

Plan

- 1 Le modèle standard
 - Approche événementielle
 - Causalité
 - Abstraction d'un calcul
- 2 Clichés (snapshots)
 - Prise de cliché
 - Utilisation des clichés
- 3 Description des algorithmes
 - Description du comportement des processus
 - Exemple : l'élection

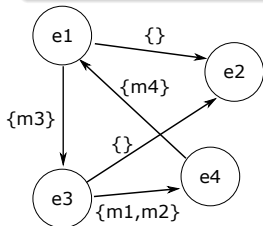


Prise de cliché (snapshot)



Définition

Objectif : Capter un état global (passé) des processus et du réseau



- Prise **instantanée** impossible
- Un site collecteur accumule
- Prise **cohérente** de clichés locaux
- Identification des messages en transit

Cliché global

Clichés locaux + Messages en transit
 $\{e_1, e_2, e_3, e_4\} + \{m_1, m_2, m_3, m_4\}$

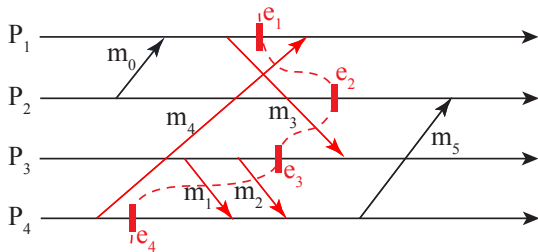
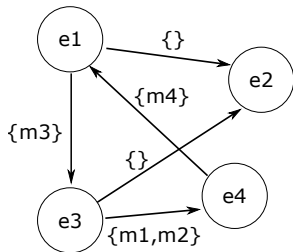
[Précis 5.1 pp.79–81]



Prise de cliché (snapshot)



Schéma temporel de la prise de cliché



Algorithme de Chandy-Lamport (1985)



- Un système existant échange des messages ;
- On **superpose** des échanges de messages dédiés pour déclencher des actions locales de sauvegarde de l'état d'un site (= un cliché local + des messages reçus) ;
- Ces états sauvegardés sont **collectés** pour construire un cliché global.

1. *Distributed Snapshots : Determining Global States of Distributed Systems*,
K. Mani Chandy and Leslie Lamport. ACM Transactions on Computer Systems,
Feb. 1985



Algorithme de Chandy-Lamport

Idée

- Construire une coupure cohérente au moyen de **marqueurs** visitant les sites.
- Les messages émis par S_j **avant** le passage du marqueur sur S_j , et reçus par S_i **après** le passage du marqueur sur S_i , sont les messages en transit de S_j vers S_i .
- Les messages reçus avant le passage du marqueur sont intégrés à l'état local du site et ne sont plus en transit.
- Les messages émis après le passage du marqueur ne sont pas dans le cliché.



Algorithme de Chandy-Lamport

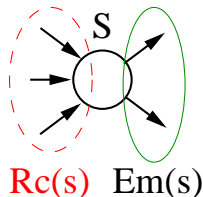


Hypothèses

- Canaux unidirectionnels et **fifo** :
 $\forall s, Rc(s)$: canaux en réception
 $Em(s)$: canaux en émission
- Réseau fortement connexe ($\forall s, s' : \exists s \rightarrow^* s'$)

Principe de l'algorithme

- Utilisation de messages **marqueurs**
- Répartition de l'évaluation : chaque site s évalue :
 - **son** cliché local ;
 - les messages considérés en transit sur ses canaux en réception $Rc(s)$



