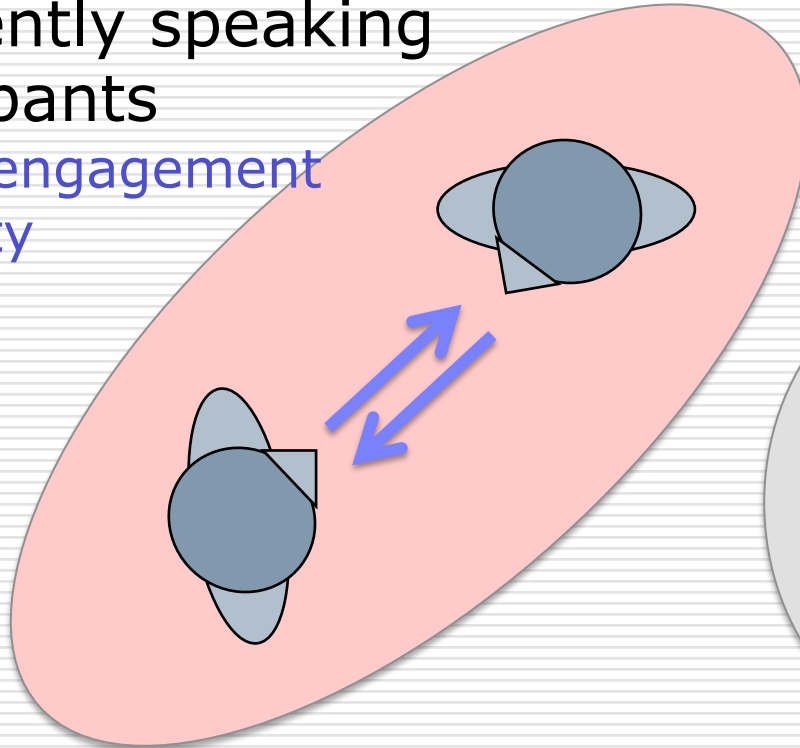


in group conversation situation,

What is engagement density

Frequently speaking
Participants

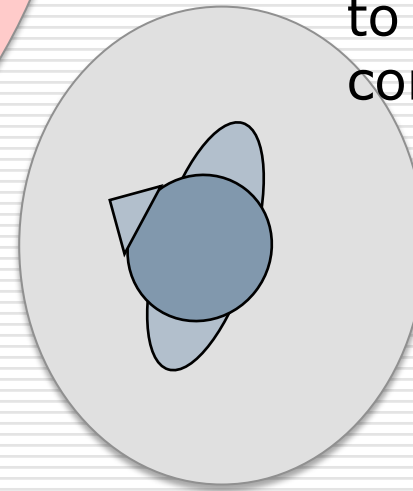
= High engagement
density



Silent participant

= Low engagement
density

= feel difficulty
to join in the
conversation

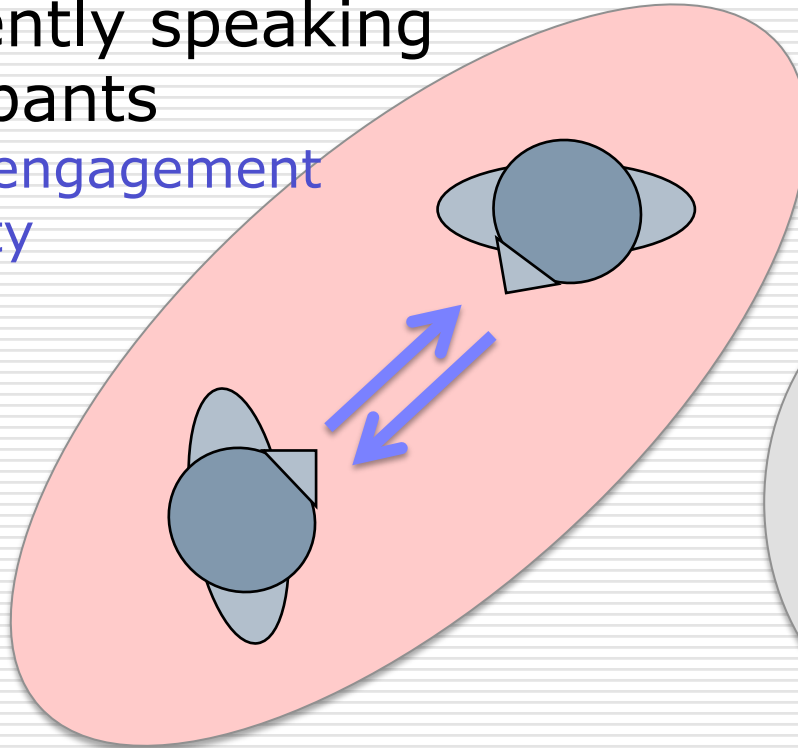


in group conversation situation,

What is engagement density

Frequently speaking
Participants

= High engagement
density

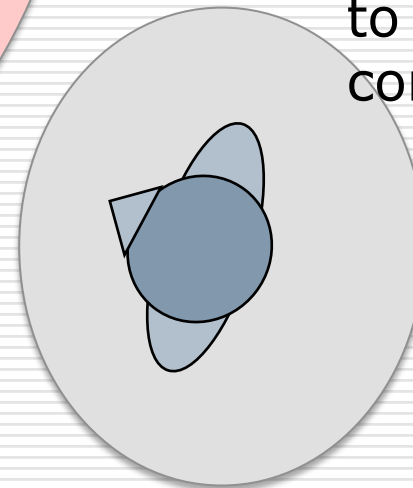


Harmonic Party

Silent participant

= Low engagement
density

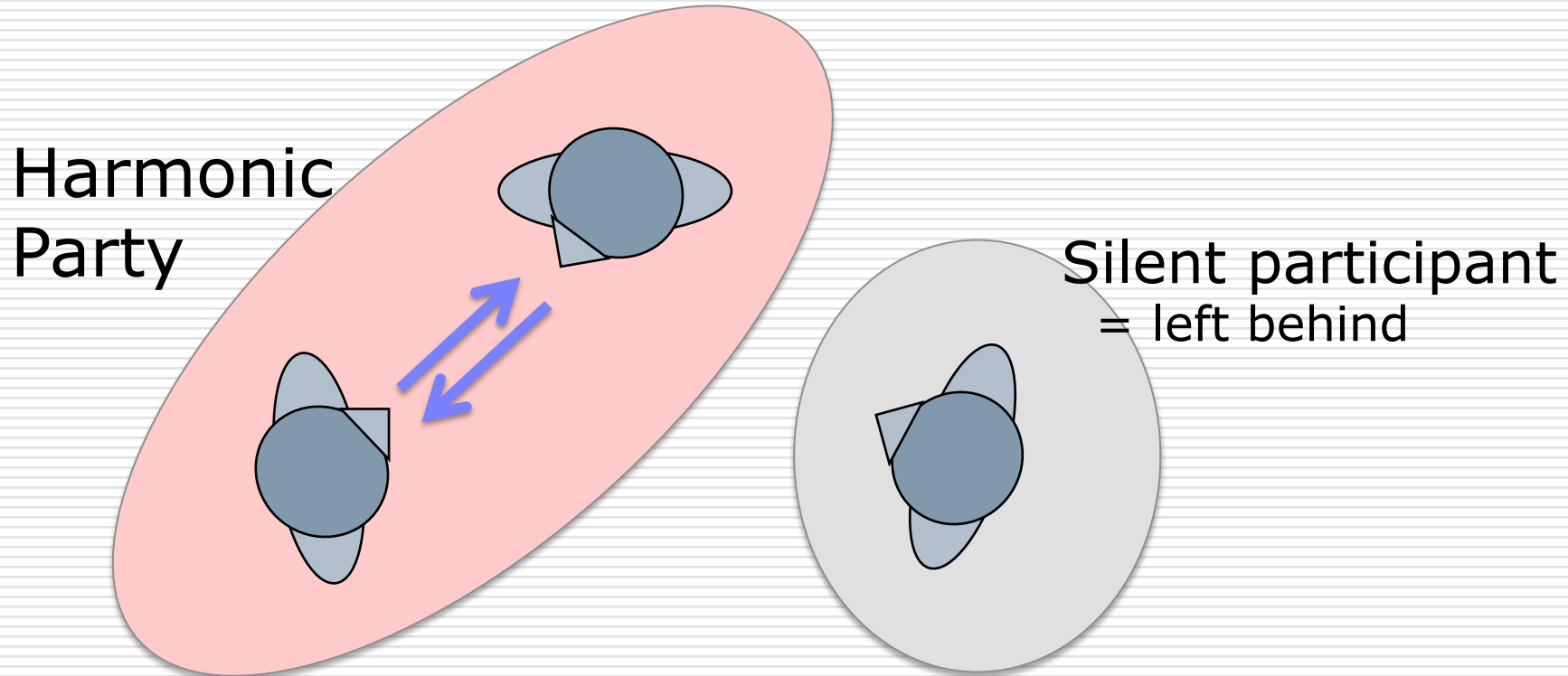
= feel difficulty
to join in the
conversation



Non-Harmonic

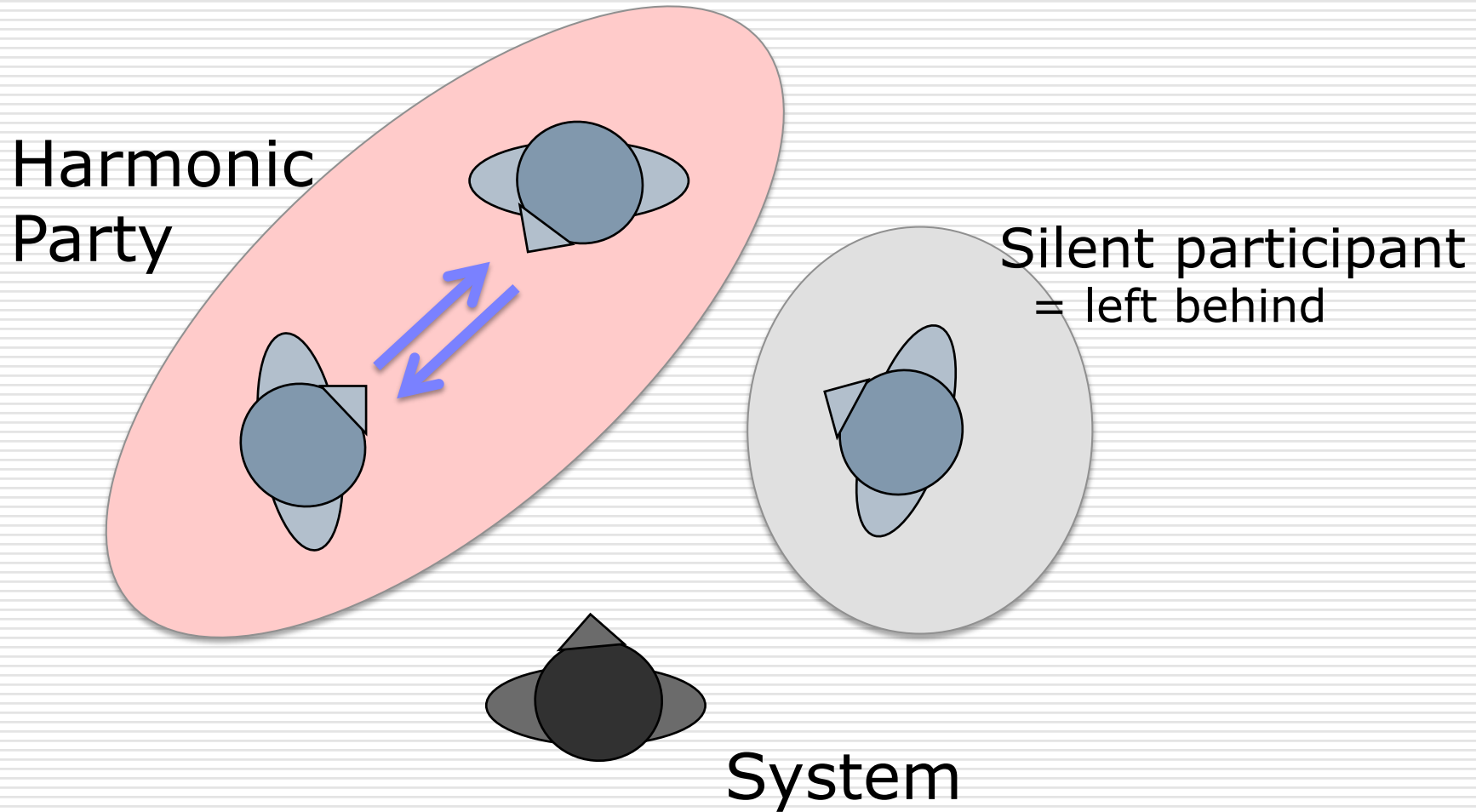
in group conversation situation,

Problem: imbalance of engagement density

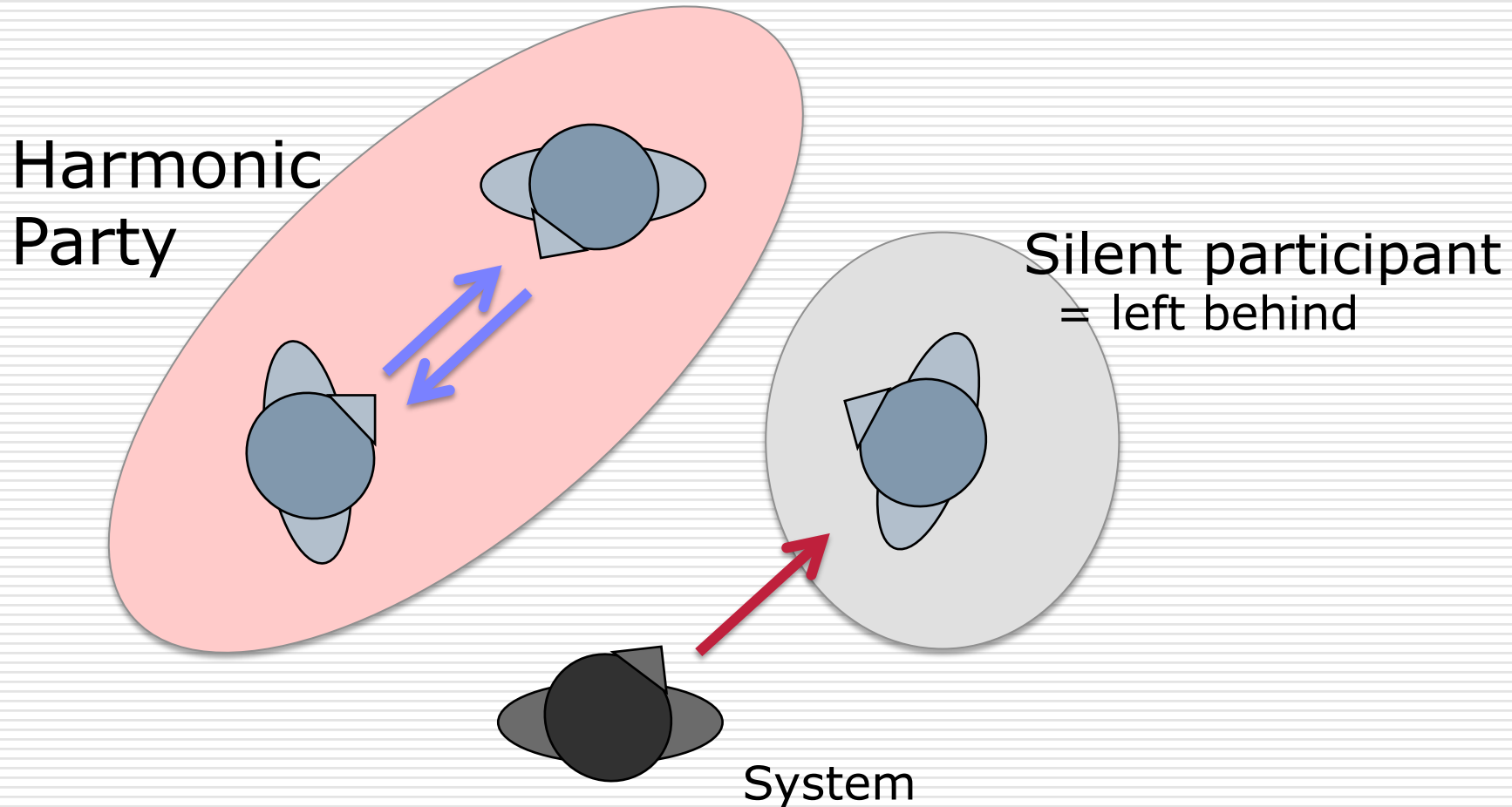


- Silent participant need help to join in the conversation.
- Harmonic party may not recognize the situation.

