Theoretical Foundations of the UML Lecture 9: Communicating Finite-State Machines

Joost-Pieter Katoen

Lehrstuhl für Informatik 2 Software Modeling and Verification Group

moves.rwth-aachen.de/teaching/ws-1718/fuml/

13. November 2017



Outline

- Introduction
- 2 Communicating Finite-State Machines
- Semantics of Communicating Finite-State Machines
- 4 Emptiness Problem for CFMs



Overview

- Introduction
- Communicating Finite-State Machines
- Semantics of Communicating Finite-State Machines
- 4 Emptiness Problem for CFMs





- Consider an MSGs as complete system specifications
 - they describe a full set of possible system scenarios



- Consider an MSGs as complete system specifications
 - they describe a full set of possible system scenarios
- Can we obtain "realisations" that exhibit precisely these scenarios?

- Consider an MSGs as complete system specifications
 - they describe a full set of possible system scenarios
- Can we obtain "realisations" that exhibit precisely these scenarios?
- Map MSGs, i.e., scenarios onto an executable model
 - model each process by a finite-state automaton
 - that communicate via unbounded directed FIFO channels



- Consider an MSGs as complete system specifications
 - they describe a full set of possible system scenarios
- Can we obtain "realisations" that exhibit precisely these scenarios?
- Map MSGs, i.e., scenarios onto an executable model
 - model each process by a finite-state automaton
 - that communicate via unbounded directed FIFO channels
- ⇒ This yields Communicating Finite-state Machines



Intuition



The need for synchronisation messages



Overview

- Introduction
- 2 Communicating Finite-State Machines
- 3 Semantics of Communicating Finite-State Machines
- 4 Emptiness Problem for CFMs



Definition

Let

- \bullet \mathcal{P} be a finite set of at least two (sequential) processes
- \bullet \mathcal{C} be a finite set of message contents



Definition

Let

- \bullet \mathcal{P} be a finite set of at least two (sequential) processes
- \bullet \mathcal{C} be a finite set of message contents

Definition (communication actions, channels)

• $Act_p^! := \{!(p,q,a) \mid q \in \mathcal{P} \setminus \{p\}, \ a \in \mathcal{C}\}$ the set of send actions by process p

Definition

Let

- \bullet \mathcal{P} be a finite set of at least two (sequential) processes
- \bullet \mathcal{C} be a finite set of message contents

- $Act_p^! := \{!(p,q,a) \mid q \in \mathcal{P} \setminus \{p\}, \ a \in \mathcal{C}\}$ the set of send actions by process p
- $Act_p^? := \{?(p, q, a) \mid q \in \mathcal{P} \setminus \{p\}, \ a \in \mathcal{C}\}$ the set of receive actions by process p

Definition

Let

- \bullet \mathcal{P} be a finite set of at least two (sequential) processes
- \bullet \mathcal{C} be a finite set of message contents

- $Act_p^! := \{!(p,q,a) \mid q \in \mathcal{P} \setminus \{p\}, \ a \in \mathcal{C}\}$ the set of send actions by process p
- $Act_p^? := \{?(p, q, a) \mid q \in \mathcal{P} \setminus \{p\}, \ a \in \mathcal{C}\}$ the set of receive actions by process p
- $\bullet \ Act_p := Act_p^! \cup Act_p^?$



Definition

Let

- be a finite set of at least two (sequential) processes P
- C be a finite set of message contents

- $Act_n! = \{!(p,q,a) \mid q \in \mathcal{P} \setminus \{p\}, \ a \in \mathcal{C}\}$ the set of send actions by process p
- $Act_n^? := \{?(p,q,a) \mid q \in \mathcal{P} \setminus \{p\}, \ a \in \mathcal{C}\}$ the set of receive actions by process p
- $Act_p := Act_p^! \cup Act_p^?$
- $Act := \bigcup_{p \in \mathcal{P}} Act_p$



Definition

Let

- \bullet \mathcal{P} be a finite set of at least two (sequential) processes
- \bullet \mathcal{C} be a finite set of message contents

- $Act_p^! := \{!(p,q,a) \mid q \in \mathcal{P} \setminus \{p\}, \ a \in \mathcal{C}\}$ the set of send actions by process p
- $Act_p^? := \{?(p, q, a) \mid q \in \mathcal{P} \setminus \{p\}, \ a \in \mathcal{C}\}$ the set of receive actions by process p
- $\bullet \ Act_p := Act_p^! \cup Act_p^?$
- $Act := \bigcup_{p \in \mathcal{P}} Act_p$
- $Ch := \{(p,q) \mid p,q \in \mathcal{P}, p \neq q\}$ "channels"



