

SPÉCIFICATION ET VÉRIFICATION DE PROGRAMMES

17 mars 2025

Allan Blanchard

Partie I

Introduction générale

Pourquoi ce cours ?

Que va-t-on y faire ?

> Avez-vous déjà développé du logiciel ?

- > Avez-vous déjà développé du logiciel ?
- > Avez-vous déjà développé du logiciel contenant des bugs ?

- > Avez-vous déjà développé du logiciel ?
- > Avez-vous déjà développé du logiciel contenant des bugs ?
- > De quelle nature étaient ces bugs ?

- > Avez-vous déjà développé du logiciel ?
- > Avez-vous déjà développé du logiciel contenant des bugs ?
- > De quelle nature étaient ces bugs ?
- > Avez-vous déjà développé du logiciel sans bugs ?

- > Avez-vous déjà développé du logiciel ?
- > Avez-vous déjà développé du logiciel contenant des bugs ?
- > De quelle nature étaient ces bugs ?
- > Avez-vous déjà développé du logiciel sans bugs ?
- > Si oui, comment le savez-vous ?

- > Avez-vous déjà développé du logiciel ?
- > Avez-vous déjà développé du logiciel contenant des bugs ?
- > De quelle nature étaient ces bugs ?
- > Avez-vous déjà développé du logiciel sans bugs ?
- > Si oui, comment le savez-vous ?

Tout programme non trivial contient (au moins) un bug (Loi de Murphy)

Anomalie logicielle

Comportement observé du logiciel qui diffère du comportement attendu

Propriétés fonctionnelles

- > Crash
- > Exception non rattrapée
- > Résultat erroné

Propriétés non-fonctionnelles

- > Mémoire consommée trop importante,
- > Temps d'exécution trop important,
- > Observation de données confidentielles
- > Non-terminaison ? Non !

Le caractère gênant des anomalies dépend

- > de leur **fréquence**,
- > de leur **dangerosité**,
- > de la facilité à les **contourner**,
- > de leur impact sur **l'utilisabilité du logiciel**,
- > de leur impact sur **l'environnement extérieur**,
- > ...

- > en 1996, destruction de la fusée Ariane 5 durant son vol inaugural
- > Cause :
 - > réutilisation d'un composant d'Ariane 4... sans respecter sa spécification
 - > débordement lors de la conversion d'un float de 64 bits vers un int 16 bits
- > Conséquence :
 - > mise hors service du composant
 - > mauvaise interprétation du pilote automatique
 - > déviation rapide de trajectoire
 - > arrachage des boosters
 - > autodestruction préventive de la fusée
- > Coût : \approx 500 millions de dollars
- > l'un des bugs les plus chers de l'histoire

- > de 1985 à 1987, l'appareil de radiothérapie Therac-25 est responsable du décès de plusieurs personnes
- > Causes :
 - > la cause première était la mauvaise architecture du système et de mauvaises pratiques de développements plutôt que les erreurs de programmation trouvées,
 - > en particulier, il n'était pas possible de tester automatiquement et proprement le système
- > Conséquence :
 - > l'utilisateur pouvait programmer un surdosage du patient sans s'en rendre compte
 - > aucun dispositif de contrôle
 - > les patients se retrouvaient alors sévèrement irradiés
- > l'un des bugs les plus graves de l'histoire

http://en.wikipedia.org/wiki/List_of_software_bugs

Catégories

- > Exploration spatiale
- > Médical
- > *tracking years*
- > Réseau électrique
- > Finance
- > Télécommunication
- > Militaire
- > Média
- > Jeux
- > Cryptographie

Nos systèmes tendent à devenir de plus en plus gros

Exemple : la taille de Linux

- > 1991 : 10 kLOC
- > 2001 : 2.4 MLOC
- > 2003 : 5.2 MLOC
- > 2009 : 11 MLOC
- > 2015 : 20.2 MLOC
- > 2020 : 27.8 MLOC

Problème de taille

- > le nombre de bugs est fonction de la taille de système ...
- > la complexité du système global aussi ...

- > coût estimé des bugs :
 - > USA 2020 : 607 milliards de dollars (CPSQ report 2020)
- > Coût pour l'utilisateur :
 - > accidents graves, voir mortels,
 - > dégâts environnementaux,
 - > destruction de matériels.
- > Coût pour le fournisseur :
 - > coût du correctif,
 - > coût du remboursement.