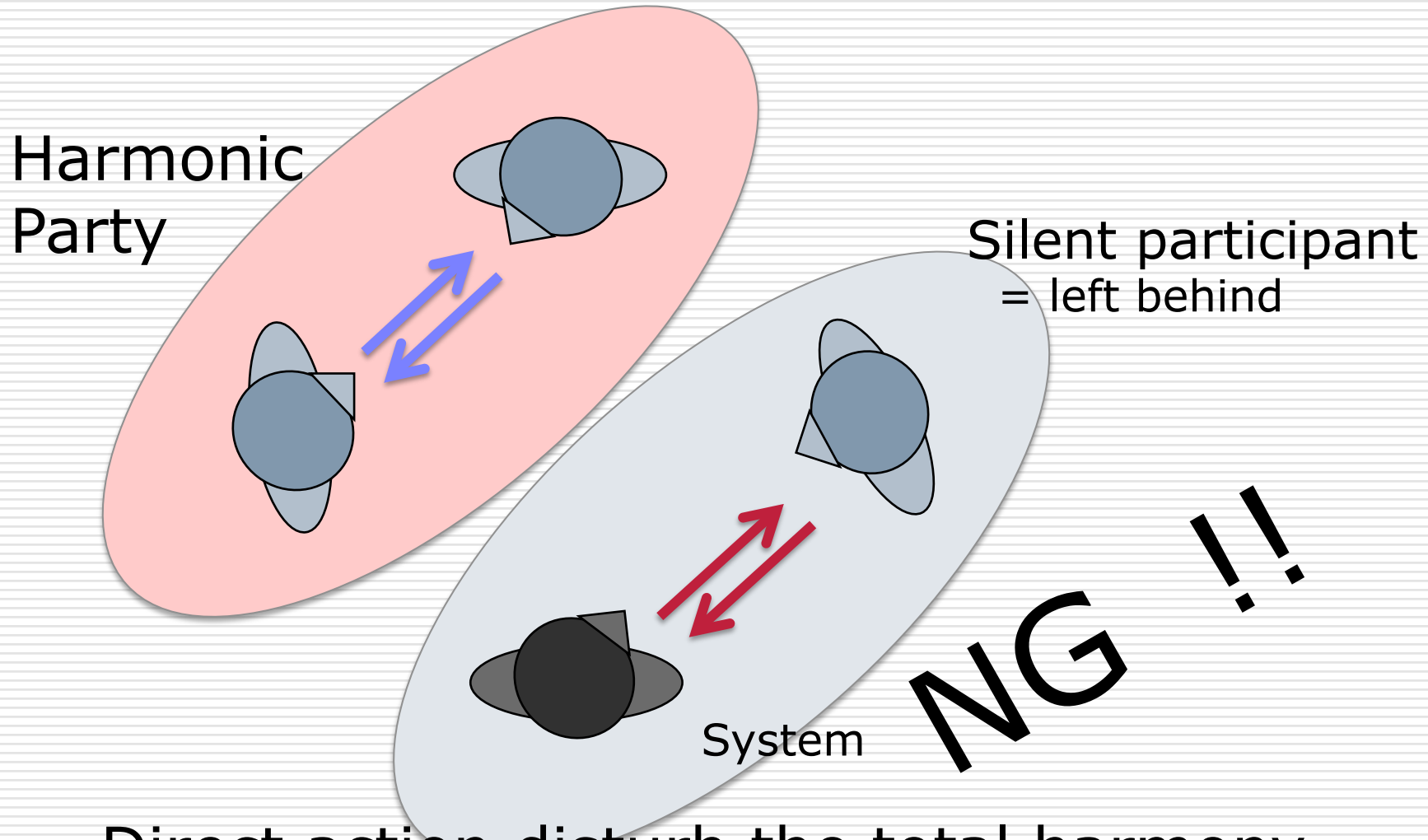
Theme: make an agent to balance engmnt. dens.



Speaking to silent participant directly...

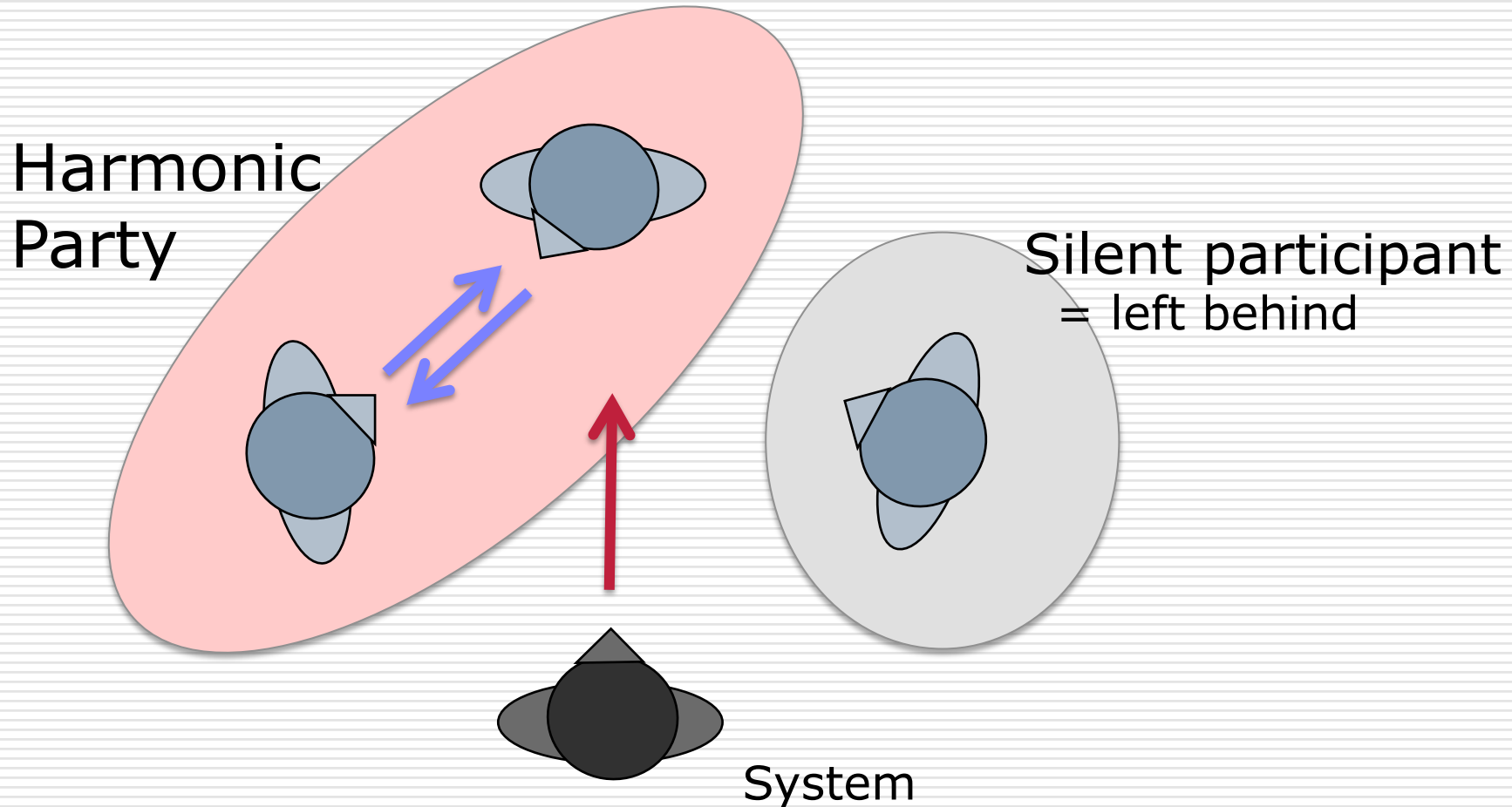


... cause another problem

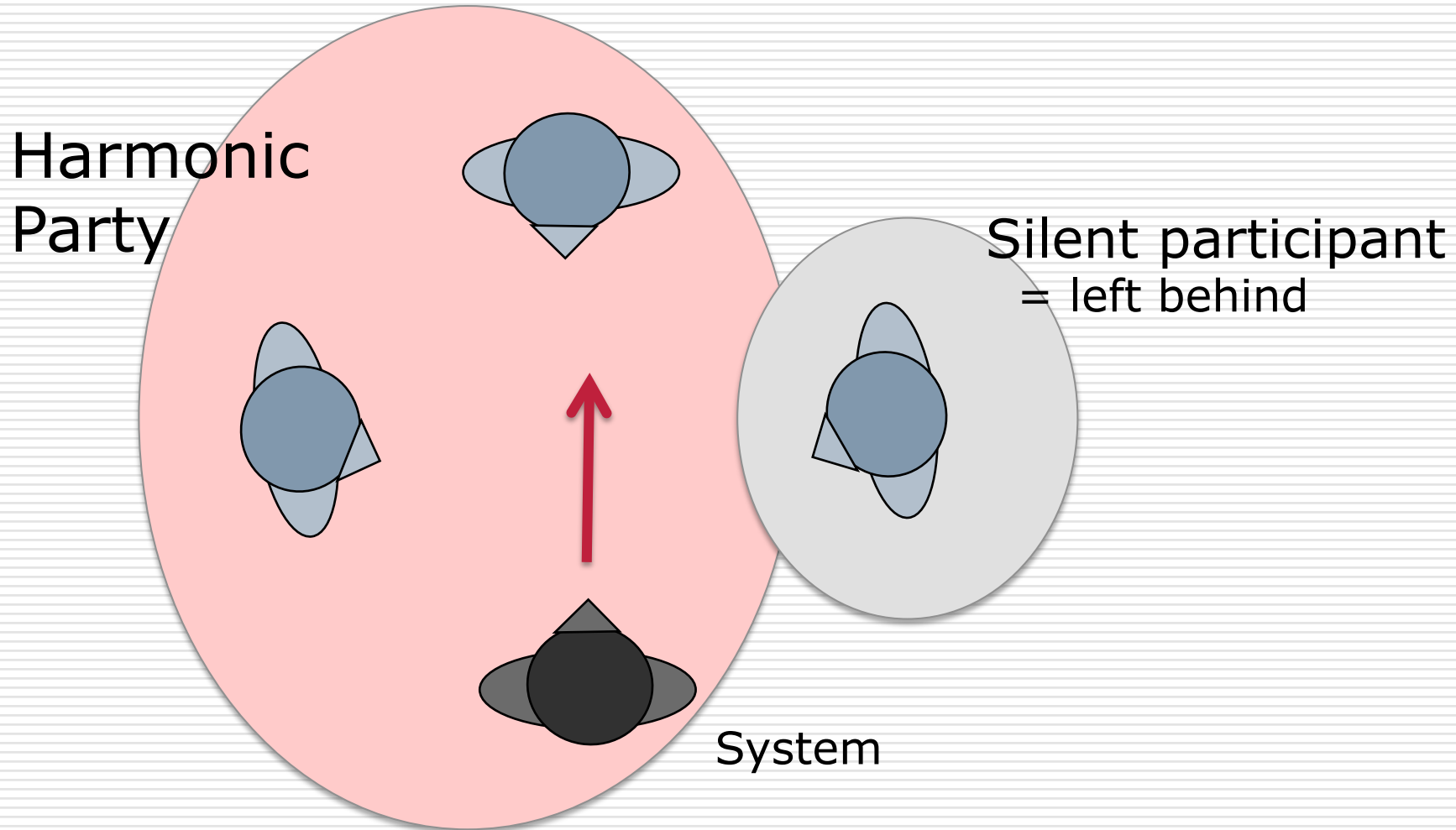


Direct action disturb the total harmony.

