Theoretical Foundations of the UML Lecture 546: Compositional Message Sequence Graphs

Joost-Pieter Katoen

Lehrstuhl für Informatik 2 Software Modeling and Verification Group

moves.rwth-aachen.de/teaching/ss-20/fuml/

May 5, 2020

Joost-Pieter Katoen Theoretical Foundations of the UML

ヘロト ヘヨト ヘヨト

э

Outline

A non-decomposable MSC

- 2 Compositional Message Sequence Charts
- 3 Compositional Message Sequence Graphs
- 4 Safe Compositional Message Sequence Graphs
- Existence of Safe Paths
 two decision poblems
 Universality of Safe Paths
 decideble

э

Compositional MSCs

Solution: drop restriction that e and m(e) belong to the same MSC (= allow for incomplete message transfer)

Definition (Compositional MSC)

 $M = (\mathcal{P}, E, \mathcal{C}, l, m, \preceq)$ is a compositional MSC (CMSC, for short) where $\mathcal{P}, E, \mathcal{C}$ and l are defined as before, and

• $m : E_! \to E_?$ is a partial, injective function such that (as before):

$$m(e) = e' \wedge l(e) = !(p,q,a) \quad \text{implies} \quad l(e') = ?(q,p,a)$$

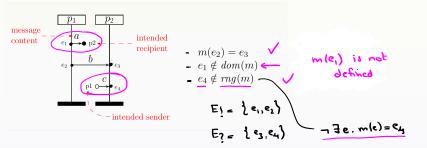
$$\stackrel{\bullet}{\preceq} = \left(\bigcup_{p \in \mathcal{P}} <_{p} \qquad \cup \qquad \{(e,m(e)) \mid e \in \underbrace{dom(m)}_{\text{domain of } m} \} \right)^{*}$$

$$\underbrace{domain \text{ of } m}_{\text{``m}(e) \text{ is defined''}}$$

Note:

An \underline{MSC} is a CMSC where m is total and bijective.

Joost-Pieter Katoen Theoretical Foundations of the UML



Joost-Pieter Katoen Theoretical Foundations of the UML



2

Paths

Let
$$G = (V, \rightarrow, v_0, F, \underline{\lambda})$$
 be a CMSG.

Definition (Path in a CMSG)

A path π of G is a finite sequence

$$\pi = u_0 u_1 \dots u_n$$
 with $u_i \in V$ $(0 \le i \le n)$ and $u_i \to u_{i+1}$ $(0 \le i < n)$

 $\lambda: \vee \longrightarrow \mathbb{CM}$

Definition (Accepting path of a CMSG)

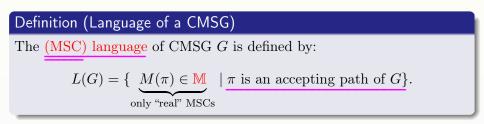
Path $\pi = u_0 \ldots u_n$ is accepting if: $u_0 = v_0$ and $u_n \in F$.

Definition (CMSC of a path)

The CMSC of a path $\pi = u_0 \dots u_n$ is:

$$M(\pi) = (\dots (\lambda(u_0) \bullet \lambda(u_1)) \bullet \lambda(u_2) \dots) \bullet \lambda(u_n)$$

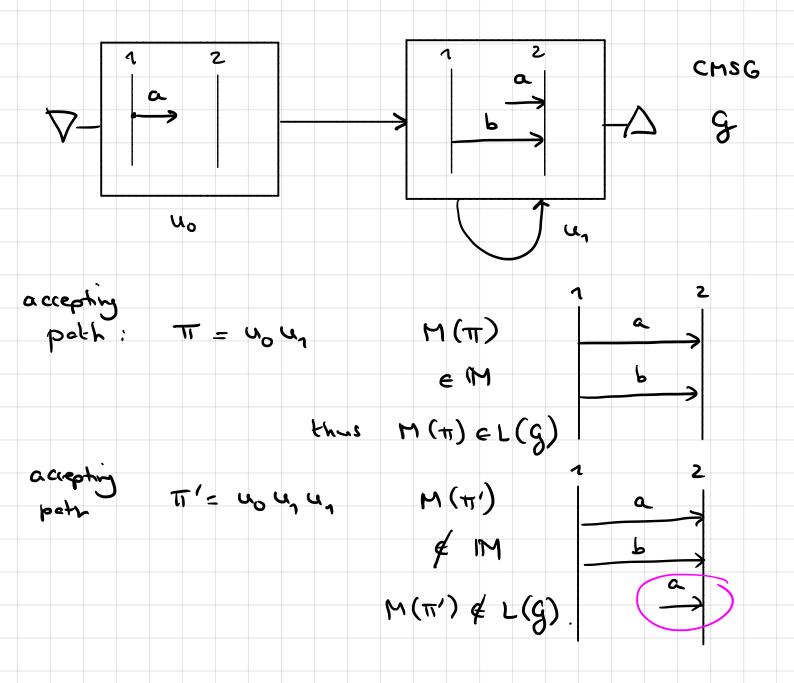
where CMSC concatenation is left associative.



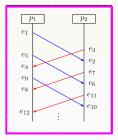
Note: Accepting paths that give rise to an CMSC (which is not an MSC) are not part of L(G).