Confiance

Avoir confiance dans les logiciels est donc un enjeu sociétal et économique majeur.

- > Idéalement, en garantissant qu'il n'a "aucune erreur"
- > Qu'est-ce-que cela veut dire ?
 - > Informellement : il fait bien ce qu'il est censé faire
 - > Formellement : à démêler dans ce cours
- > Cet idéal est-il possible ?

- > Idéalement, en garantissant qu'il n'a "aucune erreur"
- > Qu'est-ce-que cela veut dire ?
 - > Informellement : il fait bien ce qu'il est censé faire
 - > Formellement : à démêler dans ce cours
- > Cet idéal est-il possible ? **Non!** (plus de détails par la suite)
- > En pratique :
 - > en se préservant des erreurs "autant que possible"
 - > **en atteignant le rapport qualité/prix jugé optimal**
- > Comment définir l'optimal ?
 - > Logiciel "normal" : défini par le client
 - > Logiciel "critique" : imposé par des lois/normes/assurances

Définition et utilité

- > La norme DO-178 traite de la **sûreté des logiciels avioniques**.
- > Guide pour déterminer si le logiciel sera fiable dans un environnement avionique.
- > La suivre est requis pour **certifier le logiciel** pour l'utiliser en production.
- > La version DO-178C de cette norme date de 2011.
- > La première version de la DO-178 date de 1982.

(Informations de cette section extraites de Wikipédia, cette norme coûte un bras)

- A) Catastrophique – morts multiples, perte de l'avion.
- B) Dangereux – réduction importante de la sécurité, équipage dans l'incapacité d'accomplir ses tâches, blessures sérieuses ou mort de quelques occupants.
- C) Majeur – réduction significative de la sécurité, équipage en difficulté, inconfort pour les occupants et risque de blessures.
- D) Mineur – le problème est visible, mais n'implique pas de risque de sécurité.
- E) Pas d'effet

Exigences

DAL	Objectives	Failure Rate
A	66	$10^{-9}/h$
B	65	$10^{-7}/h$
C	57	$10^{-5}/h$
D	28	$10^{-3}/h$
E	0	n/a

- Constat : 1 accident grave pour 10^6 heures de vol
- Constat : 10% liés à des défaillances des systèmes avioniques
- Conclusion : seuil acceptable d'accidents catastrophiques imputables aux systèmes = 10^{-7} par heure de vol
- arbitrairement, on considère que $\approx 10^2$ de conditions de pannes peuvent être catastrophiques et qu'elles sont équiprobables
- donc, probabilité acceptable d'une condition de panne catastrophique par heure de vol = 10^{-9}
- les contraintes imposées sur le cycle de développement sont fonction de cette probabilité

Aucune contrainte de développement

- logiciel documenté ; la liste des documents fixée par la norme
- des plans doivent être établis pour fixer les méthodes de développement, de vérification, de gestion de configuration, d'assurance qualité
- traçabilité spécification du système → spécifications de haut niveau → vérifications
- tout ce qui est spécifié doit être vérifié : couverture fonctionnelle et documents doivent aussi être vérifiés
- le logiciel doit être géré en configuration, toutes les évolutions du code source doivent être justifiées
- Un service Qualité indépendant doit assurer le respect de la norme

En plus des contraintes du niveau D :

- > des règles de développement doivent être fixées (spécification, conception et codage)
- > les contraintes de traçabilité et de vérification s'appliquent à la conception et au codage
- > les exigences de bas niveau doivent être vérifiées
- > la couverture structurelle, ou couverture de code, doit être vérifiée et analysée

En plus des contraintes du niveau C :

- > La couverture de code au niveau “décision” est requise.
- > Les activités de développement et de vérification doivent être confiées à des équipes indépendantes.

En plus des contraintes du niveau B :

- > La couverture de code au niveau “condition”/“décision” est requise.

Vérifier et Valider (V & V)

Quelle est la différence entre les deux ?

Vérifier et Valider (V & V)

Quelle est la différence entre les deux ?

Vérification : “Are we building the product right ?”

Point de vue **fournisseur** : est-ce-que le logiciel fonctionne correctement ?

Validation : “Are we building the right product ?”

Point de vue **utilisateur** : est-ce-que le logiciel fait bien ce que l'on veut ?