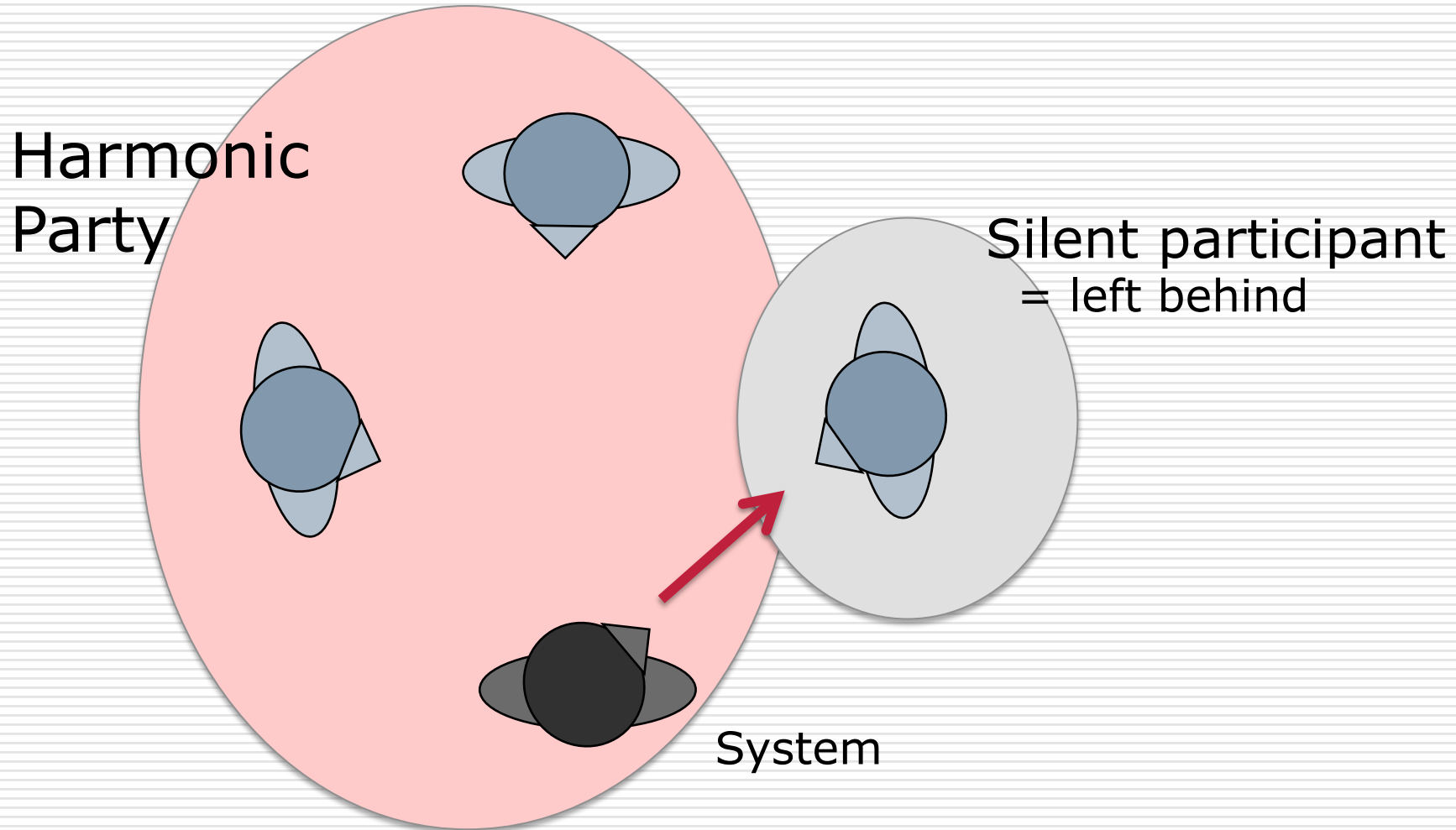
Speak to harmonic party, first



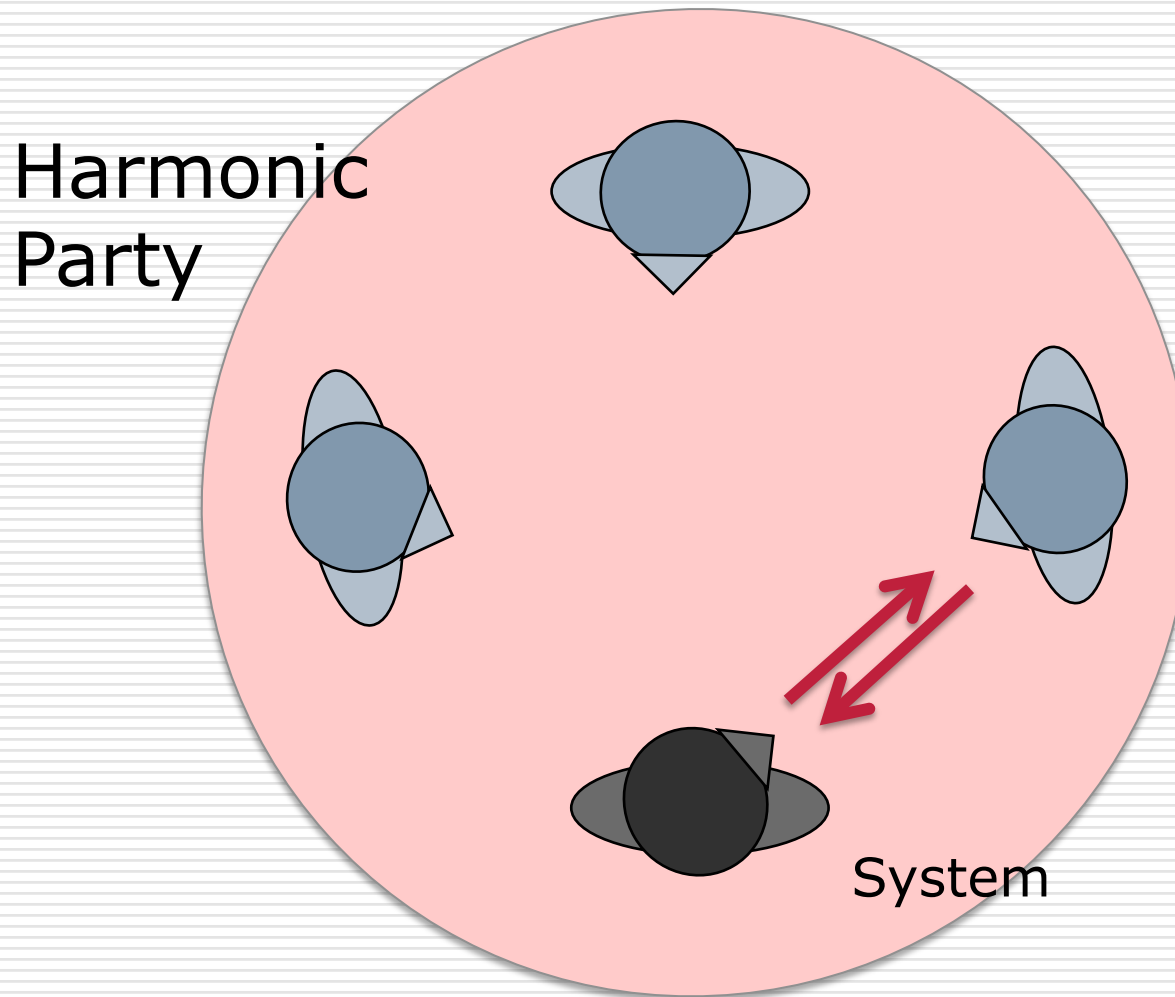
Join in the harmonic party



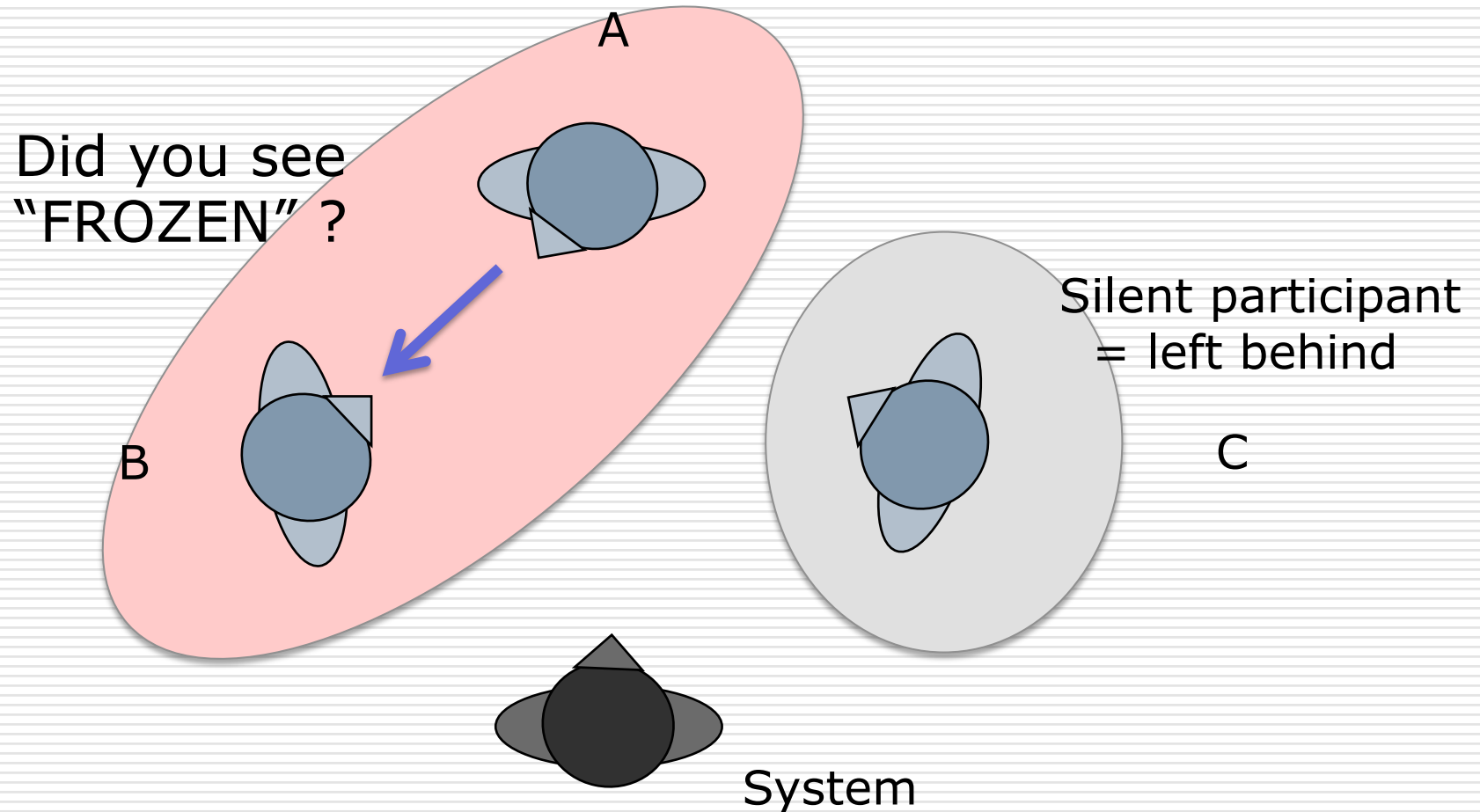
Then, speak to silent participant



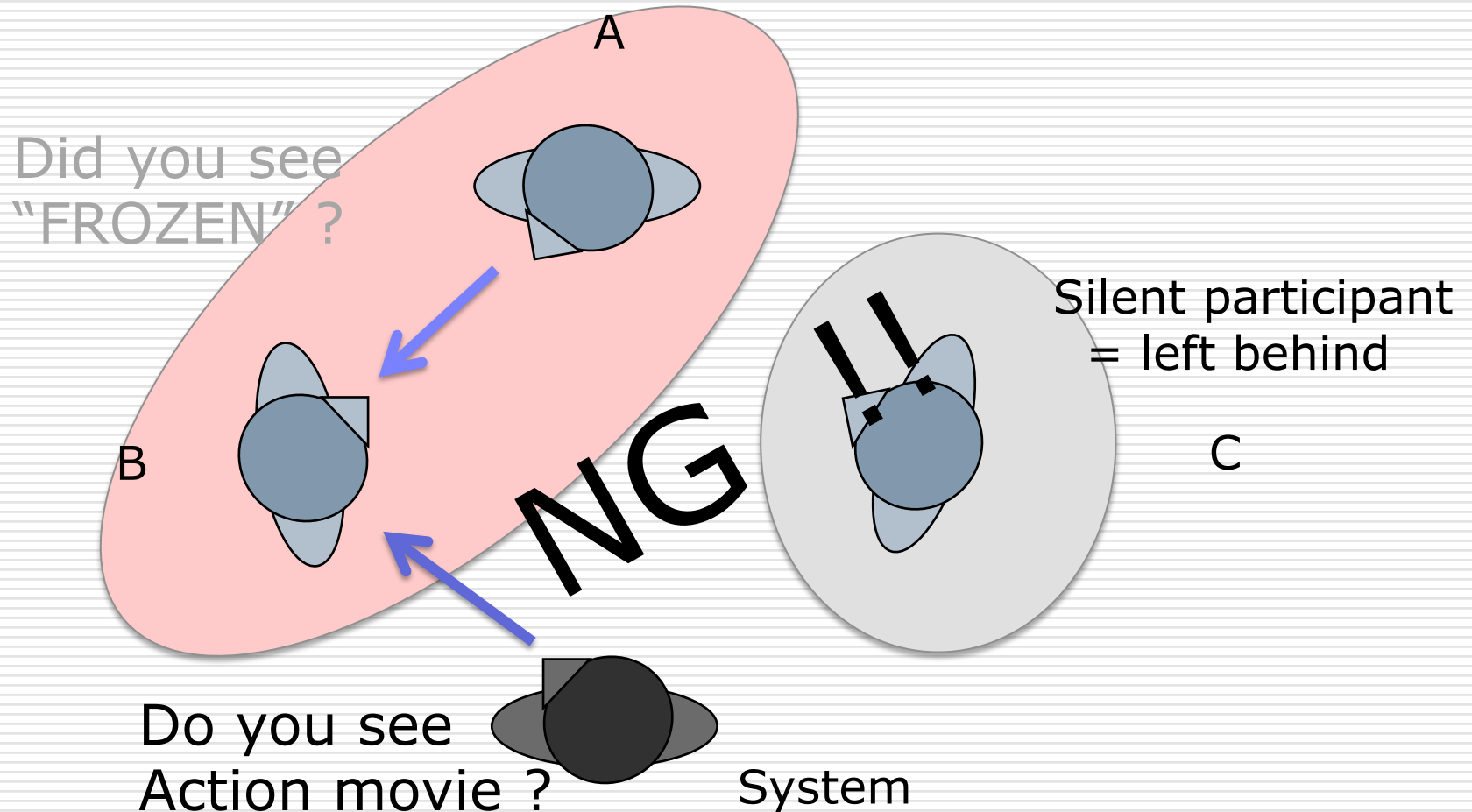
Then, all participants are in a big HP



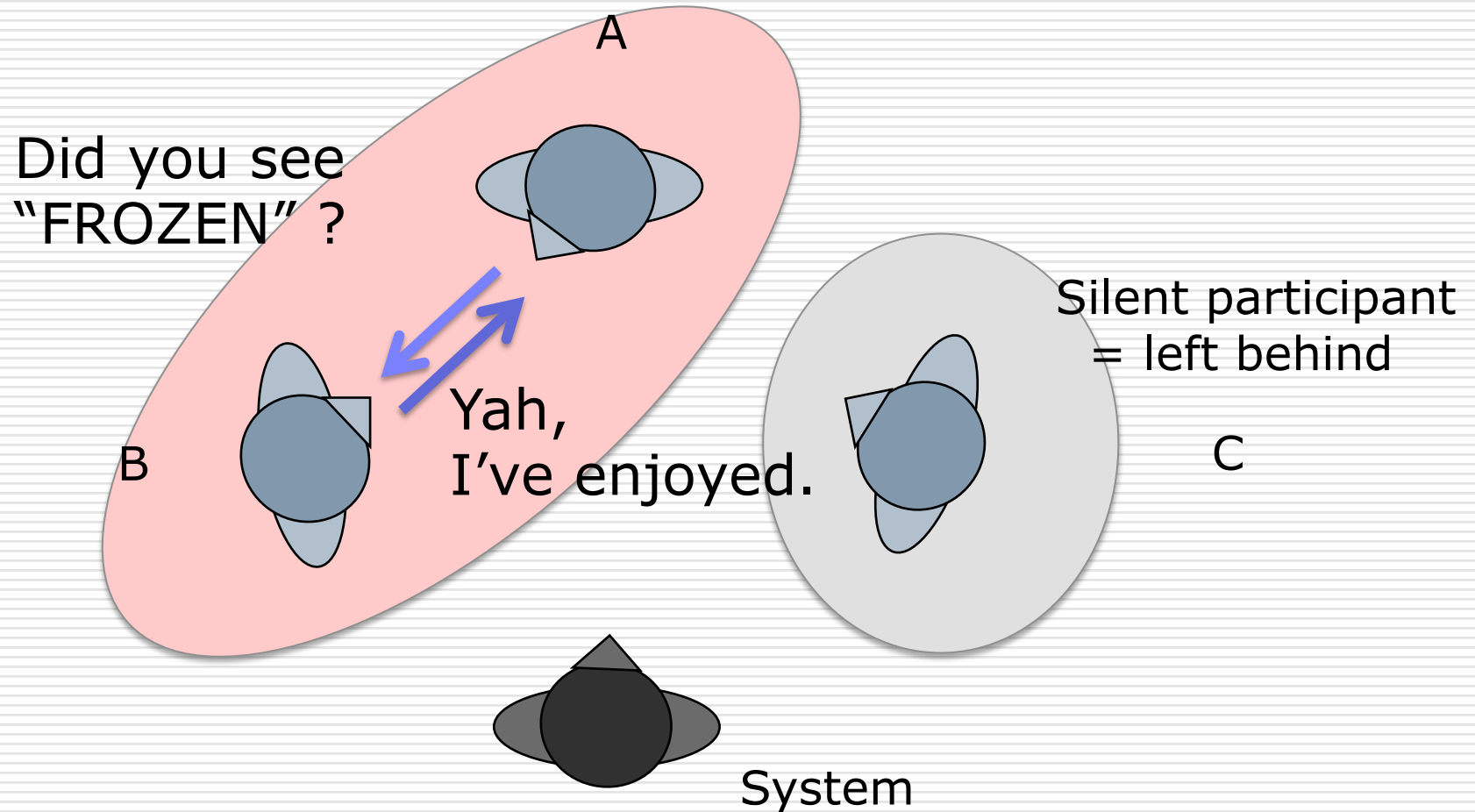
Rules to interrupt



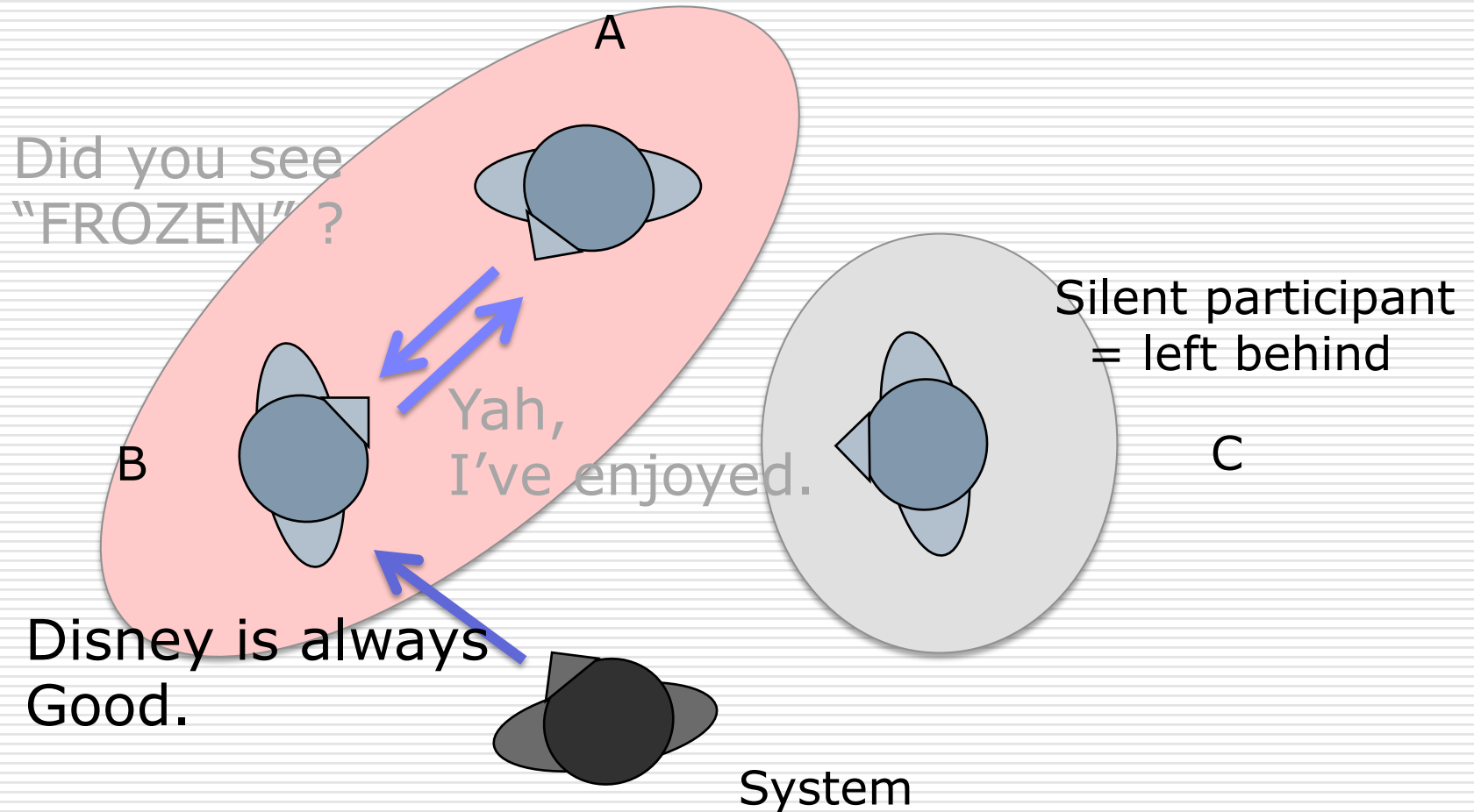
Rules to interrupt



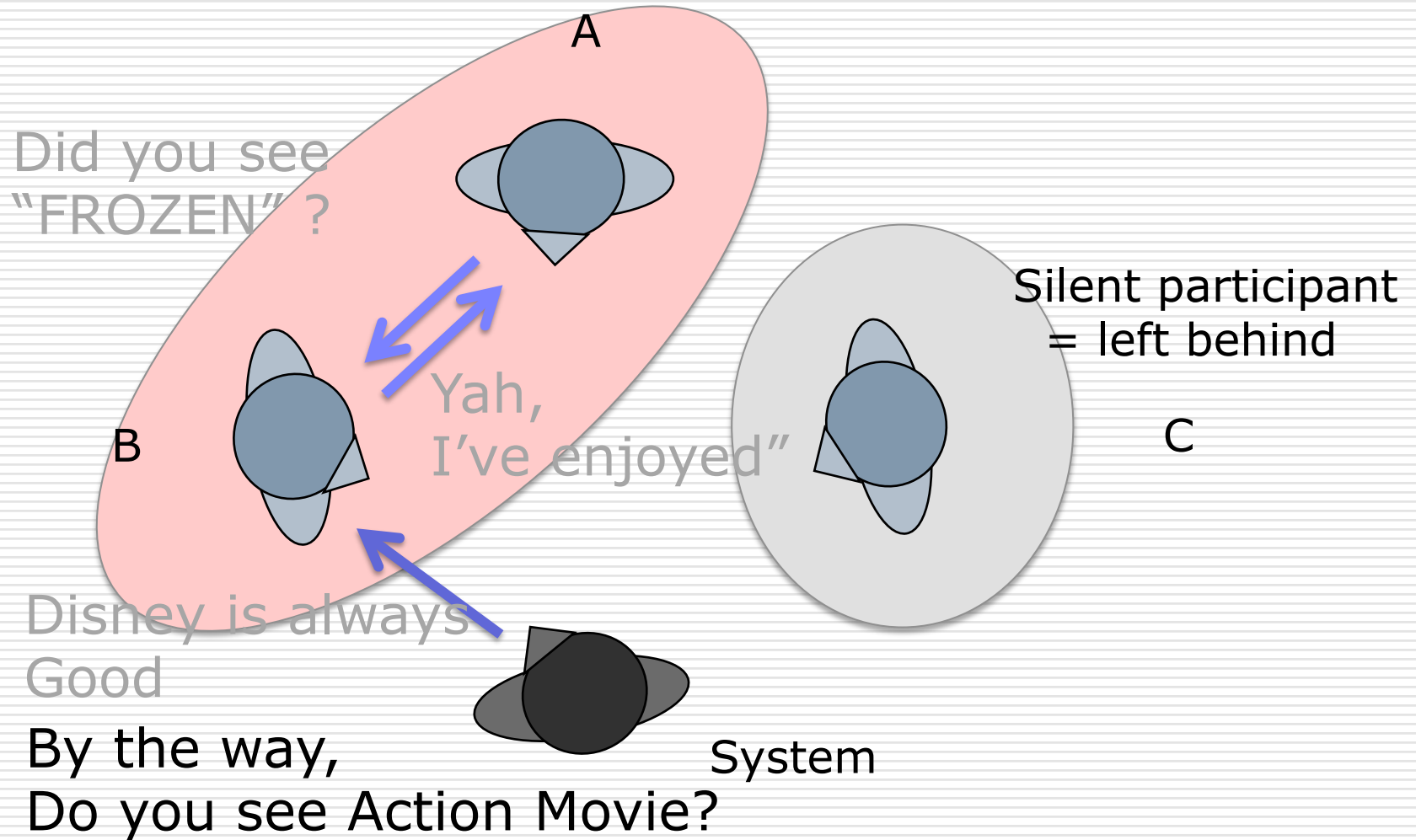
Rules to interrupt



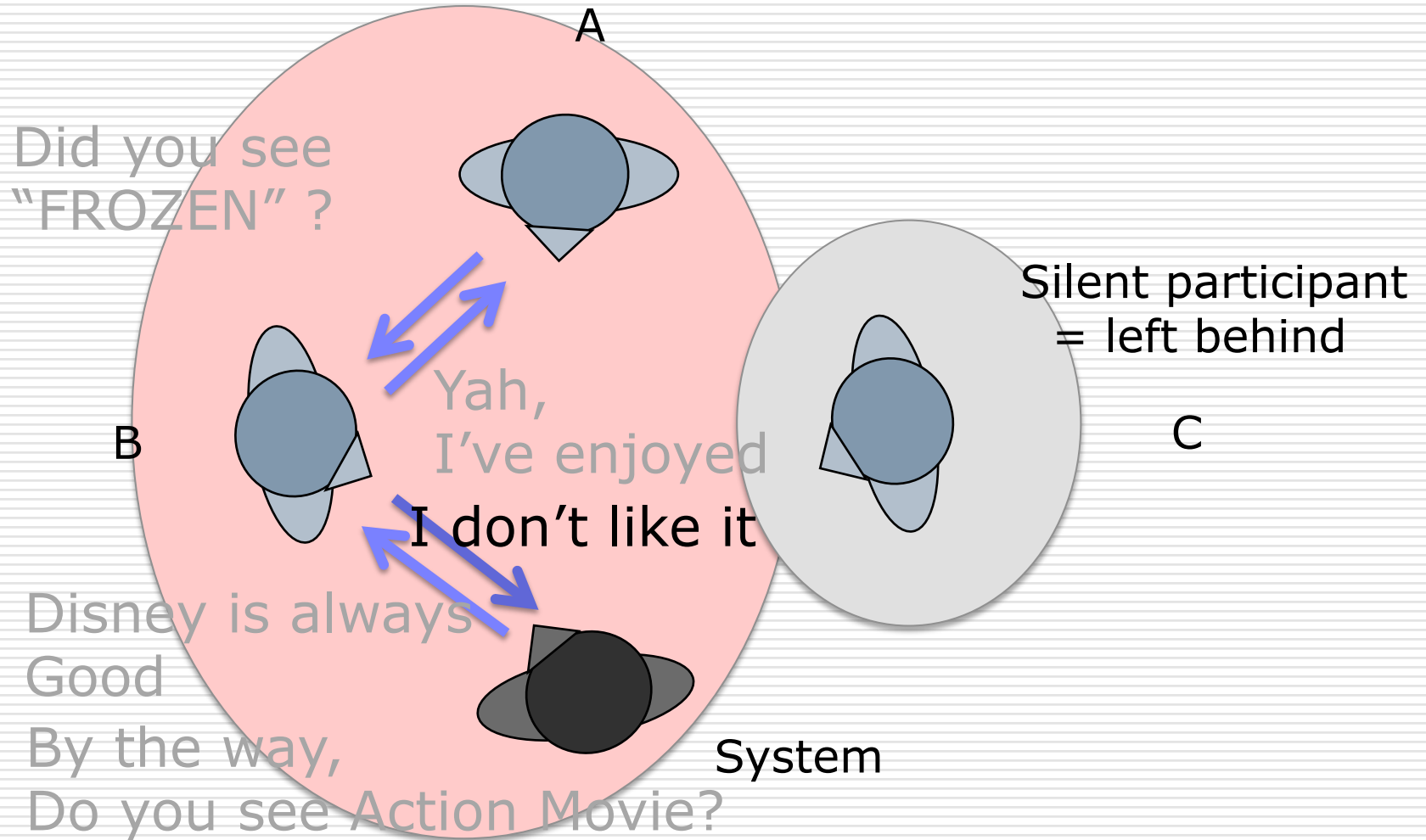