Joost-Pieter Katoen Theoretical Foundations of the UML

▲ロト ▲園 ト ▲ 臣 ト ▲ 臣 ト ● ○ ● ● ●



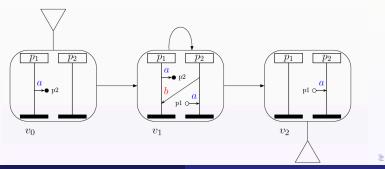
Yannakakis' example as compositional MSG



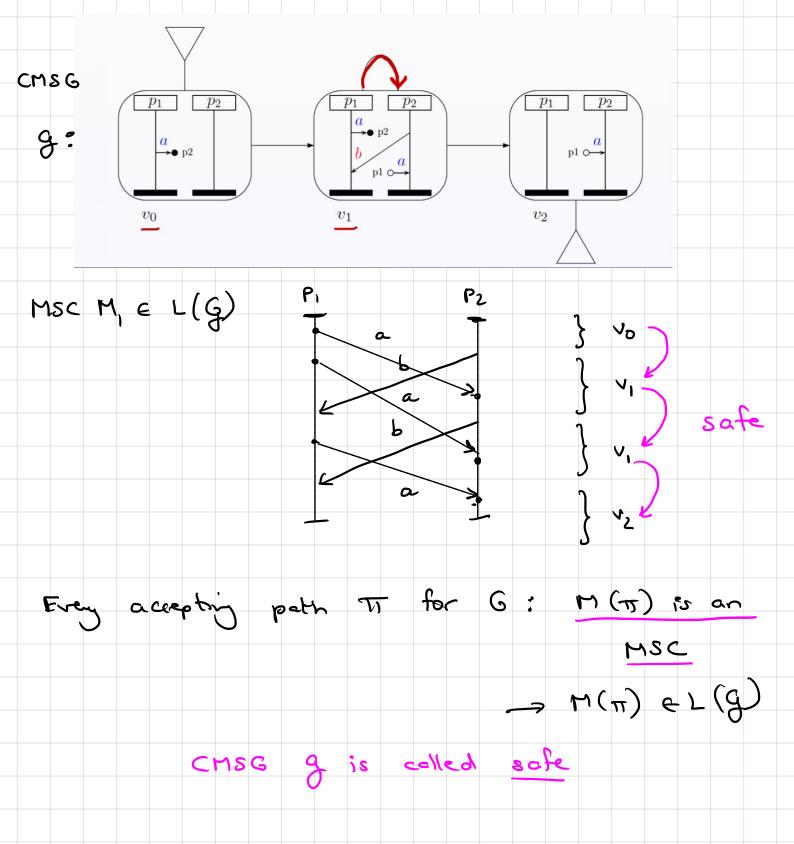
This MSC cannot be modeled for n > 1 by:

$$M = M_1 \bullet M_2 \bullet \ldots \bullet M_n \quad \text{with} \quad M_i \in \mathbb{M}$$

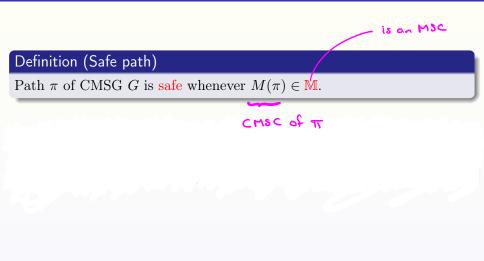
Thus it cannot be modeled by a MSG. But it can be modeled as compositional MSG:



Joost-Pieter Katoen Theoretical Foundations of the UML



Safe paths and CMSGs



Joost-Pieter Katoen Theoretical Foundations of the UML

The decision problem "does CMSG G have at least one safe, accepting path?" is <u>undecidable.</u>

Proof.

By a reduction from Post's Correspondence Problem (PCP).

... black board ...

The complement decision problem "does CMSG G have no safe, accepting path?" is <u>undecidable</u> too.

・ロト ・聞 ト ・ ヨ ト ・ ヨ ト

The decision problem "does CMSG G have at least one safe, accepting path?" is undecidable.

э

The decision problem "does CMSG G have at least one safe, accepting path?" is undecidable.

Theorem: decidability of universality of safe paths

The decision problem "are all accepting paths of CMSG G safe?" is decidable in PTIME.

The decision problem "does CMSG G have at least one safe, accepting path?" is undecidable.

Theorem: decidability of universality of safe paths

The decision problem "are all accepting paths of CMSG G safe?" is decidable in PTIME.

Proof.

Polynomial reduction to reachability problem in (non-deterministic) pushdown automata.

... see details on the next slides ...

- 4 戸 ト - 4 三 ト

Pushdown automata

Definition (Pushdown automaton)

A pushdown automaton (PDA, for short) $K = (Q, q_0, \Gamma, \Sigma, \Delta)$ with

- Q, a finite set of control states
- $q_0 \in Q$, the initial state which symbols can be put
 - Γ , a finite stack alphabet
 - Σ , a finite input alphabet a, b, c

• $\Delta \subseteq Q \times \Sigma \times \Gamma \times Q \times \Gamma^*$, the transition relation.