Definition

A communicating finite-state machine (CFM) over \mathcal{P} and \mathcal{C} is a structure

$$\mathcal{A} = (((S_p, \Delta_p))_{p \in \mathcal{P}}, \mathbb{D}, s_{init}, F)$$

where



Definition

A communicating finite-state machine (CFM) over \mathcal{P} and \mathcal{C} is a structure

$$\mathcal{A} = (((S_p, \Delta_p))_{p \in \mathcal{P}}, \mathbb{D}, s_{init}, F)$$

where

• D is a nonempty finite set of synchronization messages (or data)

We often write $s \xrightarrow{\sigma,m}_{n} s'$ instead of $(s,\sigma,m,s') \in \Delta_{p}$



Definition

A communicating finite-state machine (CFM) over \mathcal{P} and \mathcal{C} is a structure

$$\mathcal{A} = (((S_p, \Delta_p))_{p \in \mathcal{P}}, \mathbb{D}, s_{init}, F)$$

where

- D is a nonempty finite set of synchronization messages (or data)
- for each $p \in \mathcal{P}$:
 - S_p is a non-empty finite set of local states (the S_p are disjoint)
 - $\Delta_p \subseteq S_p \times Act_p \times \mathbb{D} \times S_p$ is a set of local transitions

We often write $s \xrightarrow{\sigma,m}_p s'$ instead of $(s,\sigma,m,s') \in \Delta_p$



Definition

A communicating finite-state machine (CFM) over \mathcal{P} and \mathcal{C} is a structure

$$\mathcal{A} = (((S_p, \Delta_p))_{p \in \mathcal{P}}, \mathbb{D}, s_{init}, F)$$

where

- D is a nonempty finite set of synchronization messages (or data)
- for each $p \in \mathcal{P}$:
 - S_p is a non-empty finite set of local states (the S_p are disjoint)
 - $\Delta_p \subseteq S_p \times Act_p \times \mathbb{D} \times S_p$ is a set of local transitions
- $s_{init} \in S_A$ is the global initial state
 - where $S_{\mathcal{A}} := \prod_{p \in \mathcal{P}} S_p$ is the set of global states of \mathcal{A}

We often write $s \xrightarrow{\sigma,m}_{p} s'$ instead of $(s,\sigma,m,s') \in \Delta_{p}$



Definition

A communicating finite-state machine (CFM) over \mathcal{P} and \mathcal{C} is a structure

$$\mathcal{A} = (((S_p, \Delta_p))_{p \in \mathcal{P}}, \mathbb{D}, s_{init}, F)$$

where

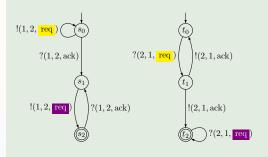
- D is a nonempty finite set of synchronization messages (or data)
- for each $p \in \mathcal{P}$:
 - S_p is a non-empty finite set of local states (the S_p are disjoint)
 - $\Delta_p \subseteq S_p \times Act_p \times \mathbb{D} \times S_p$ is a set of local transitions
- $s_{init} \in S_A$ is the global initial state
 - where $S_{\mathcal{A}} := \prod_{p \in \mathcal{P}} S_p$ is the set of global states of \mathcal{A}
- $F \subseteq S_A$ is the set of global final states

We often write $s \xrightarrow{\sigma,m}_{p} s'$ instead of $(s,\sigma,m,s') \in \Delta_{p}$