Rules to interrupt



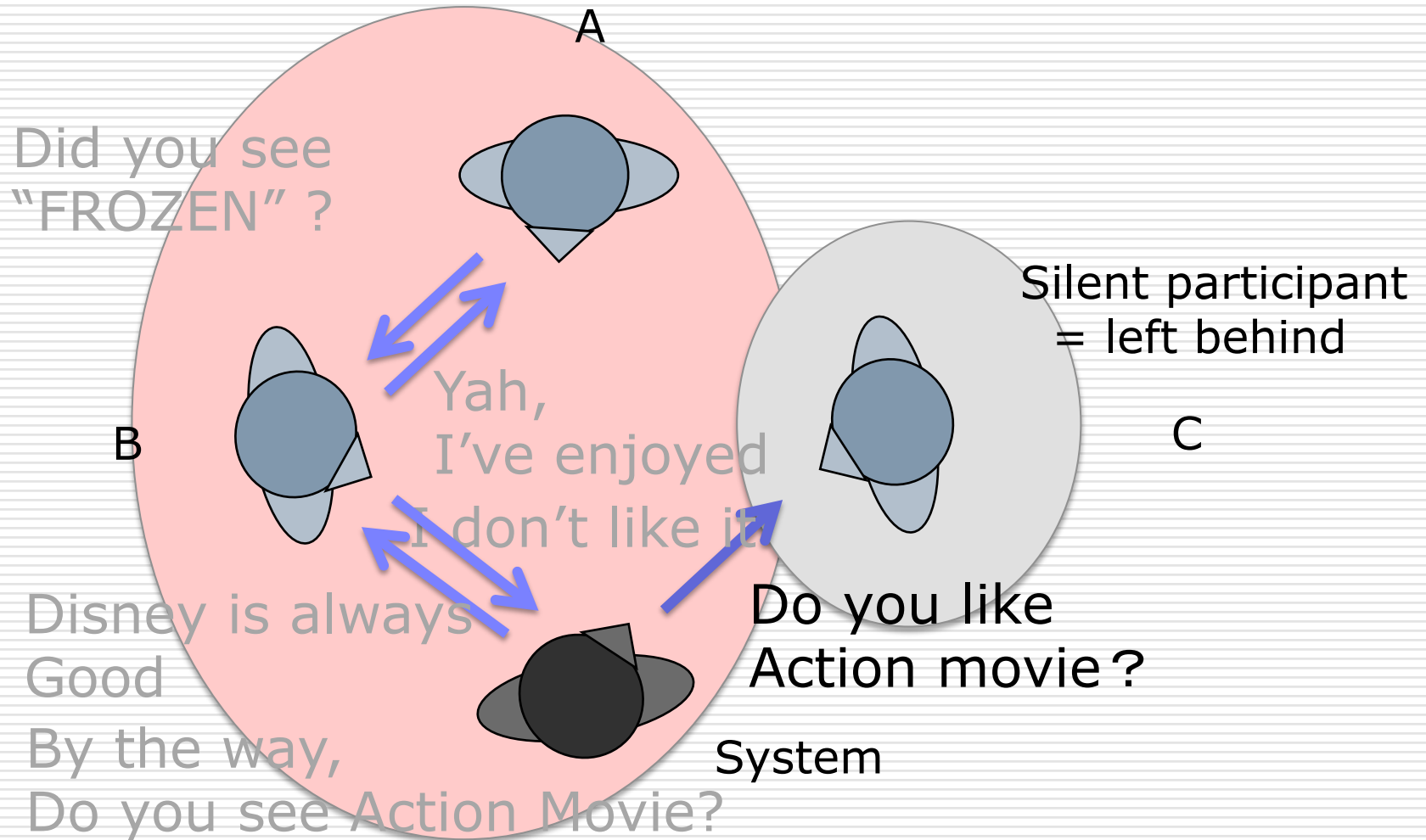
Rules to interrupt



Rules to interrupt



Rules to interrupt



Recognizing **CONVERSATIONAL MOVE**

Conversational Move

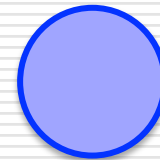
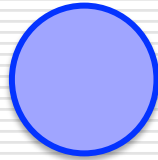
Move : Small utterance unit.

Move categories:

- **Initiation** A: I like Disney movies.
 Provide information
- **Eliciting** A: Do you like it ?
 Request someone to provide information
- **Response** B: Yes, of course.
 Response to the eliciting
- **Follow-up** A: It's good.
 ... etc.