Definition (Pushdown automaton)

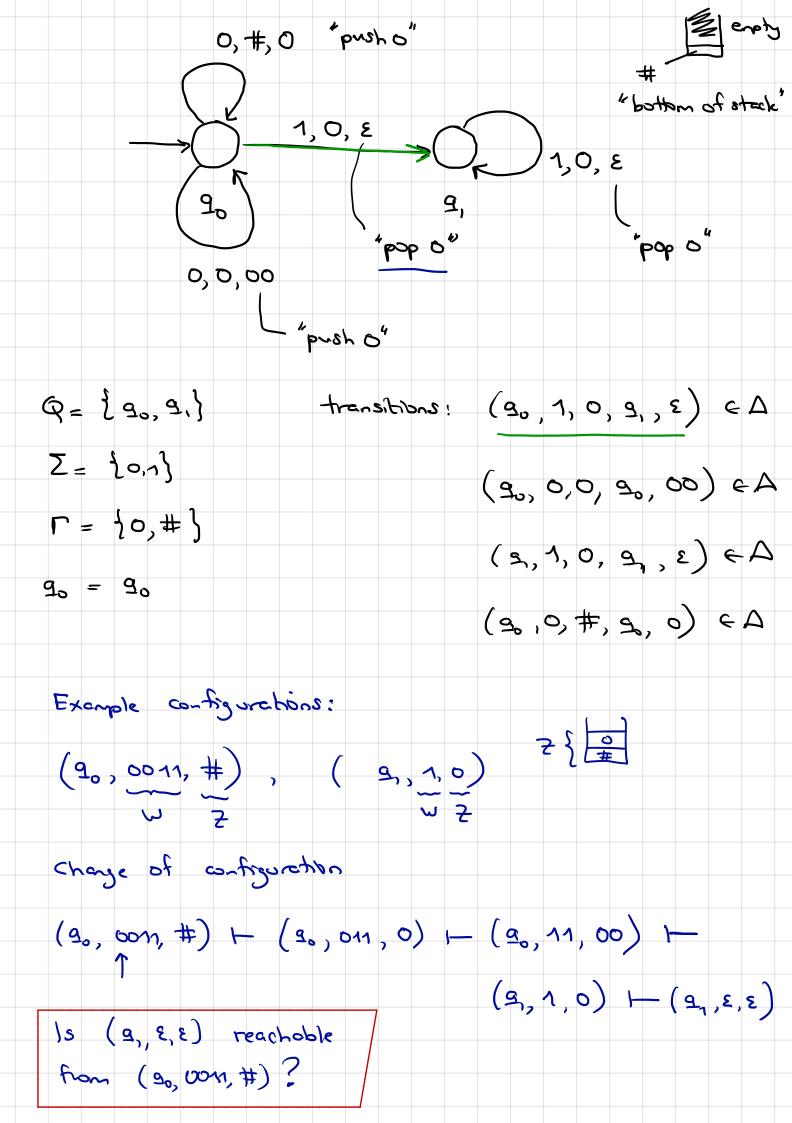
A pushdown automaton (PDA, for short) $K = (Q, q_0, \Gamma, \Sigma, \Delta)$ with

- Q, a finite set of control states
- $q_0 \in Q$, the initial state
- Γ , a finite stack alphabet
- Σ , a finite input alphabet
- $\Delta \subseteq Q \times \Sigma \times \Gamma \times Q$ (Γ^* ,) the transition relation.

Transition relation

 $(q, a, \gamma, q', \text{pop}) \in \Delta$ means: in state q, on reading input symbol a and top of stack is symbol γ , change to q' and pop γ from the stack.

・ロト ・ 母 ト ・ ヨ ト ・ ヨ ト ・ ヨ



Reachability in pushdown automata

Definition

A configuration c is a triple (state q, stack content Z, rest input w).

(atro)

Joost-Pieter Katoen Theoretical Foundations of the UML

æ

▲圖 → ▲ 国 → ▲ 国 →

Definition

A configuration c is a triple (state q, stack content Z, rest input w).

Definition

Given a transition in Δ , a (direct) successor configuration c' of c is obtained: $c \vdash c'$.

イロト 不得下 イヨト イヨト 三日

Definition

A configuration c is a triple (state q, stack content Z, rest input w).

Definition

Given a transition in Δ , a (direct) successor configuration c' of c is obtained: $c \vdash c'$.

Reachability problem

For configuration c, and initial configuration $c_0 \vdash^* c$?

<ロト < 四ト < 回ト < 回ト < 回ト = 三 回

Definition

A configuration c is a triple (state q, stack content Z, rest input w).