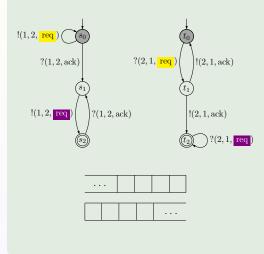


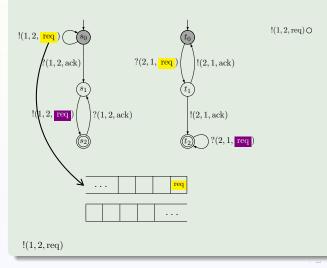
Example

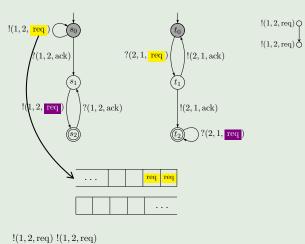


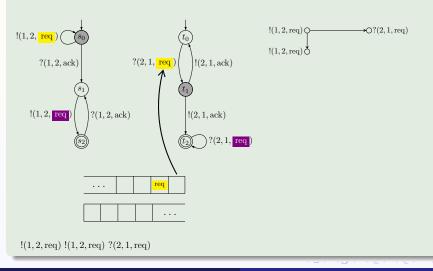
CFM \mathcal{A} over $\mathcal{P} = \{1, 2\}$ and $\mathcal{C} = \{\text{req, ack}\}$

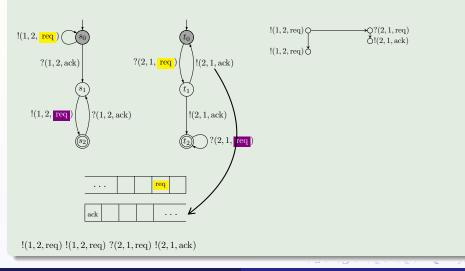
- $\bullet \mathbb{D} = \{ _, _, _ \}$
- $S_1 = \{s_0, s_1, s_2\}$
- $S_2 = \{t_0, t_1, t_2\}$
- $s_{init} = (s_0, t_0)$
- $F = \{(s_2, t_2)\}$

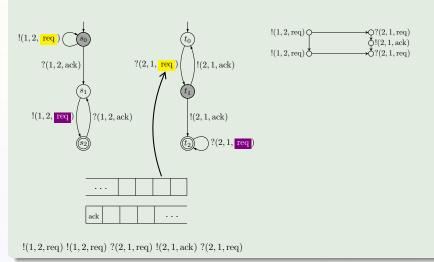


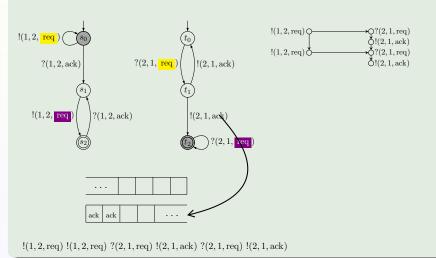


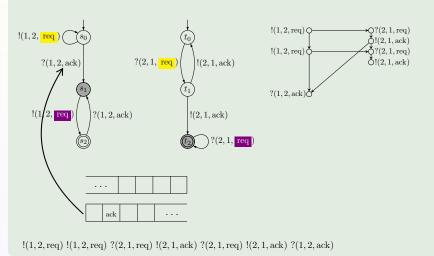


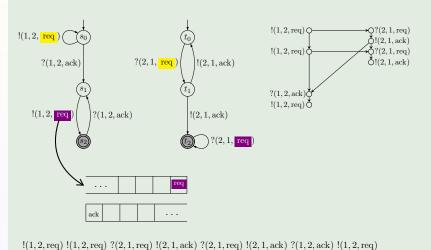


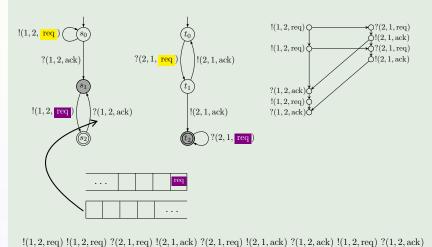




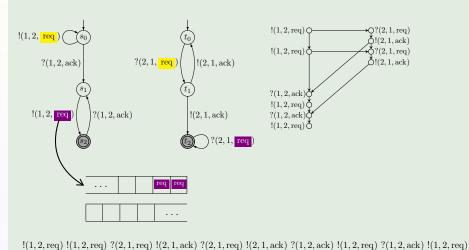






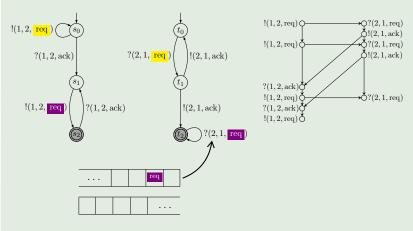


Example



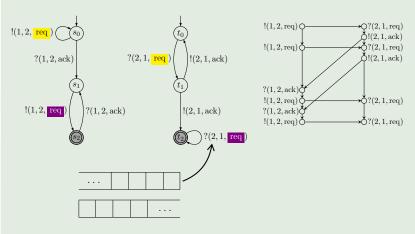
Joost-Pieter Katoen Theoretical Foundations of the UML