State transition to get initiative

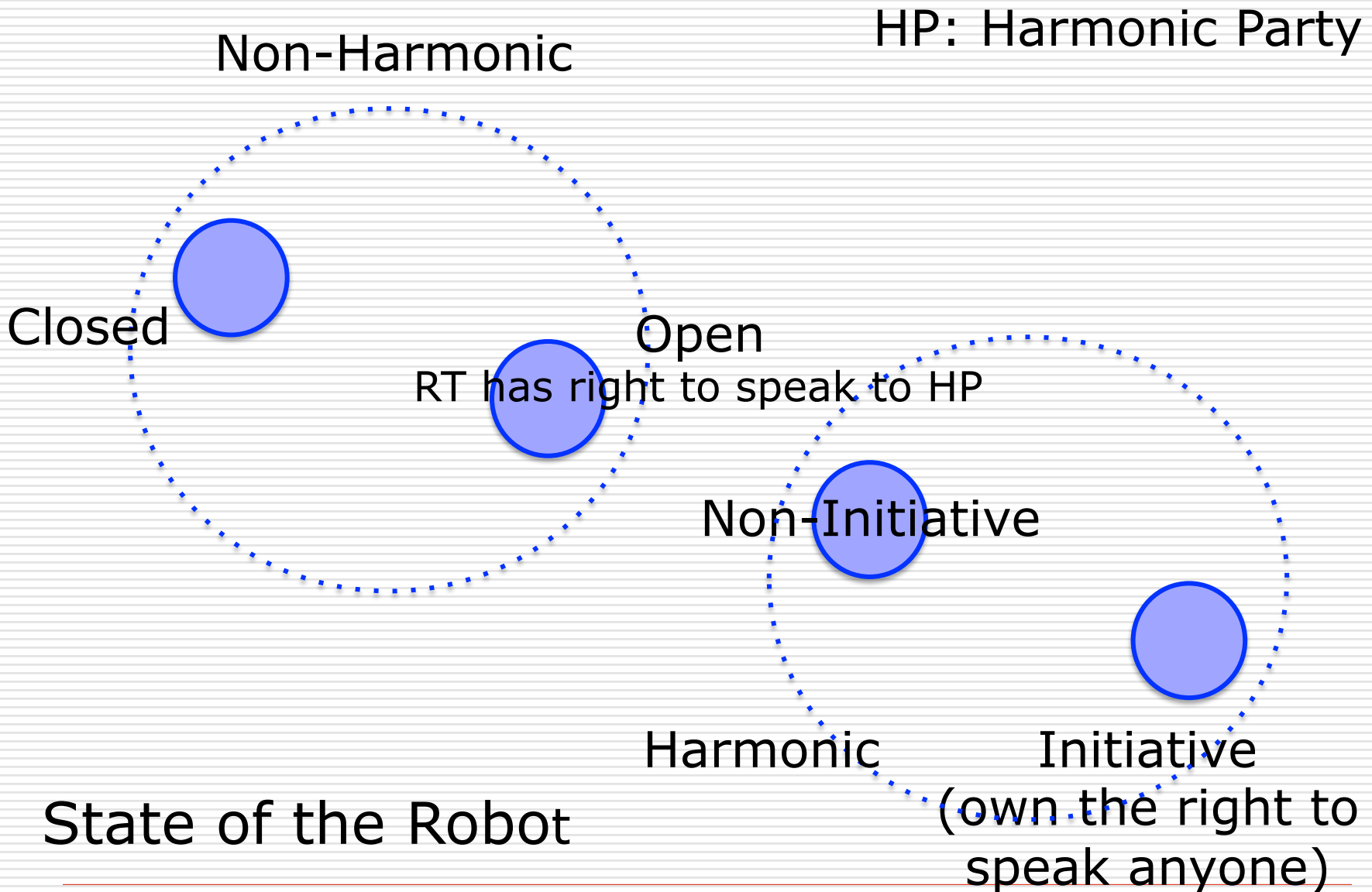
Non-Harmonic



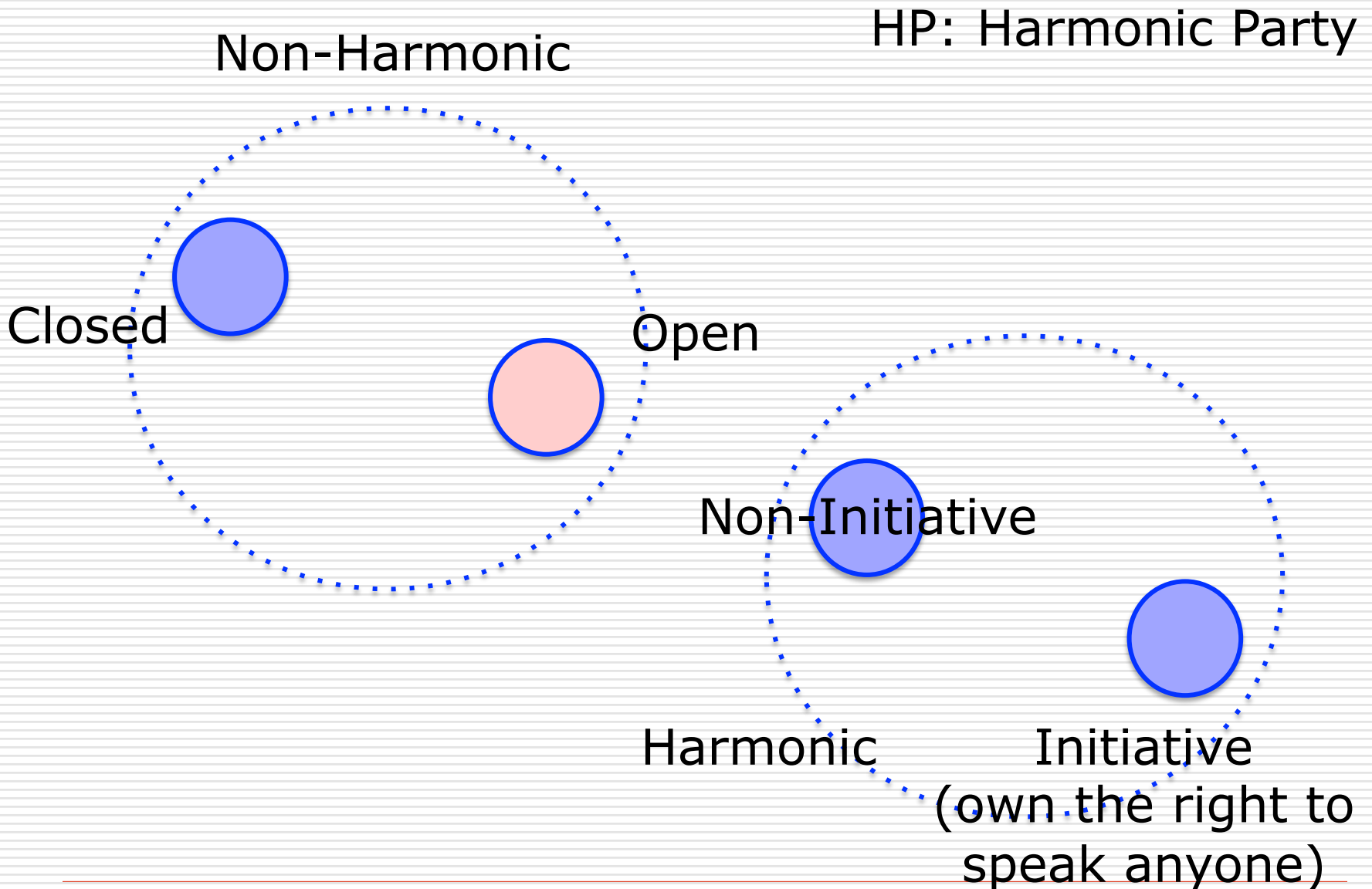
Harmonic

State of the Robot

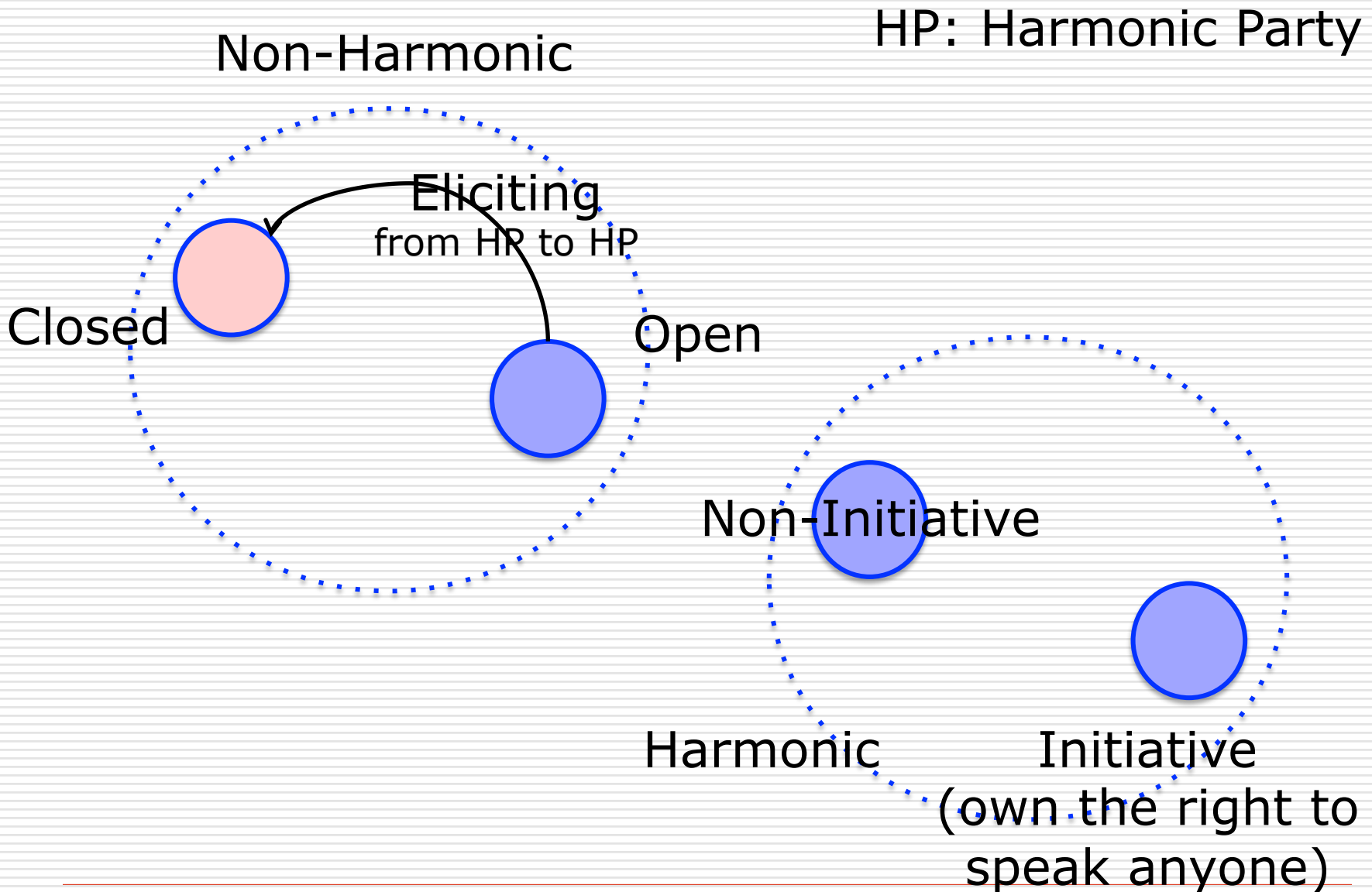
State transition to get initiative



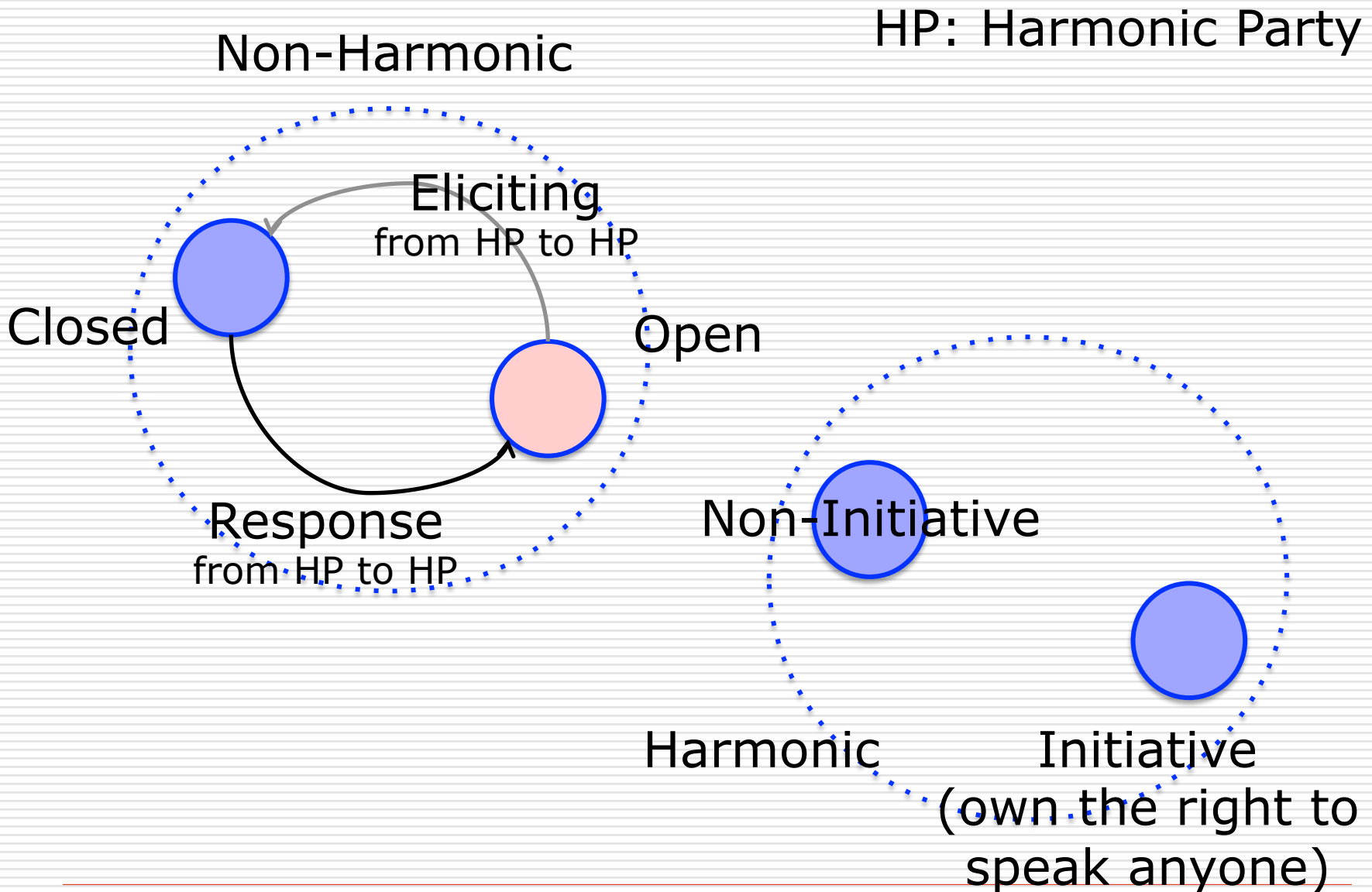
State transition to get initiative



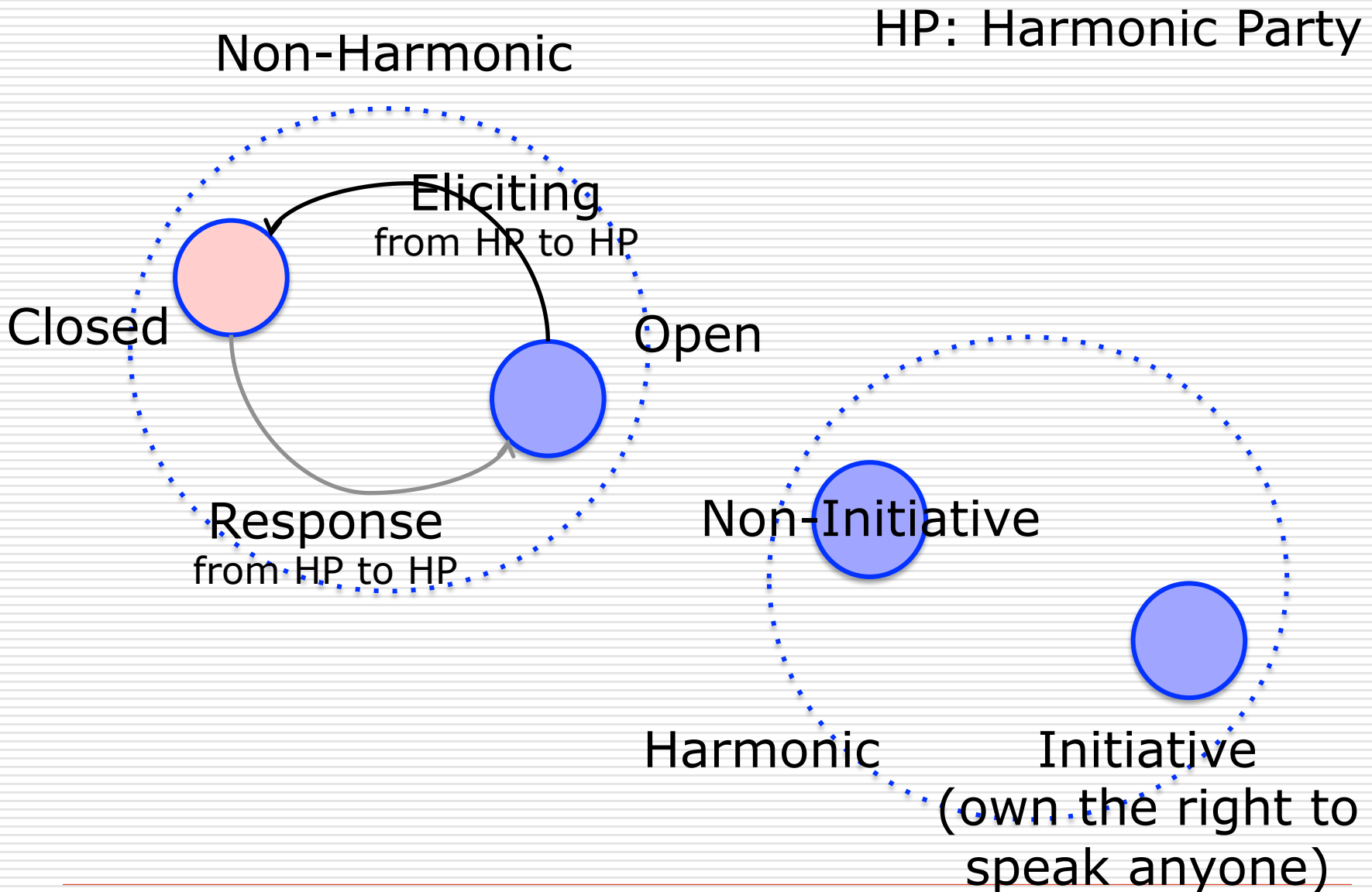
State transition to get initiative



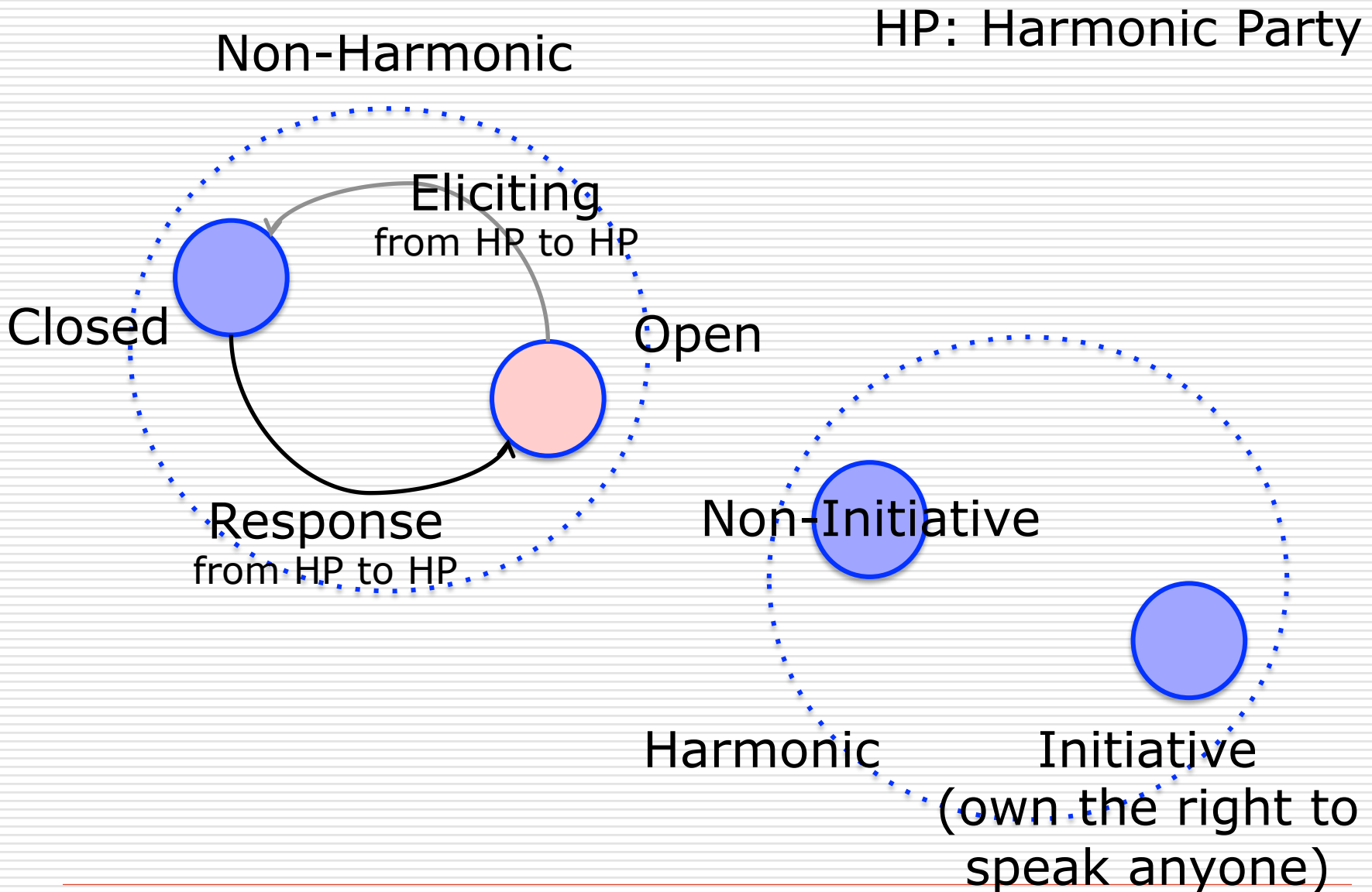
State transition to get initiative



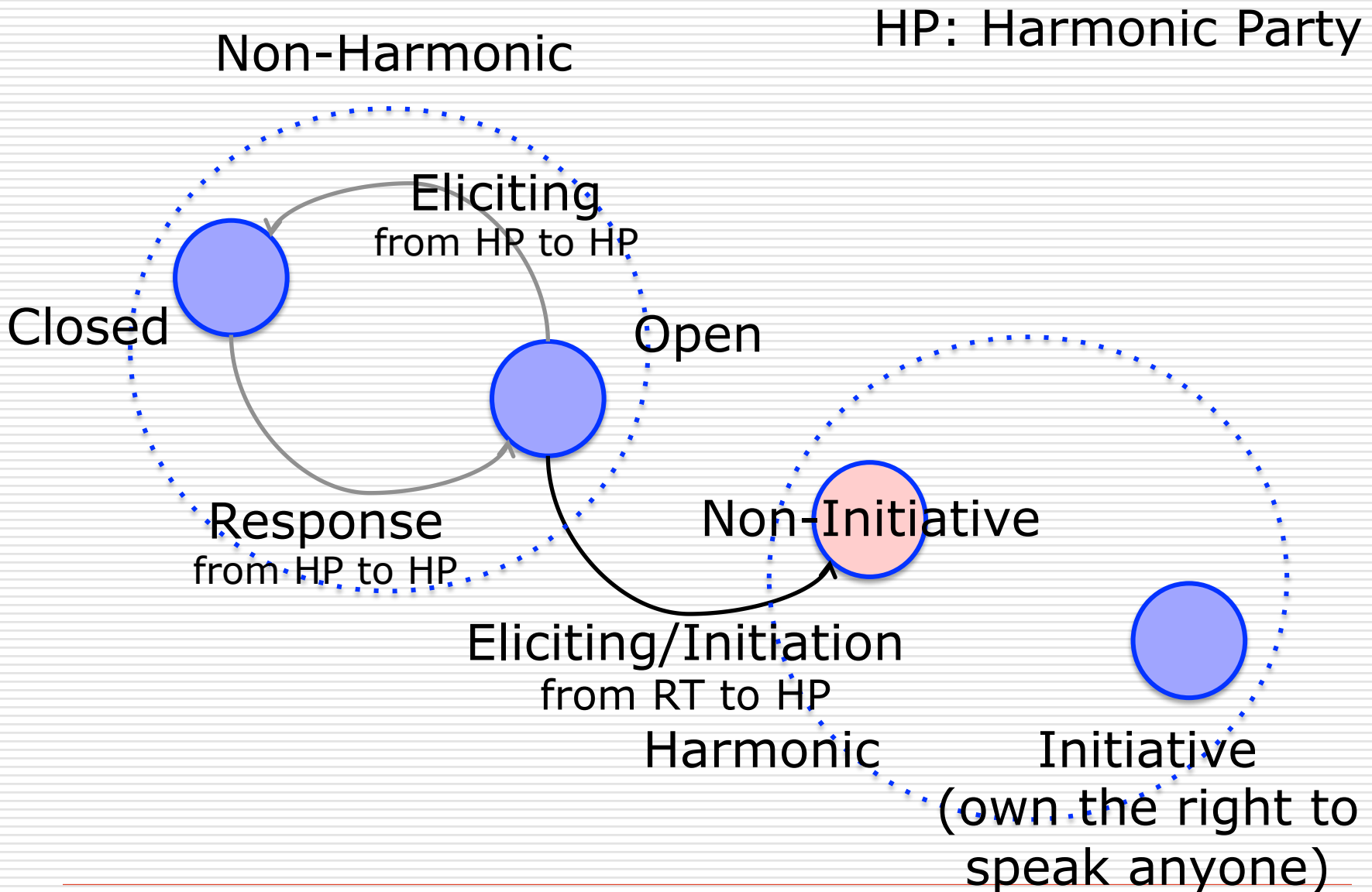