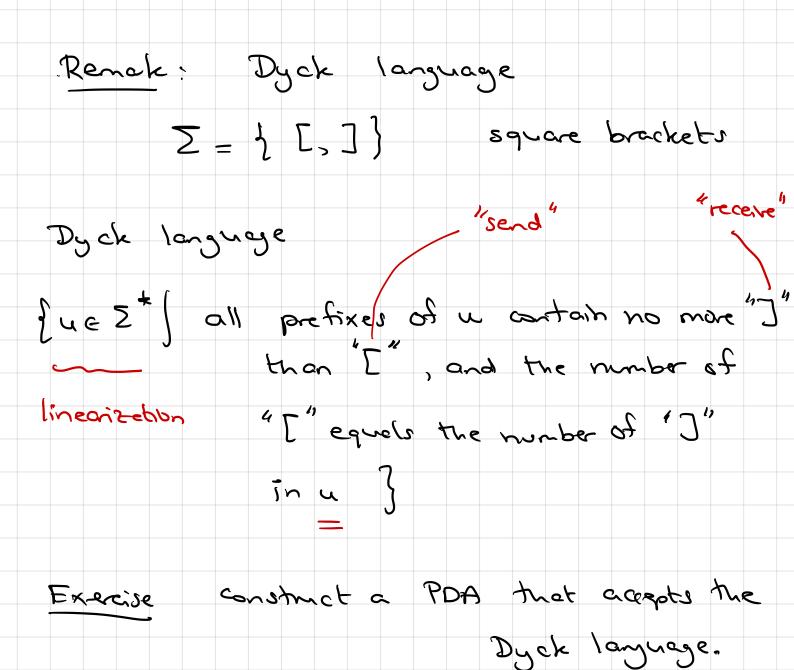
Definition

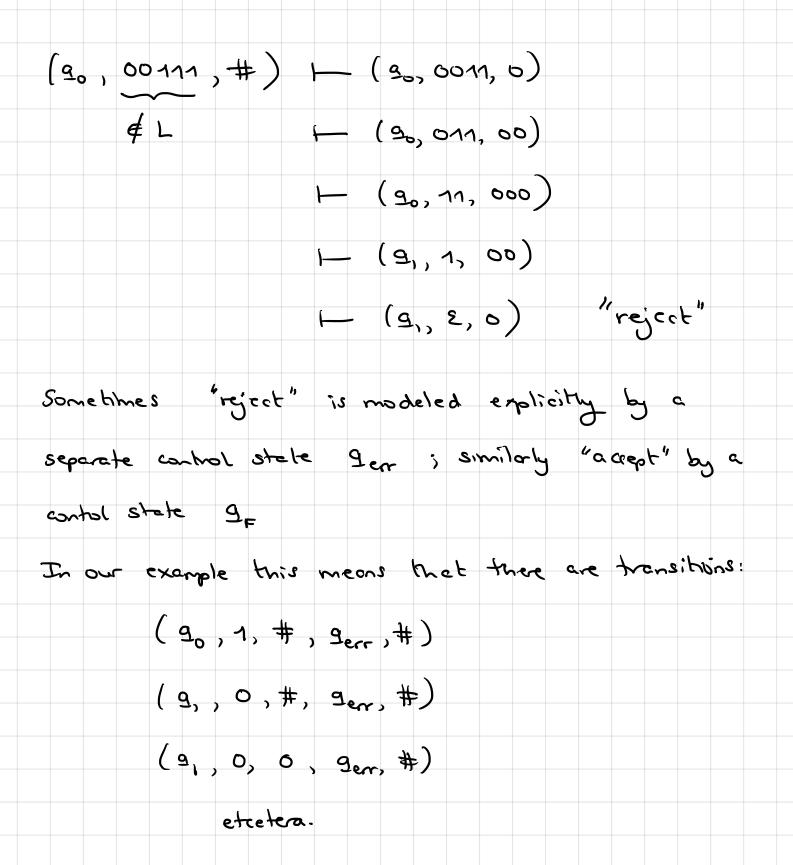
Given a transition in Δ , a (direct) successor configuration c' of c is obtained: $c \vdash c'$.

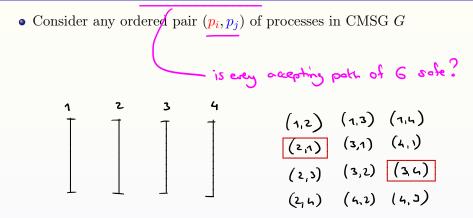
Reachability problem

For configuration c_0 , and initial configuration $c_0: c_0 \vdash^* c$?



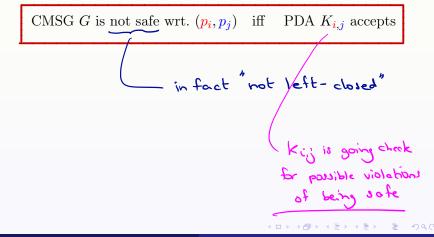


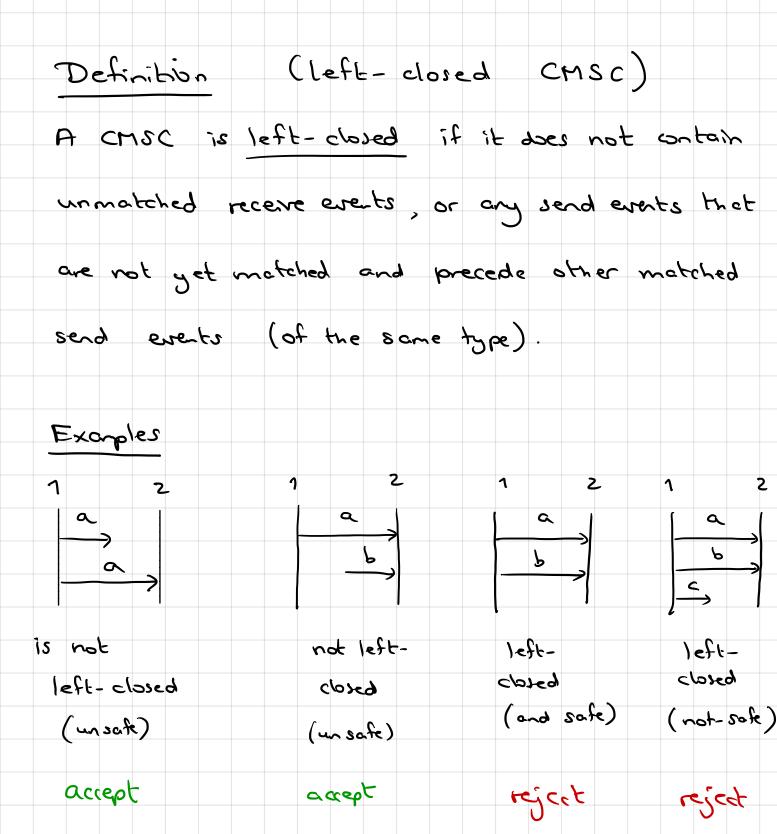




э

- Consider any ordered pair (p_i, p_j) of processes in CMSG G
- Proof idea: construct a PDA $K_{i,j} = (Q, q_0, \Gamma, \Sigma, \Delta)$ such that