Algorithme de Chandy-Lamport



Comportement d'un site s

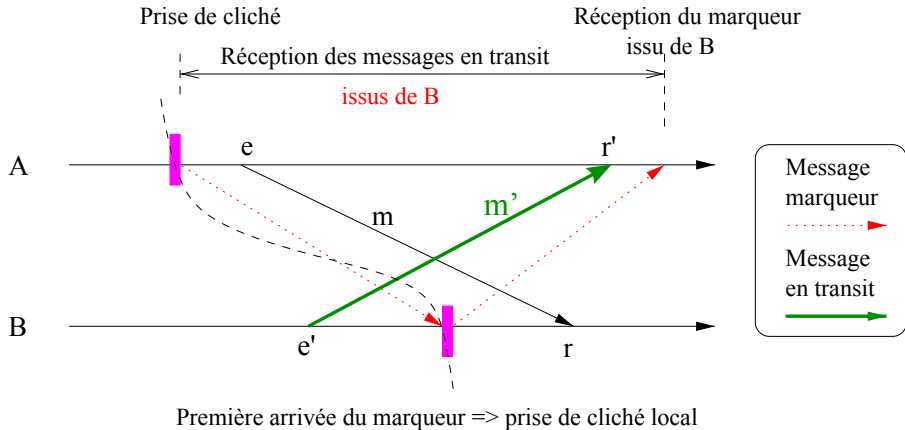
Sur réception d'un **premier marqueur** :

- 1 Prendre **son** cliché local L_s et émettre un **marqueur** sur chaque canal d'émission $c \in Em(s)$
- 2 Enregistrer dans une liste $enTransit[c]$ les messages reçus sur chaque canal de réception $c \in Rc(s)$ jusqu'à la réception d'un **marqueur** sur ce canal
- 3 Lorsqu'un marqueur a été reçu sur **tous** ses canaux de réception, communiquer au collecteur cet état partiel :
 $\langle L_s, \{enTransit[c] \mid c \in Rc(s)\} \rangle$

Déclenchement de la prise de cliché : envoi d'un message marqueur à un site quelconque.



Prise de clichés locaux et marqueurs



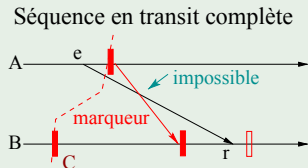
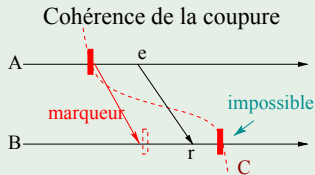
Vérifier la correction...



Propriétés

- Sûreté
 - Coupure cohérente
 - Collecte complète des messages en transit
- Vivacité
 - Tout site finit par prendre un cliché local
 - Un marqueur finit par arriver sur chaque canal de réception

Exemple



État enregistré = état *possible*

État enregistré

- Σ_{enreg} = cliché enregistré
- Σ_{init} = coupure cohérente contenant l'événement déclencheur du cliché
- Σ_{final} = coupure cohérente dans lequel le protocole de prise de cliché est terminé

Alors $\Sigma_{init} \prec \Sigma_{enreg} \prec \Sigma_{final}$

(il existe un chemin de Σ_{init} à Σ_{final} passant par Σ_{enreg} dans le treillis des coupures cohérentes)

Exemple : sur le treillis page 18, si $\Sigma_{init} = \Sigma^{11}$ et $\Sigma_{final} = \Sigma^{32}$, Σ_{enreg} peut être Σ^{11} , Σ^{21} , Σ^{12} , Σ^{31} , Σ^{22} ou Σ^{32} , et a pu ne pas être traversé dans la réalité.

Utilisation du cliché : propriété stable



Prédicat stable

Un prédicat P sur un état global E d'un système est stable ssi
 $\forall E' : E \prec E' \wedge P(E) \Rightarrow P(E')$

(exemples : le calcul est terminé, il y a eu 10 messages reçus...)

Vérification de P

Si P est un prédicat stable alors :

- $P(\Sigma_{enreg}) \Rightarrow P(\Sigma_{final})$ (et tout état ultérieur)
- $\neg P(\Sigma_{enreg}) \Rightarrow \neg P(\Sigma_{init})$ (et tout état antérieur)



Utilisation du cliché : propriété possible/certaine



Prédicat possible/certain

Pour un prédicat P :

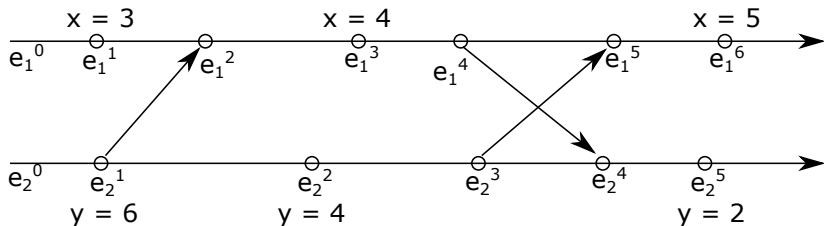
- $Pos(P)$ (possibly P) : il existe une observation cohérente (= un chemin dans le treillis) qui passe par un état où P est vrai.
- $Def(P)$ (definitely P) : toutes les observations cohérentes (= tous les chemins) passent par un état où P est vrai.