State transition to get initiative



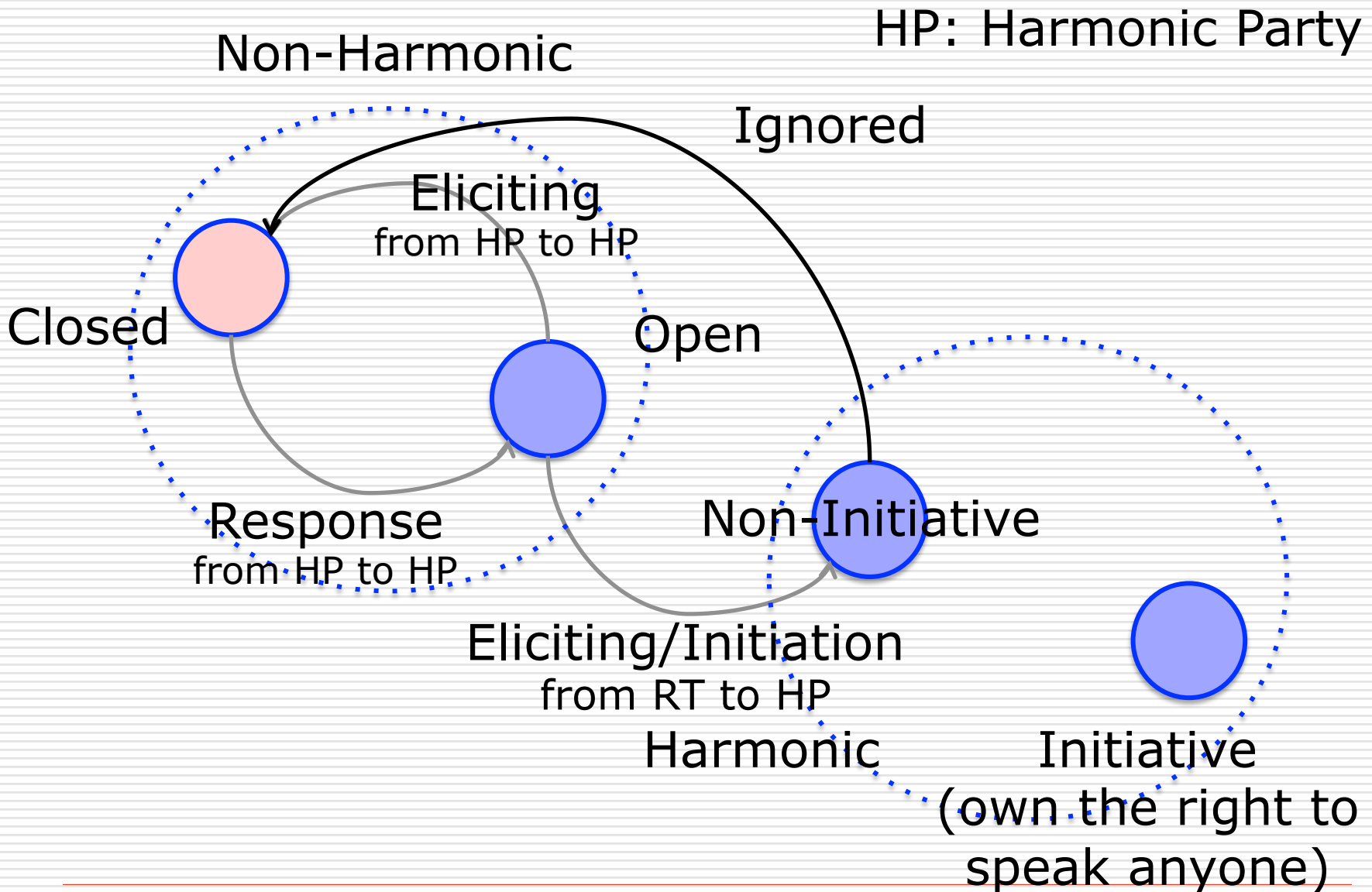
State transition to get initiative



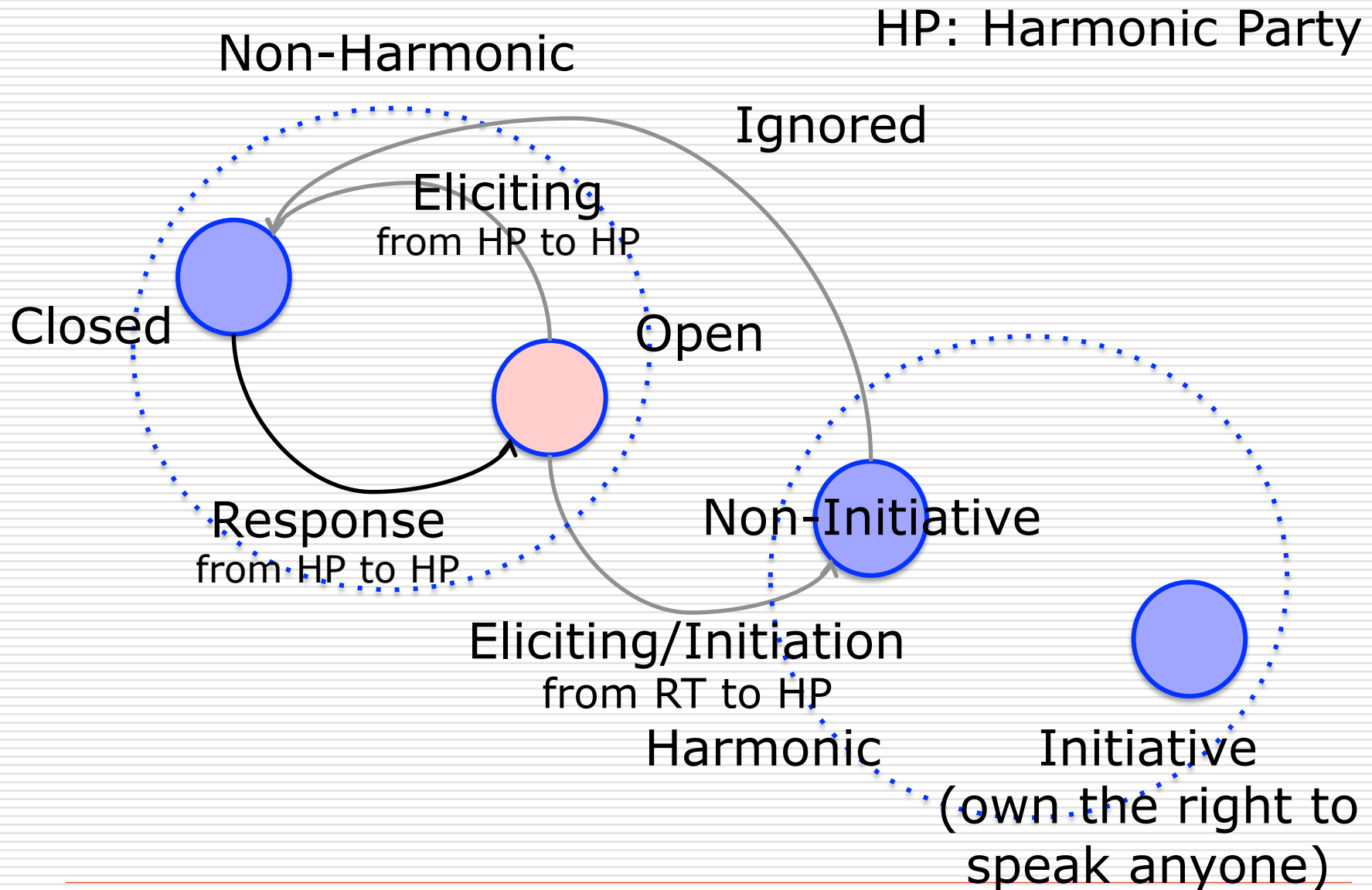
State transition to get initiative



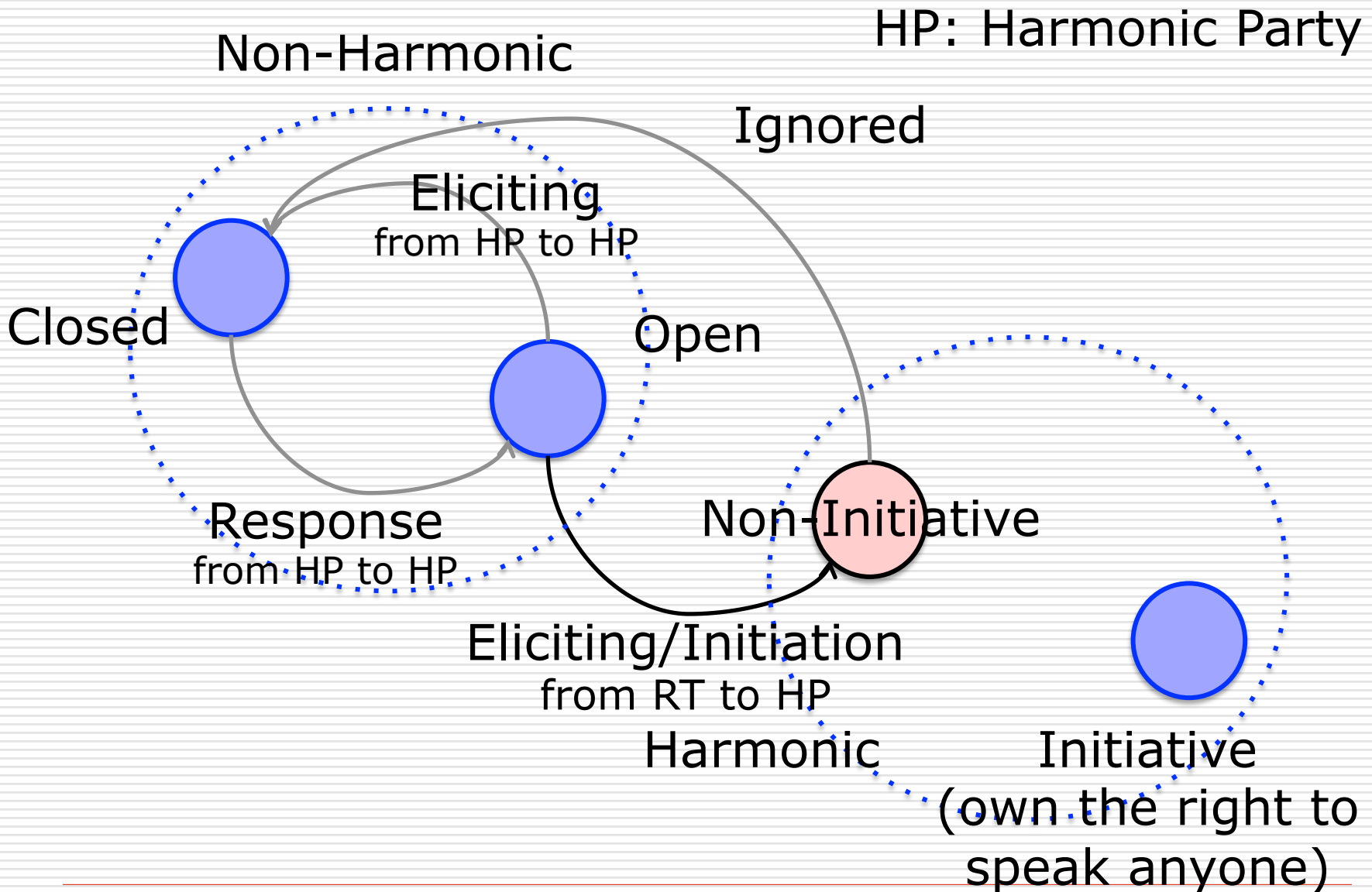
State transition to get initiative



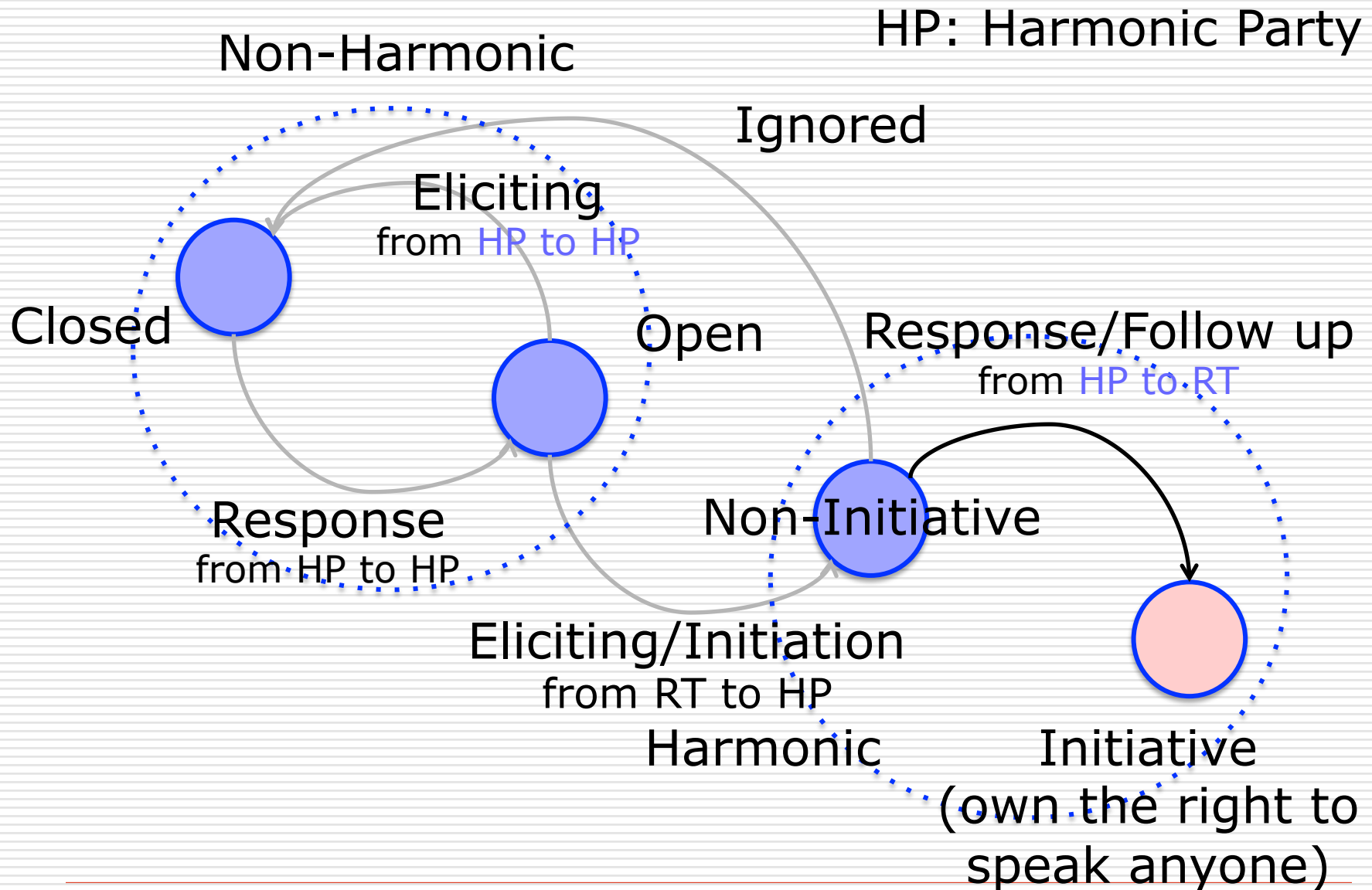
State transition to get initiative



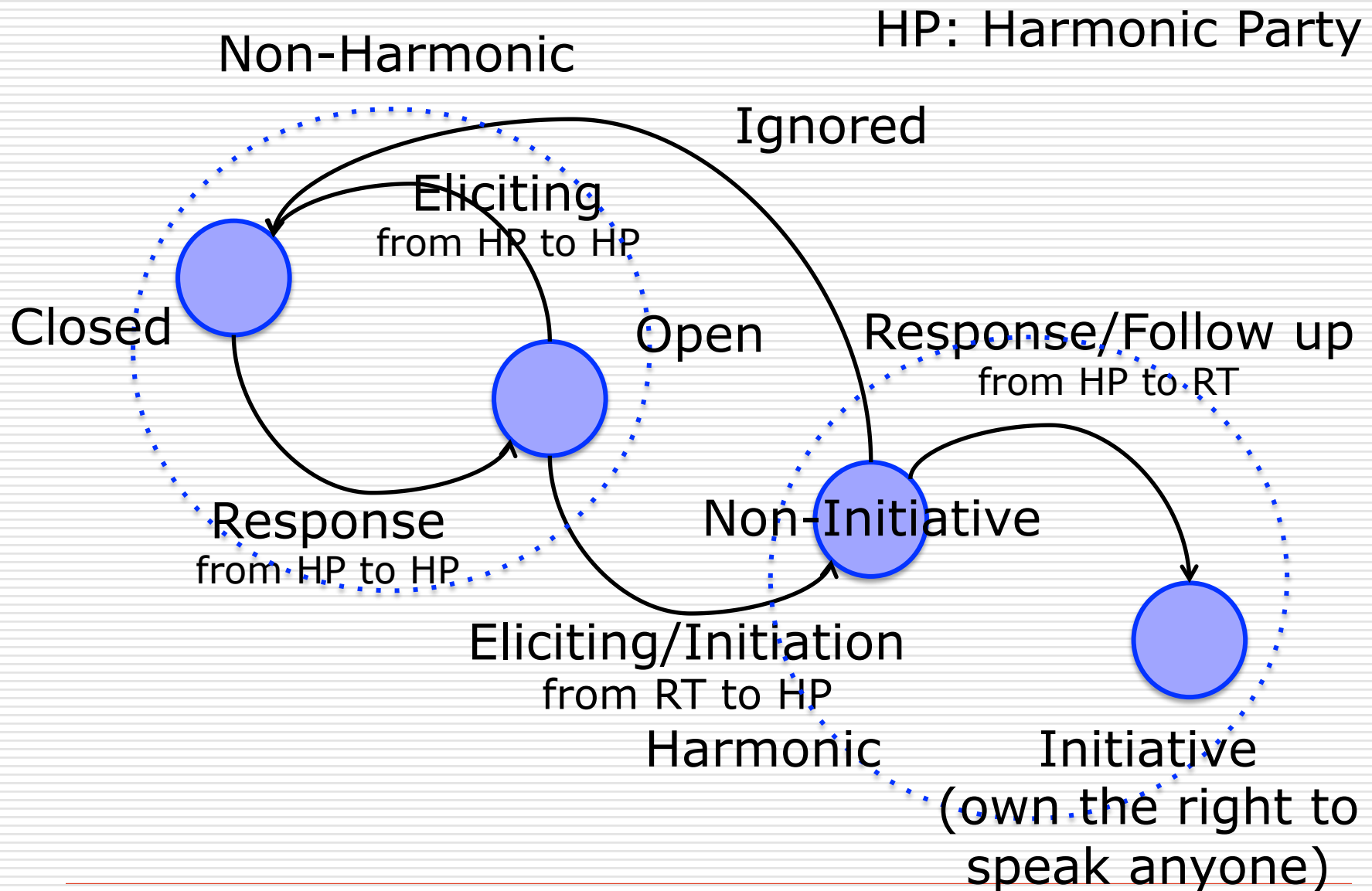
State transition to get initiative



State transition to get initiative