- Consider any ordered pair (p_i, p_j) of processes in CMSG G
- Proof idea: construct a PDA $K_{i,j} = (Q, q_0, \Gamma, \Sigma, \Delta)$ such that

CMSG G is not lc wrt. (p_i, p_j) iff PDA $K_{i,j}$ accepts

• For accepting path $u_0 \dots u_k$ in G, feed $K_{i,j}$ with out words

$$\rho_0 \dots \rho_k$$
 where $\rho_i \in Lin(\lambda(u_i))$

such that unmatched sends (of some type) precede all unmatched receipts (of the same type)

+ assume that Linearisation, because such matched events are linearisations do always exist indicated explicitly !?(p,q,a)

- Consider any ordered pair (p_i, p_j) of processes in CMSG G
- Proof idea: construct a PDA $K_{i,j} = (Q, q_0, \Gamma, \Sigma, \Delta)$ such that

CMSG G is not Lc wrt. (p_i, p_j) iff PDA $K_{i,j}$ accepts

• For accepting path $u_0 \ldots u_k$ in G, feed $K_{i,j}$ with all words

 $\rho_0 \dots \rho_k$ where $\rho_i \in Lin(\lambda(u_i))$

such that unmatched sends (of some type) precede all unmatched receipts (of the same type)

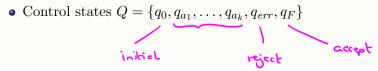
- Possible violations that $K_{i,j}$ may encounter:
- In a general second second

Let $\{a_1, \ldots, a_k\}$ be the message contents in CMSG G for (p_i, p_j) . all message in CHISG g seed from i to j, or received at j from i.

・ロト ・ 一 ト ・ ヨ ト ・ 日 ト

÷

Let $\{a_1, \ldots, a_k\}$ be the message contents in CMSG G for (p_i, p_j) . Nondeterministic PDA $K_{i,j} = (Q, q_0, \Gamma, \Sigma, \Delta)$ where:



э

Let $\{a_1, \ldots, a_k\}$ be the message contents in CMSG G for (p_i, p_j) . Nondeterministic PDA $K_{i,j} = (Q, q_0, \Gamma, \Sigma, \Delta)$ where:

• Control states $Q = \{q_0, q_{a_1}, \dots, q_{a_k}, q_{err}, q_F\}$

• Stack alphabet $\Gamma = \{1, \#\}$ 1 counts nr. of unmatched $!(p_i, p_j, a_m)$, and # is bottom of stack

3

・ 同 ト ・ ヨ ト ・ ヨ ト

Let $\{a_1, \ldots, a_k\}$ be the message contents in CMSG G for (p_i, p_j) . Nondeterministic PDA $K_{i,j} = (Q, q_0, \Gamma, \Sigma, \Delta)$ where:

- Control states $Q = \{q_0, q_{a_1}, \dots, q_{a_k}, q_{err}, q_F\}$
- Stack alphabet $\Gamma = \{1, \#\}$ 1 counts nr. of unmatched $!(p_i, p_j, a_m)$, and # is bottom of stack

- possible symbols in the linearisations

• Input alphabet $\Sigma = \begin{cases} \text{unmatched action } !(p_i, p_j, a_m) \\ \text{unmatched action } ?(p_j, p_i, a_m) \\ \text{matched actions } !?(p_i, p_j, a_m) \end{cases} \checkmark$

Let $\{a_1, \ldots, a_k\}$ be the message contents in CMSG G for (p_i, p_j) . Nondeterministic PDA $K_{i,j} = (Q, q_0, \Gamma, \Sigma, \Delta)$ where:

• Control states $Q = \{q_0, q_{a_1}, \dots, q_{a_k}, q_{err}, q_F\}$

• Stack alphabet
$$\Gamma = \{1, \#\}$$

1 counts nr. of unmatched $!(p_i, p_j, a_m)$, and $\#$ is bottom of stack

- Input alphabet $\Sigma = \begin{cases} \text{unmatched action } !(p_i, p_j, a_m) \\ \text{unmatched action } ?(p_j, p_i, a_m) \\ \text{matched action } !?(p_i, p_j, a_m) \end{cases}$
- Transition function Δ is described on next slide

< 日 > < 同 > < 回 > < 回 > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ >

Safeness of CMSGs (2)

• Initial configuration is $(q_0, \#, w)$

• w is linearization of actions at p_i and p_j on an accepting path of G

э

- 4 戸 ト - 4 三 ト

Safeness of CMSGs (2)

Initial configuration is (q₀, #, w)
w is linearization of actions at p_i and p_j on an accepting path of G
On reading !(p_i, p_j, a_m) in q₀, push 1 on stack
nondeterministically move to state q_{am} or stay in q₀
"seen an unmatched send from p; to p; with write q_m"