Example



 $!{(1,2,\mathrm{req})} \ !{(1,2,\mathrm{req})} \ ?{(2,1,\mathrm{req})} \ !{(2,1,\mathrm{ack})} \ ?{(2,1,\mathrm{req})} \ !{(2,1,\mathrm{ack})} \ ?{(1,2,\mathrm{ack})} \ !{(1,2,\mathrm{req})} \ ?{(1,2,\mathrm{ack})} \ ?{(1,2,\mathrm{ack})} \ !{(1,2,\mathrm{req})} \ ?{(1,2,\mathrm{ack})} \ !{(1,2,\mathrm{ack})} \ !{(1,2,\mathrm{ack})} \ ?{(1,2,\mathrm{ack})} \ !{(1,2,\mathrm{ack})} \ ?{(1,2,\mathrm{ack})} \ !{(1,2,\mathrm{ack})} \ ?{(1,2,\mathrm{ack})} \ ?{(1,2,\mathrm{a$

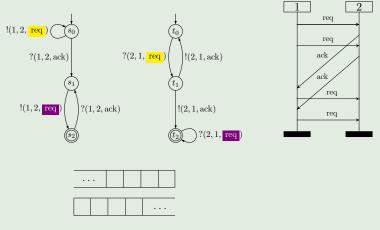
Example



 $!{(1,2,\mathrm{req})} \ !{(1,2,\mathrm{req})} \ ?{(2,1,\mathrm{req})} \ !{(2,1,\mathrm{ack})} \ ?{(2,1,\mathrm{req})} \ !{(2,1,\mathrm{ack})} \ ?{(1,2,\mathrm{ack})} \ !{(1,2,\mathrm{req})} \ ?{(1,2,\mathrm{ack})} \ !{(1,2,\mathrm{ack})} \ !{(1,2,\mathrm{a$

Communicating finite-state machines

Example



 $!{(1,2,\mathrm{req})} \ !{(1,2,\mathrm{req})} \ ?{(2,1,\mathrm{req})} \ !{(2,1,\mathrm{ack})} \ ?{(2,1,\mathrm{req})} \ !{(2,1,\mathrm{ack})} \ ?{(1,2,\mathrm{ack})} \ !{(1,2,\mathrm{req})} \ ?{(1,2,\mathrm{ack})} \ ?{(1,2,\mathrm{ack})} \ !{(1,2,\mathrm{ack})} \ ?{(1,2,\mathrm{ack})} \ !{(1,2,\mathrm{ack})} \ ?{(1,2,\mathrm{ack})} \ !{(1,2,\mathrm{ack})} \ ?{(1,2,\mathrm{ack})} \ ?{(1,2,\mathrm{a$

Overview

- Introduction
- Communicating Finite-State Machines
- 3 Semantics of Communicating Finite-State Machines
- 4 Emptiness Problem for CFMs



Theoretical Foundations of the UML

Let $\mathcal{A} = (((S_p, \Delta_p))_{p \in \mathcal{P}}, \mathbb{D}, s_{init}, F)$ be a CFM over \mathcal{P} and \mathcal{C} .

Definition (configurations)

Configurations of A: $Conf_A := S_A \times \{ \eta \mid \eta : Ch \to (\mathcal{C} \times \mathbb{D})^* \}$



Let $\mathcal{A} = (((S_p, \Delta_p))_{p \in \mathcal{P}}, \mathbb{D}, s_{init}, F)$ be a CFM over \mathcal{P} and \mathcal{C} .

Definition (configurations)

Configurations of $A: Conf_A := S_A \times \{ \eta \mid \eta : Ch \to (\mathcal{C} \times \mathbb{D})^* \}$

Definition (global step)

 $\Longrightarrow_{\mathcal{A}} \subseteq Conf_{\mathcal{A}} \times Act \times \mathbb{D} \times Conf_{\mathcal{A}}$ is defined as follows:

Let $\mathcal{A} = (((S_p, \Delta_p))_{p \in \mathcal{P}}, \mathbb{D}, s_{init}, F)$ be a CFM over \mathcal{P} and \mathcal{C} .