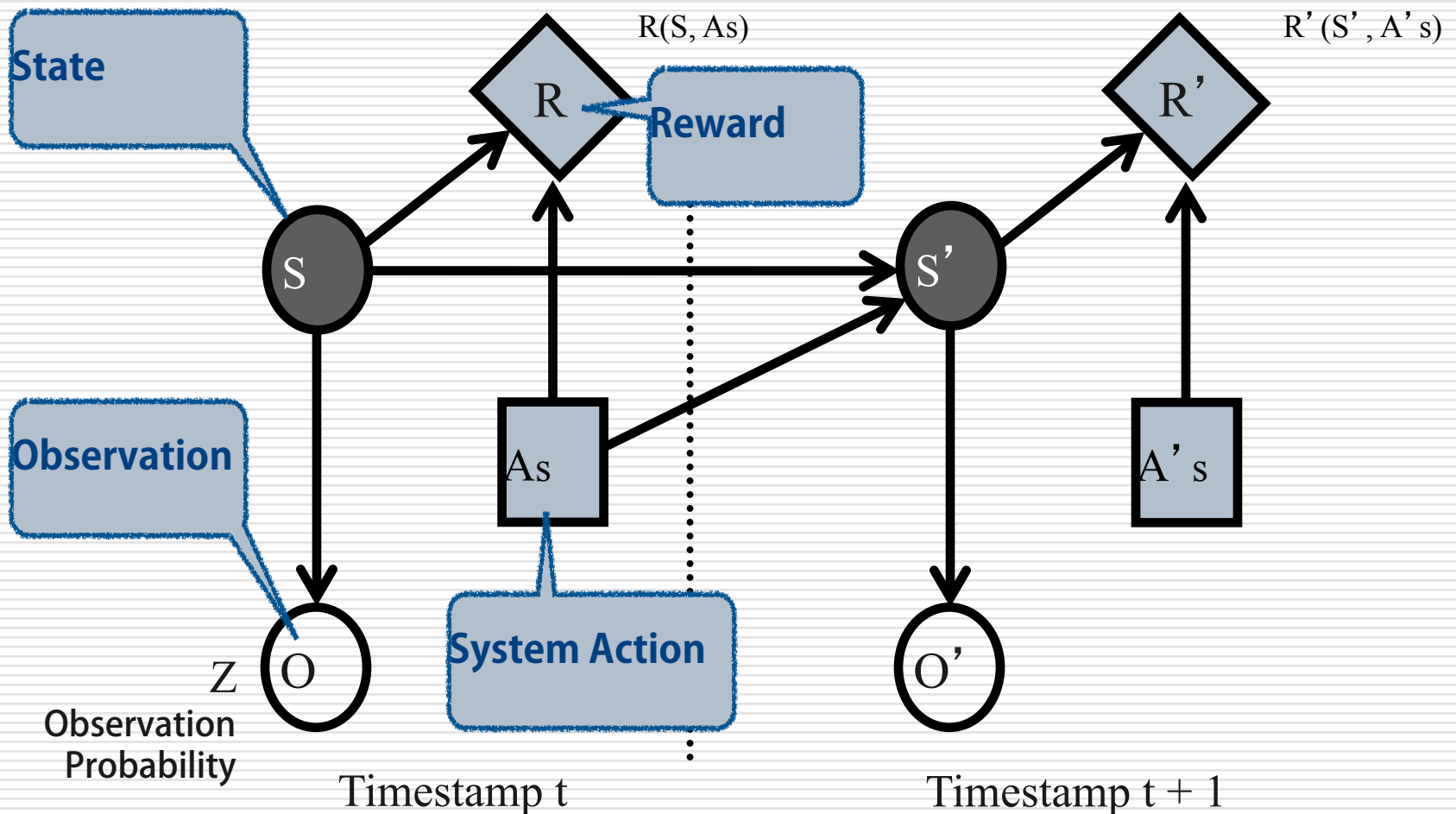


State transition



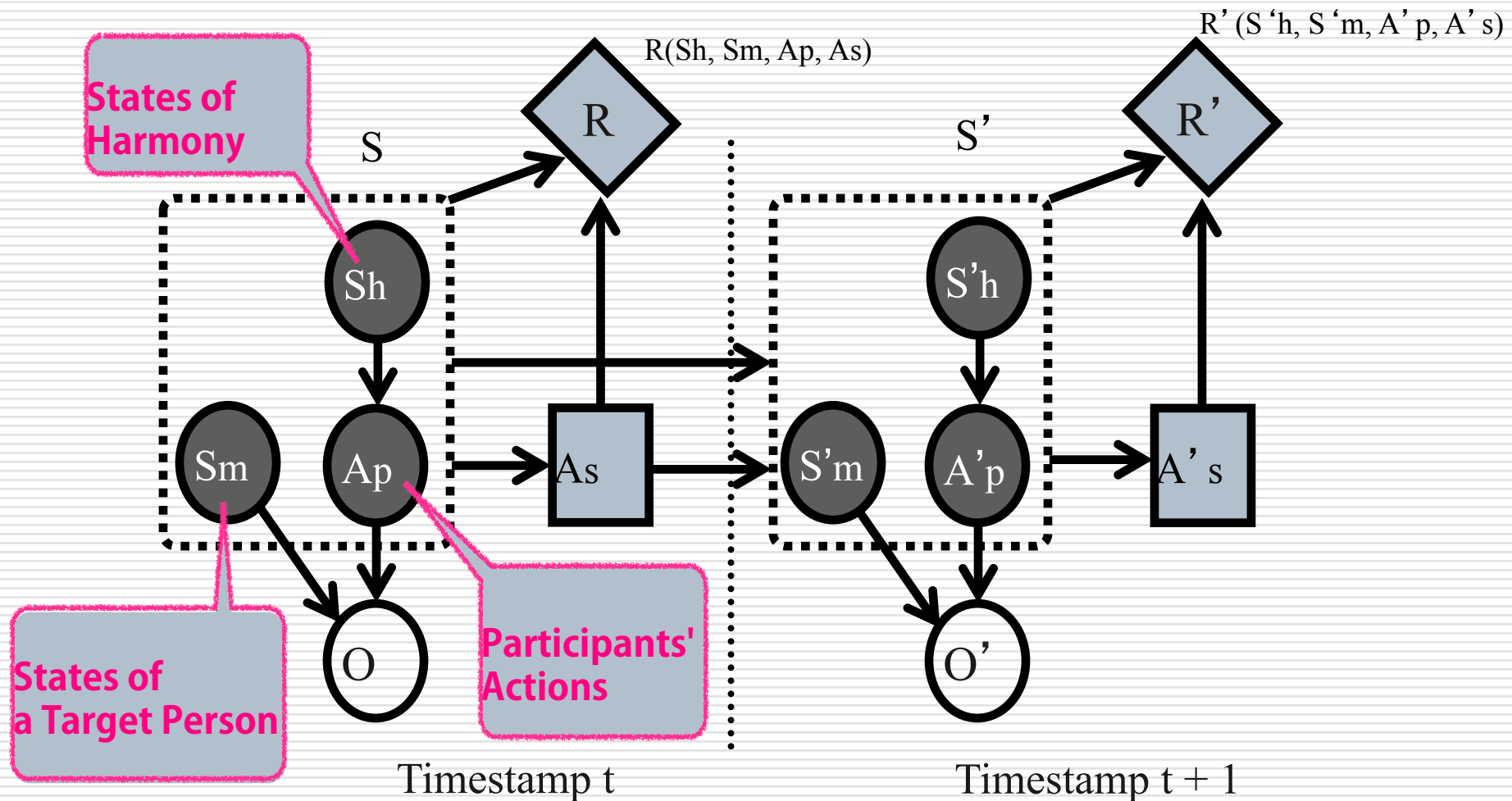
Control Strategies

POMDP: Partially Observable Markov Decision Process



Control Strategies

POMDP: Partially Observable Markov Decision Process



Belief State Update

Observation model

Motivation model

Participants' action model

$$b'(s'_m, a'_m) = \eta \cdot P(o' | s'_m, a'_p, a_s) \cdot P(s'_m | a_s) \cdot \sum_{a_p} P(a'_p | s'_h, a_p, a_s) \cdot \sum_{s_h} P(s'_h | s_h, a_p, a_s) \cdot b(s_m, a_m)$$