÷

- Initial configuration is $(q_0, \#, w)$
 - w is linearization of actions at p_i and p_j on an accepting path of G
- On reading $!(p_i, p_j, a_m)$ in q_0 , push 1 on stack
 - nondeterministically move to state q_{a_m} or stay in q_0 p = rdiy unmetched send $p; \rightarrow p'_i$
- On reading $?(p_j, p_i, a_m)$ in q_0 , proceed as follows:
 - if 1 is on stack, pop it
 - otherwise, i.e., if stack is empty, accept (i.e., move to q_F)

not left-closed

÷

▲御▶ ▲注▶ ▲注▶

- Initial configuration is $(q_0, \#, w)$
 - w is linearization of actions at p_i and p_j on an accepting path of G
- On reading $!(p_i, p_j, a_m)$ in q_0 , push 1 on stack
 - nondeterministically move to state q_{a_m} or stay in q_0
- On reading $?(p_j, p_i, a_m)$ in q_0 , proceed as follows:
 - if 1 is on stack, pop it
 - otherwise, i.e., if stack is empty, accept (i.e., move to q_F)
- On reading matched send $\underline{!?}(p_i, p_j, a_k)$ in q_0
 - stack empty (i.e., equal to #)? ignore input; otherwise, accept

・ロト ・ 雪 ト ・ ヨ ト ・ ヨ ト

- Initial configuration is $(q_0, \#, w)$
 - w is linearization of actions at p_i and p_j on an accepting path of G
- On reading $!(p_i, p_j, a_m)$ in q_0 , push 1 on stack
 - nondeterministically move to state q_{a_m} or stay in q_0
- On reading $?(p_j, p_i, a_m)$ in q_0 , proceed as follows:
 - if 1 is on stack, pop it
 - otherwise, i.e., if stack is empty, accept (i.e., move to q_F)
- On reading matched send $!?(p_i, p_j, a_k)$ in q_0
 - stack empty (i.e., equal to #)? ignore input; otherwise, accept
- Ignore the following inputs in state q_0 :
 - matched send events $!?(p_j, p_i, a_k)$, and
 - unmatched sends or receipts not related to p_i and p_j

- Initial configuration is $(q_0, \#, w)$
 - w is linearization of actions at p_i and p_j on an accepting path of G
- On reading !(p_i, p_j, a_m) in q₀, push 1 on stack
 nondeterministically move to state q_{a_m} or stay in q₀
- On reading (p_i, p_i, a_m) in q_0 , proceed as follows:
 - if 1 is on stack, pop it
 - otherwise, i.e., if stack is empty, accept (i.e., move to q_F)

$\mathbf{X} \bullet$ On reading matched send $!?(\mathbf{p}_i, \mathbf{p}_i, a_k)$ in q_0 • stack empty (i.e., equal to #)? ignore input; otherwise, accept

- Ignore the following inputs in state q_0 :
 - matched send events $!?(p_i, p_i, a_k)$, and
 - unmatched sends or receipts not related to p_i and p_j static sends,
- Remaining input w empty? Accept, if stack-non-empty; else rejecta

The behaviour in state q_{a_m} for $0 < m \leq k$:

• Ignore all actions except $?(p_j, p_i, a_\ell)$ for all $0 < \ell \leq k$

The behaviour in state q_{a_m} for $0 < m \leq k$:

- Ignore all actions except $?(p_j, p_i, a_\ell)$ for all $0 < \ell \leq k$
- On reading ?(p_j, p_i, a_ℓ) (for some 0 < ℓ ≤ k) in state q_{a_m} do:
 if 1 is on top of stack, pop it