Vérification

- $P(\Sigma_{enreg}) \Rightarrow Pos(P)$ mais pas l'inverse...
- $\neg Pos(P) \Rightarrow Def(\neg P)$ mais pas l'inverse...



Exemple de vérification de propriétés



- $x - y \geq 0$?
(en supposant x croissant, y décroissant \Rightarrow propriété stable)
- $Pos(x = y - 2)$?
- $Def(x = y)$?

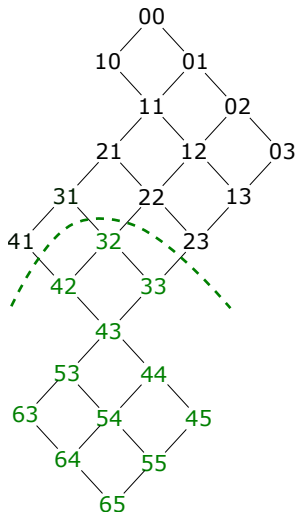
(d'après Lorenzo Alvisi)



Exemple de vérification



Propriété stable



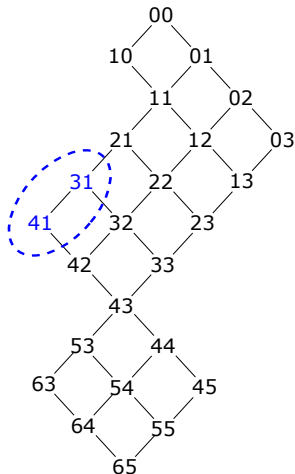
- $x - y \geq 0$?
(sous l'hypothèse $x \uparrow, y \downarrow$)
- N'importe quel cliché Σ_{enreg} obtenu après Σ^{32} permet de le vérifier



Exemple de vérification



Possibilité



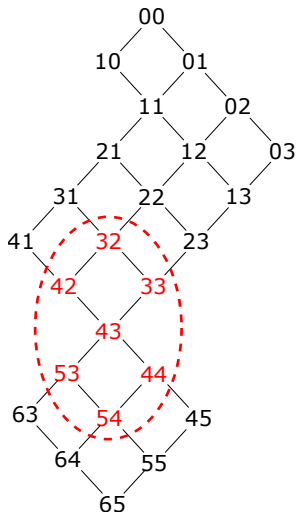
- $Pos(x = y - 2)$?
- $x = y - 2$ est vrai dans les états cohérents Σ^{31} et Σ^{41}
- Détecté uniquement si $\Sigma_{enreg} \in \{\Sigma^{31}, \Sigma^{41}\}$
- Σ_{enreg} pas nécessairement survenu dans la réalité



Exemple de vérification



Certitude



- $Def(x = y)$?
- Vrai
- $Pos(x = y)$ pas détecté si on capture un état antérieur à Σ^{32} ou postérieur à Σ^{54}
- La capture d'un état (p.e. Σ^{42}) ne permet pas de conclure



Utilisation du cliché : propriété possible/certaine



Principe de la vérification

- Un processus moniteur M collecte tous les états locaux
- M construit le treillis des coupures cohérentes (à partir d'un codage complet de la relation de causalité, cf chapitre suivant)
- Pour évaluer $Pos(P)$: parcourir le treillis depuis l'état initial, niveau par niveau, et s'arrêter au premier état où P est vrai. Aucun état $\Rightarrow \neg Pos(P)$.
- Pour évaluer $Def(P)$: parcourir le treillis depuis l'état initial, niveau par niveau, en ne développant que les états vérifiant $\neg P$. Si plus d'état, alors $Def(P)$; si état final atteint (et $\neg P$ dans cet état) alors $\neg Def(P)$.
- Explosion combinatoire : pour N sites ayant chacun au plus m états, possiblement m^N coupures cohérentes.