- > Besoin de définir ce que fait un programme, *i.e.* sa **sémantique**
- > besoin de définir ce qu'est censé faire un programme, *i.e.* sa **spécification**
- > **comment garantir qu'un programme** (caractérisé par sa sémantique) **satisfait sa spécification ?**

Analyse de programmes

Déterminer certaines propriétés d'un programme, en particulier des propriétés sémantiques sur son comportement à l'exécution.

Dynamique

L'analyseur peut exécuter le programme

Statique

L'analyseur ne peut pas exécuter le programme

peut exécuter le programme

- > **test** (chapitre II)
- > **monitoring** (chapitre IV)
- > **avantage** : prend en compte le contexte d'exécution
- > **inconvenient** : subit/influe sur le contexte d'exécution

observe le code du programme, mais ne l'exécute pas

Quel code ?

- > Source (ici, C)
- > Assembleur
- > Binaire

Quelles analyses ?

- > **interprétation abstraite** (chapitre V)
- > **méthode déductive** (chapitre VI)
- > **inconvenient** ? : ne prend pas en compte le contexte d'exécution, voire de compilation
- > **avantage** : ne nécessite pas/n'influe pas sur le contexte d'exécution

Usage majeur historique des analyses statiques

Les analyses statiques sont très utilisées dans les compilateurs.

Exemples

- > Propagation de constantes,
- > élimination de sous-expressions communes,
- > élimination de code mort,
- > variables “vivantes” et allocation de registres...

- > vérifier et valider un programme est **long et compliqué**
- > vu la taille des systèmes actuels, **impossible à faire manuellement**

Besoin d'automatisation et d'outillage

Quel outil parfait pour garantir qu'un programme P satisfait sa spécification S ?

Quel outil parfait pour garantir qu'un programme P satisfait sa spécification S ?

- 1 } Donner P et S en entrée à un outil O
- 2 } appuyer sur "entrée" (analyse automatique)
- 3 } O regarde P et S , idéalement sans exécuter P
- 4 } O donne instantanément la réponse
 - > oui : P satisfait S
 - > non : il y a un problème à tel endroit P par rapport à tel point de S
- 5 } être sûr que la réponse donnée par O est correcte

Quel outil parfait pour garantir qu'un programme P satisfait sa spécification S ?

- 1 } Donner P et S en entrée à un outil O
- 2 } appuyer sur "entrée" (analyse automatique)
- 3 } O regarde P et S , idéalement sans exécuter P
- 4 } O donne instantanément la réponse
 - > oui : P satisfait S
 - > non : il y a un problème à tel endroit P par rapport à tel point de S
- 5 } être sûr que la réponse donnée par O est correcte

Problème : L'univers n'est pas indifférent à l'intelligence, il lui est activement hostile

Une analyse statique automatique exacte est impossible.

Théorème de Rice (1953)

Toute propriété

- > **extensionnelle** (qui dépend uniquement de la sémantique du programme)
- > **non triviale** (ni toujours vraie, ni toujours fausse)

est

- > **indécidable**

Exemple : problème de l'arrêt

Il est impossible de déterminer statiquement si un programme quelconque termine (en temps fini) quelles que soient ses données d'entrée.

- > **Test** : exécute le programme pour trouver des bugs. (Chapitre II)
- > **Model-checking** : travailler non pas sur le programme, mais sur un modèle du programme (Non traité dans ce cours)
- > **Interprétation abstraite** : effectuer des approximations de la sémantique concrète. (Chapitre V)
- > **Méthodes déductives** : raisonner de manière précise, par déduction, sur le programme et extraire les problèmes “durs” (Chapitre VI)

Chaque technique a ses avantages...

Et au moins un gros inconvénient

> Que cherche-t-on à faire ?

- > Que cherche-t-on à faire ?
 - > **Vérifier** : est-ce-que le logiciel fonctionne correctement ?
 - > **Valider** : est-ce-que le logiciel fait bien ce qu'il est censé faire ?
- > Comment cherche-t-on à le faire ?