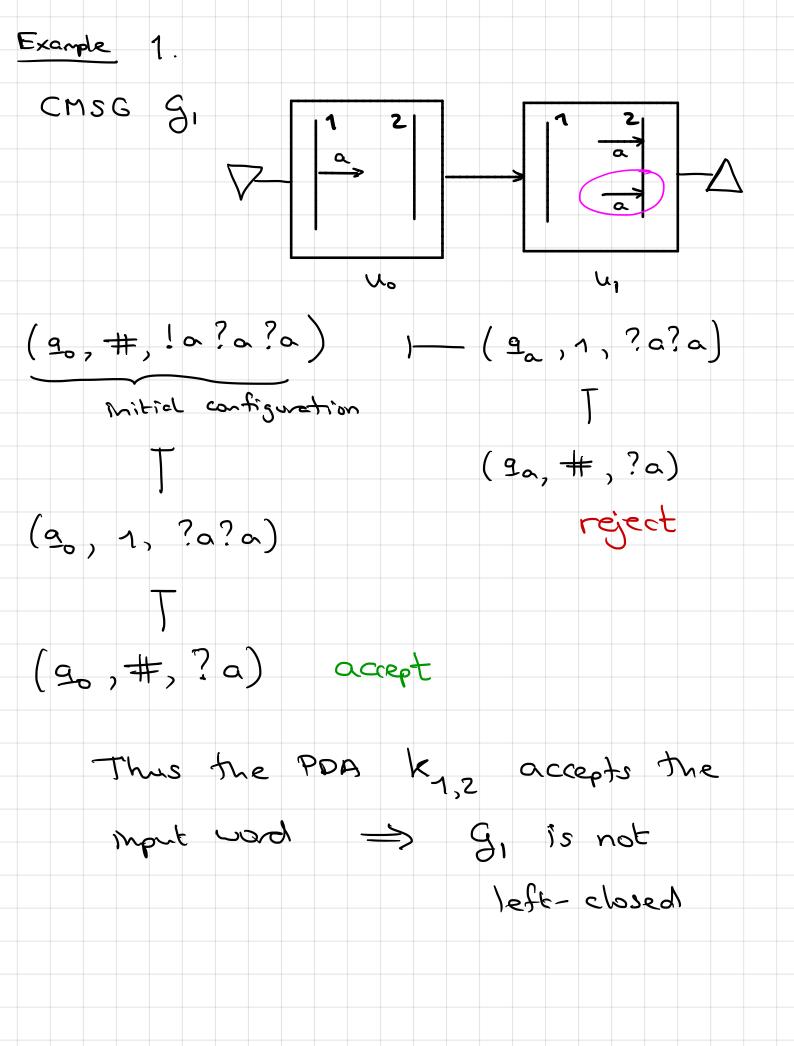
イロト 不得下 イヨト イヨト 三日

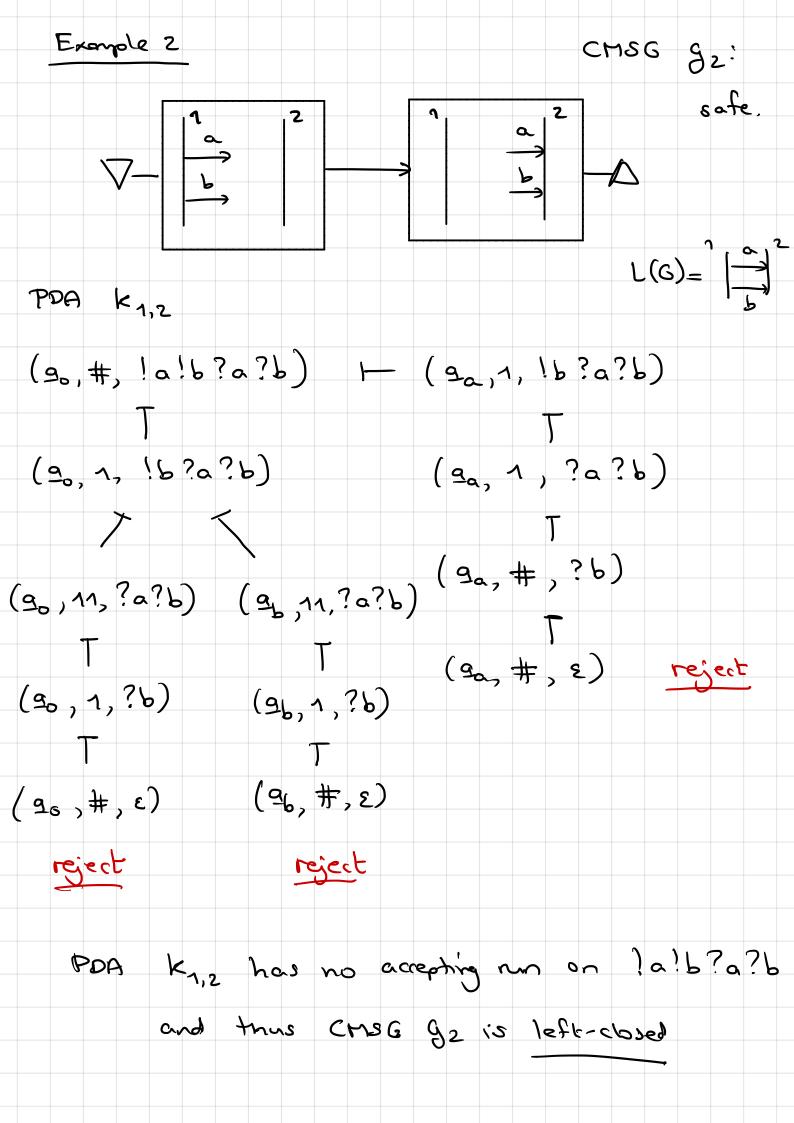
The behaviour in state q_{a_m} for $0 < m \leq k$:

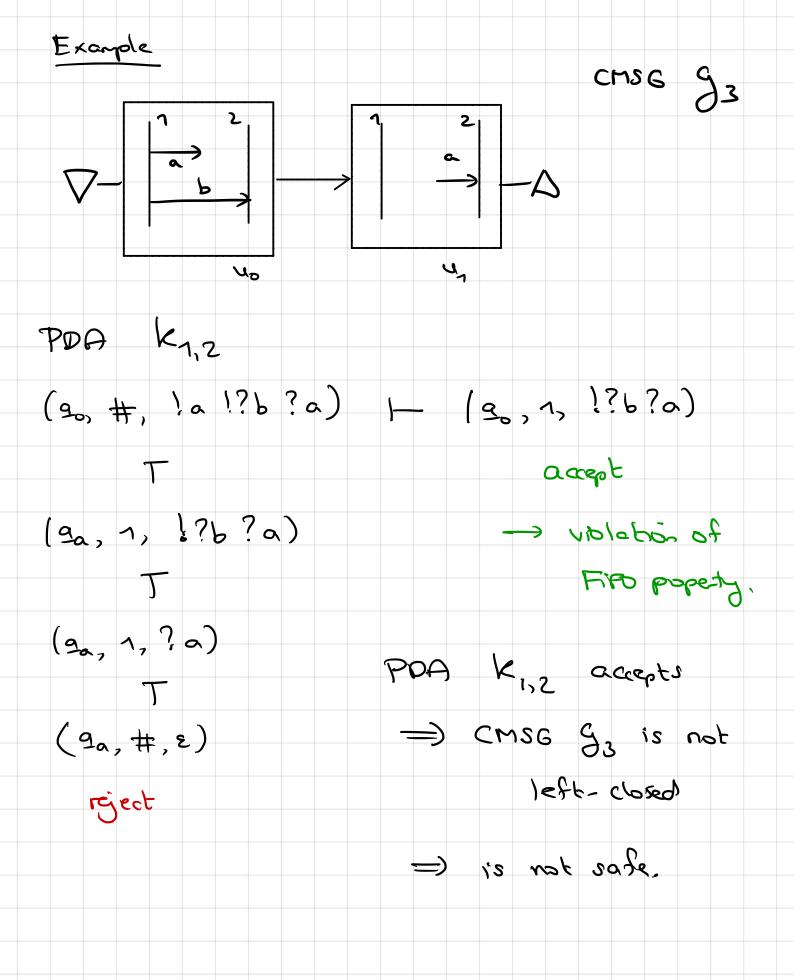
- Ignore all actions except $?(p_j, p_i, a_\ell)$ for all $0 < \ell \leq k$
- On reading ?(p_j, p_i, a_ℓ) (for some 0 < ℓ ≤ k) in state q_{a_m} do:
 if 1 is on top of stack, pop it
- \rightarrow If stack is empty:
 - if last receive differs from a_m , accept
 - otherwise reject, while ignoring the rest (if any) of the input