Harmony model

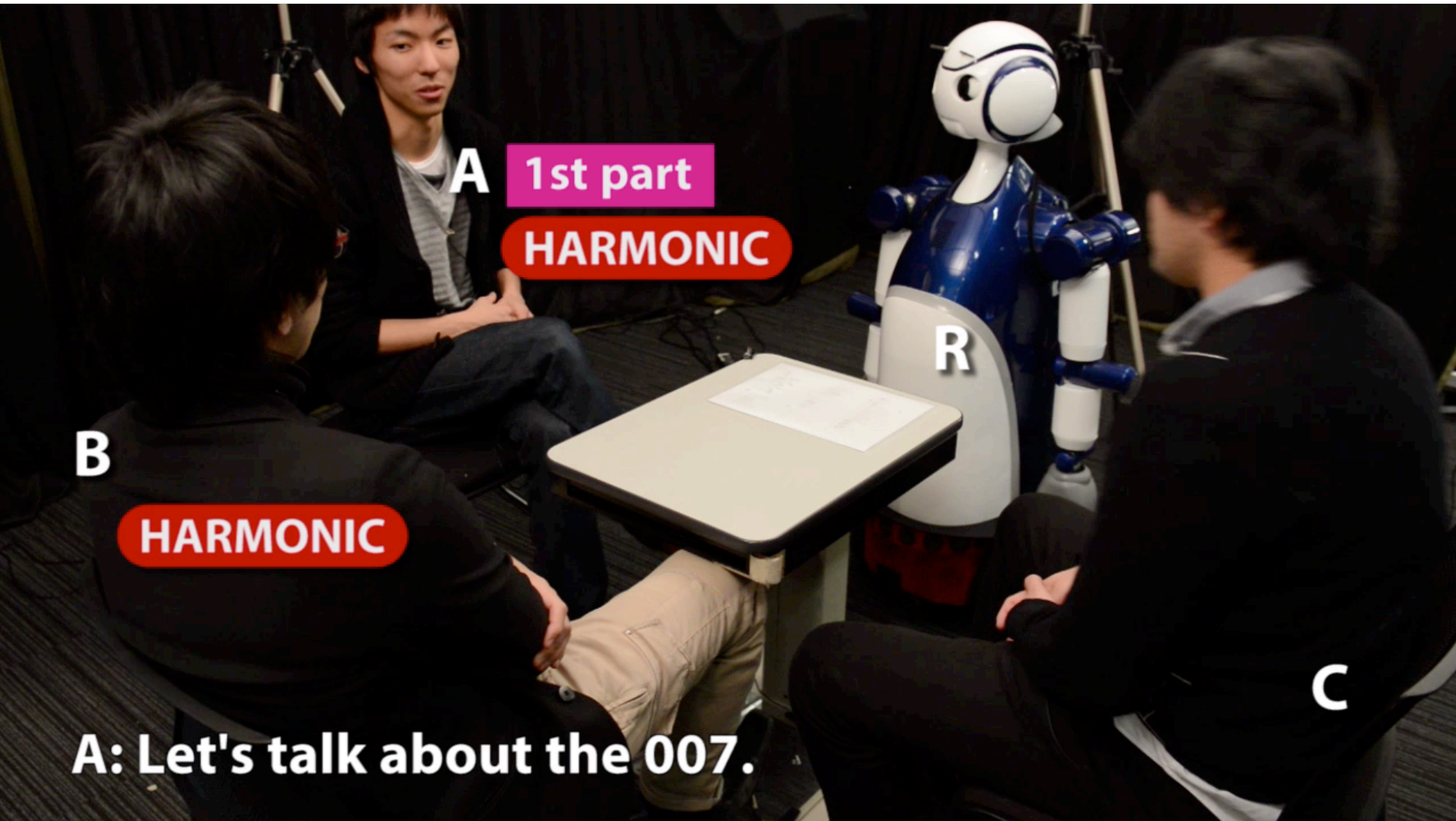
s_h : State of harmony

s_m : State of a target participant

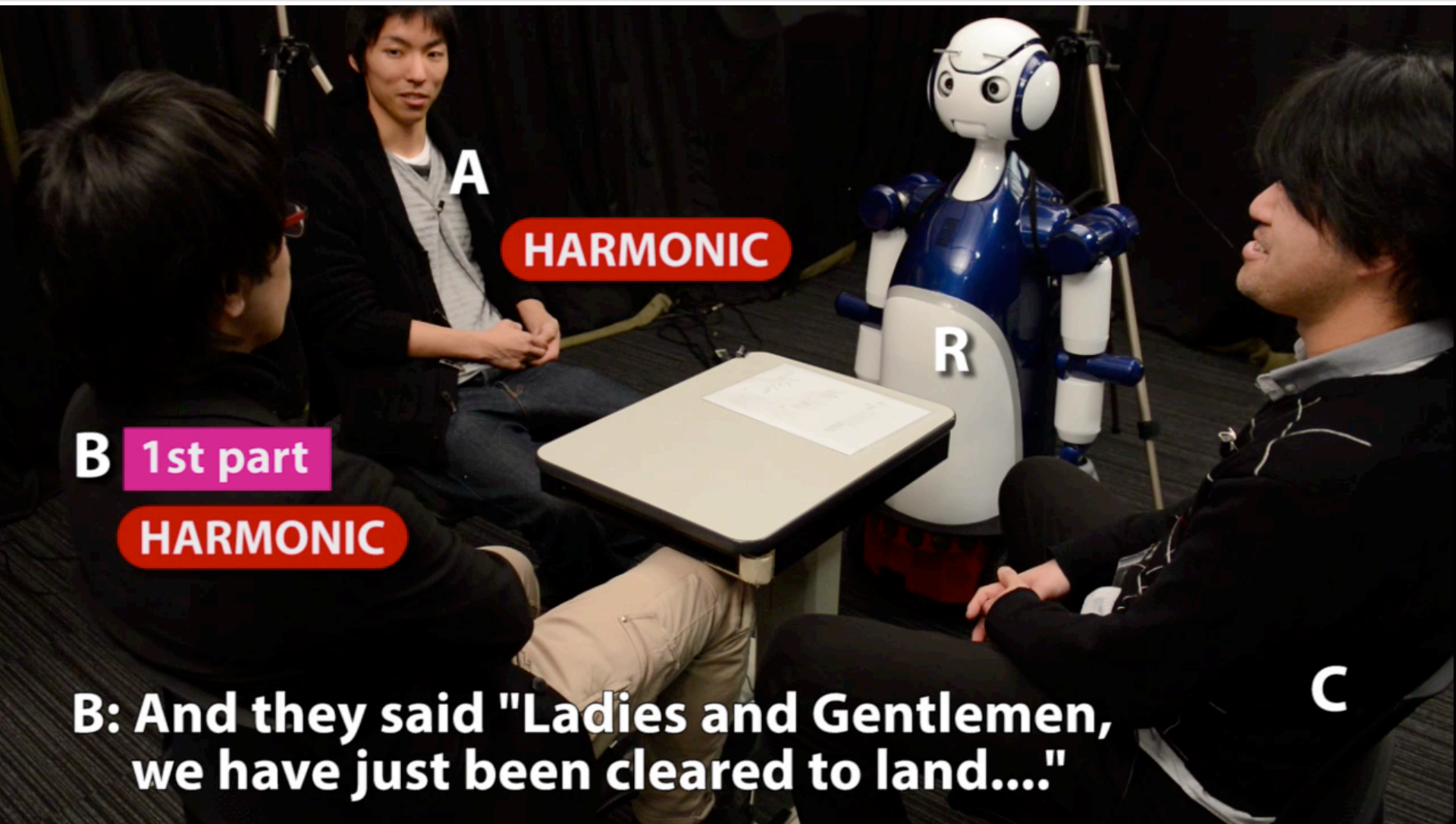
a_s : System action

a_p : Participants' action

All participants be “HARMONIC”



All participants be "HARMONIC"



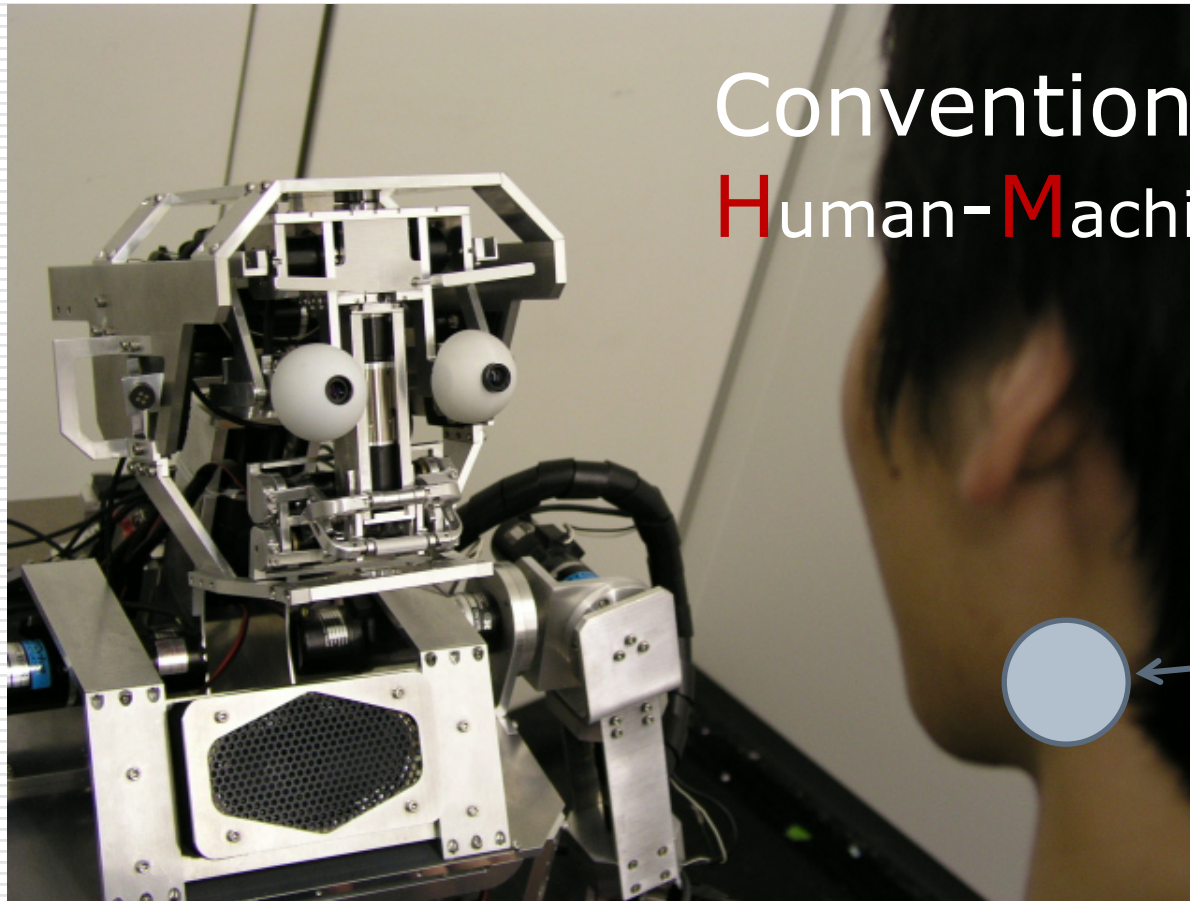
Why Robot?

Role of Body

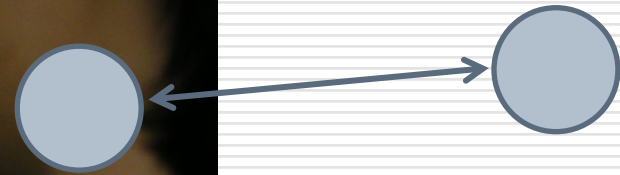
Role of eye gaze

Role of para-linguistic information

Old style HMC



Conventional
Human-**M**achine **C**onversation



2 node communication
with ***FIXED CHANNEL***

Group conversation

Human-Machine Conversation
in a group situation



MULTI node communication
with *PUBLIC NETWORK*

Conversation is COMMUNICATION

*Conversational
Communication
in a group*

*Data
Communication
in a network*

Linguistic Info.

Message

Facilitation:
Is situation good?

Application

Turn control:
Who has right to speak?

Session Protocol

Reliable & speedy transmission:
Did transmission succeed/not?

Transport

Participation structure:
Who join in the conversation?
Who is speaker? Who is addressee?

Data link

Conversation is COMMUNICATION

*Conversational
Communication
in a group*

*Data
Communication
in a network*

Linguistic Info.

Message

*Visual Information of BODY
is important to keep PROTOCOL*

Facilitation:

Is situation good?

Turn centre

Eye gaze
Body

Reli

*Conversation needs human-like BODY
as the physical-layer protocol*

Participation struct

Who is in conversation?
Who is speaker? Who is addressee?

Body & Face directions

Data link

Visit our facebook page at

[https://www.facebook.com/](https://www.facebook.com/PerceptualComputingLab)

[PerceptualComputingLab](https://www.facebook.com/PerceptualComputingLab)

... and push “like it”

REFERENCES:

<Overall>

Tetsunori Kobayashi, Shinya Fujie, “Conversational Robots: An approach to conversation protocol issues that utilizes the paralinguistic information available in a robot-human setting,” [Acoustical Science and Technology](#), 34(2), pp. 64-72 (2013).

<Day care center domo>

Y. Matsuyama, S. Fujie, H. Taniyama, and T. Kobayashi, "Psychological evaluation of a group communication activation robot in a party game," *Proc. Interspeech 2010*, pp. 3046-3049, Sept. 2010.

<Enjoyable conversation>

Y. Matsuyama, A. Saito, S. Fujie, and T. Kobayashi, "Automatic Expressive Opinion Sentence Generation for Enjoyable Conversational Systems," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2015.

<4 people conv.>

Y. Matsuyama, I. Akiba, S. Fujie, and T. Kobayashi, "Four-participant group conversation: A facilitation robot controlling engagement density as the fourth participant," *Elsevier, Computer Speech & Language*, 33(1), pp.1-24, Sep. 2015.



Group 1
SPK

Group 1
ADR

Did you watch movies
recently?

Group 1
SPT



Use of Eye Gaze