- > Que cherche-t-on à faire ?
 - > **Vérifier** : est-ce-que le logiciel fonctionne correctement ?
 - > **Valider** : est-ce-que le logiciel fait bien ce qu'il est censé faire ?
- > Comment cherche-t-on à le faire ?
 - > En **analysant** le programmes
 - > En **outillant** ce processus d'analyse pour qu'il se rapproche de l'idéal :
 - > correct
 - > exact
 - > automatique (et instantané)
 - > statique

INTRODUCTION

TEST DE PROGRAMMES

SPÉCIFICATION

VÉRIFICATION À L'EXÉCUTION

ANALYSE STATIQUE

VÉRIFICATION DÉDUCTIVE

COMBINAISONS DE TECHNIQUES

CONCLUSION

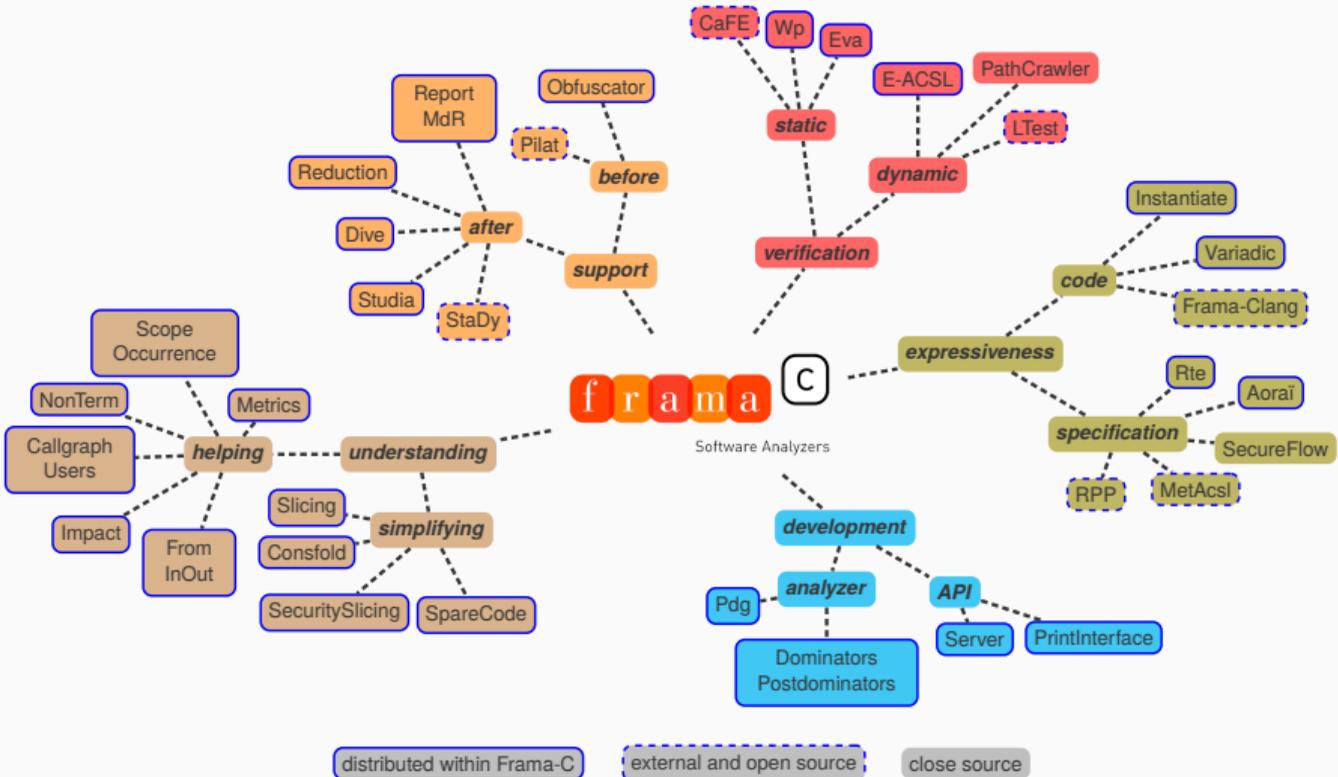
- > C
 - > le langage est tout tordu et plein de pièges
- > uthash
 - > un outil de test header-only
- > gcovr
 - > un outil d'analyse de couverture pour GCC
- > Frama-C
 - > une suite d'outils pour l'analyse de code C



Plateforme dédiée à l'analyse de code C

- > développée depuis 2004,
- > utilisée industriellement (THALES, Airbus, ANSSI, ...),
- > open-source,
- > extensible, basée sur une architecture à plug-ins,
- > principalement utilisée pour de l'analyse formelle.