Definition (configurations)

Configurations of $A: Conf_{\mathcal{A}} := S_{\mathcal{A}} \times \{ \eta \mid \eta : Ch \to (\mathcal{C} \times \mathbb{D})^* \}$

Definition (global step)

 $\Longrightarrow_{\mathcal{A}} \subseteq Conf_{\mathcal{A}} \times Act \times \mathbb{D} \times Conf_{\mathcal{A}}$ is defined as follows:

- sending a message: $((\overline{s}, \eta), !(p, q, a), m, (\overline{s}', \eta')) \in \Longrightarrow_{\mathcal{A}} if$
 - $\bullet \ (\overline{s}[p],!(p,q,a),m,\overline{s}'[p]) \in \Delta_{p}$
 - $\eta' = \eta[(p, q) := (a, m) \cdot \eta((p, q))]$
 - $\overline{s}[r] = \overline{s}'[r]$ for all $r \in \mathcal{P} \setminus \{p\}$

Let $\mathcal{A} = (((S_p, \Delta_p))_{p \in \mathcal{P}}, \mathbb{D}, s_{init}, F)$ be a CFM over \mathcal{P} and \mathcal{C} .

Definition (configurations)

Configurations of $A: Conf_{\mathcal{A}} := S_{\mathcal{A}} \times \{ \eta \mid \eta : Ch \to (\mathcal{C} \times \mathbb{D})^* \}$

Definition (global step)

 $\Longrightarrow_{\mathcal{A}} \subseteq Conf_{\mathcal{A}} \times Act \times \mathbb{D} \times Conf_{\mathcal{A}}$ is defined as follows:

- sending a message: $((\overline{s}, \eta), !(p, q, a), m, (\overline{s}', \eta')) \in \Longrightarrow_{\mathcal{A}} if$
 - $(\overline{s}[p], !(p, q, a), m, \overline{s}'[p]) \in \Delta_p$
 - $\eta' = \eta[(p, q) := (a, m) \cdot \eta((p, q))]$
 - $\overline{s}[r] = \overline{s}'[r]$ for all $r \in \mathcal{P} \setminus \{p\}$
- receipt of a message: $((\overline{s}, \eta), ?(p, q, a), m, (\overline{s}', \eta')) \in \Longrightarrow_{\mathcal{A}} if$
 - $(\overline{s}[p], ?(p, q, a), m, \overline{s}'[p]) \in \Delta_p$
 - $\eta((q, \mathbf{p})) = w \cdot (a, m) \neq \epsilon \text{ and } \eta' = \eta[(q, \mathbf{p}) := w]$
 - $\overline{s}[r] = \overline{s}'[r]$ for all $r \in \mathcal{P} \setminus \{p\}$

Example



Linearizations of a CFM

Let $\mathcal{A} = (((S_p, \Delta_p))_{p \in \mathcal{P}}, \mathbb{D}, s_{init}, F)$ be a CFM over \mathcal{P} and \mathcal{C} .

Definition (accepting runs)

A run ρ of CFM \mathcal{A} on word $w = \sigma_1 \dots \sigma_n \in Act^*$ is an alternating sequence $\rho = \gamma_0 m_1 \gamma_1 \dots \gamma_{n-1} m_n \gamma_n$ such that

- $\gamma_0 = (s_{init}, \eta_{\varepsilon})$ with η_{ε} mapping any channel to ε



Linearizations of a CFM

Let $\mathcal{A} = (((S_n, \Delta_n))_{n \in \mathcal{P}}, \mathbb{D}, s_{init}, F)$ be a CFM over \mathcal{P} and \mathcal{C} .

Definition (accepting runs)

A run ρ of CFM \mathcal{A} on word $w = \sigma_1 \dots \sigma_n \in Act^*$ is an alternating sequence $\rho = \gamma_0 m_1 \gamma_1 \dots \gamma_{n-1} m_n \gamma_n$ such that

- $\gamma_{i-1} \xrightarrow{\sigma_i, m_i} A \gamma_i$ for any $i \in \{1, \dots, n\}$

The run ρ is accepting if $\gamma_n \in F \times \{\eta_{\varepsilon}\}.$



Linearizations of a CFM

Let $\mathcal{A} = (((S_n, \Delta_n))_{n \in \mathcal{P}}, \mathbb{D}, s_{init}, F)$ be a CFM over \mathcal{P} and \mathcal{C} .