not left-closed

- 4 同 ト - 4 目 ト - 4 目 ト







It follows: PDA $K_{i,j}$ accepts iff CMSG G is not **lc** wrt. (p_i, p_j)

 $\implies \text{CMSG } G \text{ is not } \mathbf{L} \mathsf{c} \text{ wrt. } (p_i, p_j) \text{ iff configuration } \underbrace{(q_F, \cdot, \cdot)}_{\mathbf{Q} \in \mathbf{C}} \text{ is not } \mathbf{L} \mathsf{c}$

3

▲圖 ▶ ▲ 国 ▶ ▲ 国 ▶

It follows: PDA $K_{i,j}$ accepts iff CMSG G is not Lc wrt. (p_i, p_j)

- \implies CMSG G is not **lc** wrt. (p_i, p_j) iff configuration (q_F, \cdot, \cdot) is reachable.

э

It follows: PDA $K_{i,j}$ accepts iff CMSG G is not c wrt. (p_i, p_j)

- \implies CMSG G is not (p_i, p_j) iff configuration (q_F, \cdot, \cdot) is reachable.
- ⇒ reachability of a configuration in a PDA is in PTIME, hence checking safeness wrt. (p_i, p_j) is in PTIME.

Time complexity

The worst-case time complexity of checking whether CMSG G is safe is in $\mathcal{O}(k^2 \cdot N^2 \cdot L \cdot |E|^2)$ where $k = |\mathcal{P}|, N = |V|$, and $L = |\mathcal{C}|$.

・ 何 ト ・ ヨ ト ・ ヨ ト … ヨ

It follows: PDA $K_{i,j}$ accepts iff CMSG G is not safe wrt. (p_i, p_j)

- \implies CMSG G is not safe wrt. (p_i, p_j) iff configuration (q_F, \cdot, \cdot) is reachable.
- ⇒ reachability of a configuration in a PDA is in PTIME, hence checking safeness wrt. (p_i, p_j) is in PTIME.

Time complexity

The worst-case time complexity of checking whether CMSG G is safe is in $\mathcal{O}(k^2 \cdot N^2 \cdot L \cdot |E|^2)$ where $k = |\mathcal{P}|, N = |V|$, and $L = |\mathcal{C}|$.

Proof.

Checking reachability in PDA $K_{i,j}$ is in $\mathcal{O}(L \cdot |E|^2)$.

It follows: PDA $K_{i,j}$ accepts iff CMSG G is not safe wrt. (p_i, p_j)

- \implies CMSG G is not safe wrt. (p_i, p_j) iff configuration (q_F, \cdot, \cdot) is reachable.
- ⇒ reachability of a configuration in a PDA is in PTIME, hence checking safeness wrt. (p_i, p_j) is in PTIME.

Time complexity

The worst-case time complexity of checking whether CMSG G is safe is in $\mathcal{O}(k^2 \cdot N^2 \cdot L \cdot |E|^2)$ where $k = |\mathcal{P}|, N = |V|$, and $L = |\mathcal{C}|$.

Proof.

Checking reachability in PDA $K_{i,j}$ is in $\mathcal{O}(L \cdot |E|^2)$. The number of PDAs is k^2 , as we consider ordered pairs in \mathcal{P} .

It follows: PDA $K_{i,j}$ accepts iff CMSG G is not safe wrt. (p_i, p_j)

- \implies CMSG G is not safe wrt. (p_i, p_j) iff configuration (q_F, \cdot, \cdot) is reachable.
- ⇒ reachability of a configuration in a PDA is in PTIME, hence checking safeness wrt. (p_i, p_j) is in PTIME.

Time complexity

The worst-case time complexity of checking whether CMSG G is safe is in $\mathcal{O}(k^2 \cdot N^2 \cdot L \cdot |E|^2)$ where $k = |\mathcal{P}|, N = |V|$, and $L = |\mathcal{C}|$.

Proof.

Checking reachability in PDA $K_{i,j}$ is in $\mathcal{O}(L \cdot |E|^2)$. The number of PDAs is k^2 , as we consider ordered pairs in \mathcal{P} . The number of paths in the CMSG G for each pair that need to be checked is in $\mathcal{O}(N^2)$, as a single traversal for each loop in G suffices.