Definition (accepting runs)

A run ρ of CFM \mathcal{A} on word $w = \sigma_1 \dots \sigma_n \in Act^*$ is an alternating sequence $\rho = \gamma_0 m_1 \gamma_1 \dots \gamma_{n-1} m_n \gamma_n$ such that

- **1** $\gamma_0 = (s_{init}, \eta_{\varepsilon})$ with η_{ε} mapping any channel to ε
- $\gamma_{i-1} \xrightarrow{\sigma_i, m_i} A \gamma_i$ for any $i \in \{1, \dots, n\}$

The run ρ is accepting if $\gamma_n \in F \times \{\eta_{\varepsilon}\}.$

Definition (linearization of a CFM)

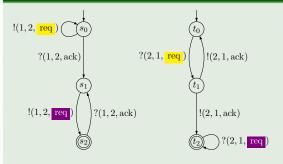
The (word) language of CFM \mathcal{A} is defined by:

 $Lin(\mathcal{A}) := \{ w \in Act^* \mid \text{there is an accepting run of } \mathcal{A} \text{ on } w \}$



Linearizations of an example CFM

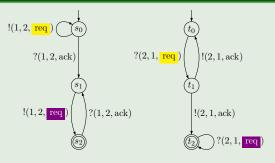
Example



CFM \mathcal{A} over $\{1,2\}$ and $\{\text{req}, \text{ack}\}$

Linearizations of an example CFM

Example



CFM \mathcal{A} over $\{1,2\}$ and $\{req,ack\}$

$$Lin(\mathcal{A}) = \left\{ w \in Act^* \mid \text{there is } n \geqslant 1 \text{ such that:} \right.$$

$$w \upharpoonright 1 = !(1, 2, \text{req}))^n \ (?(1, 2, \text{ack}) \ !(1, 2, \text{req}))^n$$

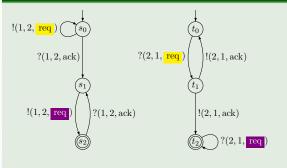
 $w \upharpoonright 2 = (?(2, 1, \text{req}) \ !(2, 1, \text{ack}))^n \ (?(2, 1, \text{req}))^n$

for any $u \in Pref(w)$ and $(p,q) \in Ch$:

$$\sum_{a \in C} |u|_{!(p,q,a)} - \sum_{a \in C} |u|_{?(q,p,a)} \geqslant 0$$

Linearizations of an example CFM

Example

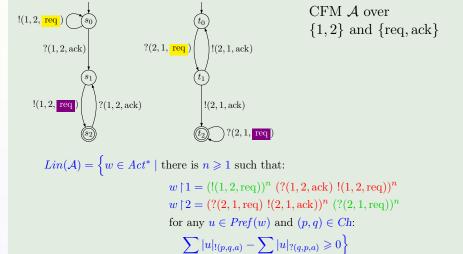


CFM \mathcal{A} over $\{1,2\}$ and $\{\text{req}, \text{ack}\}$

- \bullet !(1, 2, req) and !(2, 1, ack) are always independent.
- \bullet !(1, 2, req) and ?(1, 2, ack) are always dependent.
- \bullet !(1, 2, req) and ?(2, 1, req) are sometimes independent.
- → non-regular (word) languages

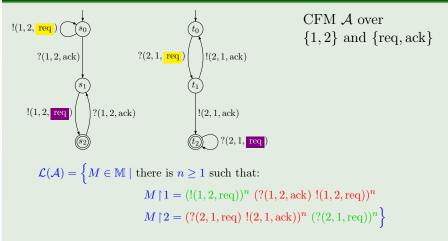
Linearizations and MSCs of an example CFM

Example



Linearizations and MSCs of an example CFM

Example



Overview

- Introduction
- Communicating Finite-State Machines
- 3 Semantics of Communicating Finite-State Machines
- 4 Emptiness Problem for CFMs



Elementary questions are undecidable for CFMs

Emptiness of CFMs is undecidable

[Brand & Zafiropulo 1983]

The following problem is undecidable (even if C is a singleton):

INPUT: CFM \mathcal{A} over processes \mathcal{P} and message contents \mathcal{C}

QUESTION: Is $\mathcal{L}(\mathcal{A})$ empty?



Elementary questions are undecidable for CFMs

Emptiness of CFMs is undecidable

[Brand & Zafiropulo 1983]

The following problem is undecidable (even if C is a singleton):

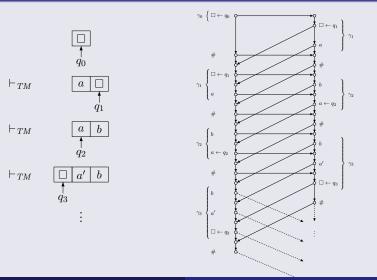
INPUT: CFM \mathcal{A} over processes \mathcal{P} and message contents \mathcal{C}

QUESTION: Is $\mathcal{L}(\mathcal{A})$ empty?

Proof (sketch)

Reduction from the halting problem for Turing machine $TM = (Q, \Sigma, \Delta, \square, q_0, q_f)$ to emptiness for a CFM with two processes. Build CFM $\mathcal{A} = ((\mathcal{A}_1, \mathcal{A}_2), \mathbb{D}, s_{init}, F)$ over $\{1, 2\}$ and some singleton set \mathcal{C} such that $\mathcal{L}(\mathcal{A}) \neq \emptyset$ iff TM can reach q_f , i.e., TM accepts.

- Process 1 sends current configurations to process 2
- ullet Process 2 chooses successor configurations and sends them to 1
- $\bullet \ \mathbb{D} = \Big((\Sigma \cup \{\Box\}) \times (Q \cup \{_\}) \Big) \cup \{\#\}$



- Left or standstill transition: Process 2 may just wait for a symbol containing a state of TM and to alter it correspondingly. In the example, the left-moving transition (q_2, a, a', L, q_3) is applied so that process 2
 - ullet sends b unchanged back to process 1
 - detects (receives) $a \leftarrow q_2$
 - sends a' to process 1 entering a state indicating that the symbol to be sent next has to be equipped with q_3
 - receives # so that the symbol $\square \leftarrow q_3$ has to be inserted before returning #

- Left or standstill transition: Process 2 may just wait for a symbol containing a state of TM and to alter it correspondingly. In the example, the left-moving transition (q_2, a, a', L, q_3) is applied so that process 2
 - \bullet sends b unchanged back to process 1
 - detects (receives) $a \leftarrow q_2$
 - sends a' to process 1 entering a state indicating that the symbol to be sent next has to be equipped with q_3
 - receives # so that the symbol $\square \leftarrow q_3$ has to be inserted before returning #
- Right transition: Process 2 has to guess what the position right before the head is. For example, provided process 2 decided in favor of (q_2, a, a', R, q_3) while reading b, it would have to
 - send $b \leftarrow q_3$ instead of just b, entering some state $t(a \leftarrow q_2)$
 - receive $a \leftarrow q_2$ (no other symbol can be received in state $t(a \leftarrow q_2)$)
 - send a' back to process 1

Proof (contd.)

• Introduce local final states s_f and t_f , one for process 1 and one for process 2, respectively (i.e., $F = \{(s_f, t_f)\}$ and \mathcal{A} is locally accepting).

- Introduce local final states s_f and t_f , one for process 1 and one for process 2, respectively (i.e., $F = \{(s_f, t_f)\}$ and \mathcal{A} is locally accepting).
- At any time, process 1 may switch into s_f , in which arbitrary and arbitrarily many messages can be received to empty channel (2,1).



- Introduce local final states s_f and t_f , one for process 1 and one for process 2, respectively (i.e., $F = \{(s_f, t_f)\}$ and \mathcal{A} is locally accepting).
- At any time, process 1 may switch into s_f , in which arbitrary and arbitrarily many messages can be received to empty channel (2,1).
- Process 2 is allowed to move into t_f and to empty the channel (1,2) as soon as it receives a letter $c \leftarrow q_f$ for some c.



- Introduce local final states s_f and t_f , one for process 1 and one for process 2, respectively (i.e., $F = \{(s_f, t_f)\}$ and \mathcal{A} is locally accepting).
- At any time, process 1 may switch into s_f , in which arbitrary and arbitrarily many messages can be received to empty channel (2,1).
- Process 2 is allowed to move into t_f and to empty the channel (1,2) as soon as it receives a letter $c \leftarrow q_f$ for some c.
- As process 2 modifies a configuration of TM locally, finitely many states are sufficient in A.