FRAMEWORK FOR MODULAR ANALYSIS OF C CODE



Partie II

Test de programmes

Comment tester ?

Comment évaluer des tests ?

> Avez-vous déjà **testé un logiciel** ?

- > Avez-vous déjà **testé un logiciel** ?
- > Était-ce **le vôtre** ou celui d'un **tiers** ?

- > Avez-vous déjà **testé un logiciel** ?
- > Était-ce **le vôtre** ou celui d'un **tiers** ?
- > Quel(s) étai(en)t votre/vos **but(s)** ?

- > Avez-vous déjà **testé un logiciel** ?
- > Était-ce **le vôtre** ou celui d'un **tiers** ?
- > Quel(s) étai(en)t votre/vos **but(s)** ?
- > Comment vous y preniez-vous ?
 - > Au hasard ?
 - > Avec une **méthodologie** particulière ? Laquelle ?

- > Avez-vous déjà **testé un logiciel** ?
- > Était-ce **le vôtre** ou celui d'un **tiers** ?
- > Quel(s) étai(en)t votre/vos **but(s)** ?
- > Comment vous y preniez-vous ?
 - > Au hasard ?
 - > Avec une **méthodologie** particulière ? Laquelle ?
- > **Combien de temps** y avez-vous consacré (comparativement au reste du développement) ?

Selon G. Myers

Tester, c'est exécuter le programme dans l'intention d'y trouver des anomalies ou des défauts.
(G. Myers (The Art of Software Testing, 1979))

Selon l'IEEE

Le test est l'exécution ou l'évaluation d'un système ou d'un composant, par des moyens automatiques ou manuels, pour vérifier qu'il répond à ses spécifications ou identifier les différences entre les résultats attendus et les résultats obtenus.

Selon G. Myers

Tester, c'est exécuter le programme dans l'intention d'y trouver des anomalies ou des défauts.
(G. Myers (The Art of Software Testing, 1979))

Selon l'IEEE

Le test est l'exécution ou l'évaluation d'un système ou d'un composant, par des moyens automatiques ou manuels, pour vérifier qu'il répond à ses spécifications ou identifier les différences entre les résultats attendus et les résultats obtenus.

Ce que le test n'est pas

Testing can only reveal the presence of errors but never their absence.
(E. W. Dijkstra (Notes on Structured Programming, 1972))

- > marché de plus de 40 milliards d'euros en 2020
- > estimation à plus de 60 milliards d'ici à 2027
- > > 50% du temps du cycle de développement d'un logiciel critique (parfois > 90%)
- > en moyenne 30% pour un logiciel "standard"

Le test coûte cher

Plus un bug est corrigé tard, plus le correctif est cher, empiriquement :

en phase de développement	coût 1
en phase d' intégration (bug de conception)	coût 10
en phase de recette (bug de spécification)	coût 100
en phase d' exploitation	coût > 1000

Nécessité de **détecter les problèmes au plus tôt** du cycle de développement

Un bug finit par coûter très cher

1) Construire la qualité du produit

- > lors de la phase de conception, du codage,
- > en partie par les développeurs,
- > objectif : trouver rapidement le plus de bugs possible.

2) Démontrer la qualité du produit à un tiers

- > une fois le produit terminé,
- > idéalement : par une équipe dédiée,
- > objectif : convaincre un organisme de certification, le client ...

Cas de test

Scénario à exécuter pour tout ou partie du système, composition :

- > caractéristiques de l'entrée correspondante,
- > estimation de l'**oracle**, le résultat attendu.

On dérive :

- > une **donnée de test** représentative de l'entrée attendue,
- > un **oracle concret** correspondant.

Finalement :

- > on **exécute** le test,
- > on **compare*** le résultat avec l'oracle concret (Verdict : succès/échec)

* Remarque : l'égalité stricte n'est pas forcément nécessaire

Définir l'oracle est un problème difficile

“Bons” cas

- > le programme ne plante pas
- > programmation par contrat
- > test dos à dos : tester avec un logiciel similaire

Suite et script

- > Suite de tests : ensemble de tests du système
- > Script de test : code qui lance les différents tests et collecte les résultats

Test de programme

Analyse dynamique !

Le but des tests est de trouver des bugs

Ne pas tester en supposant qu'aucun bug ne va être trouvé !

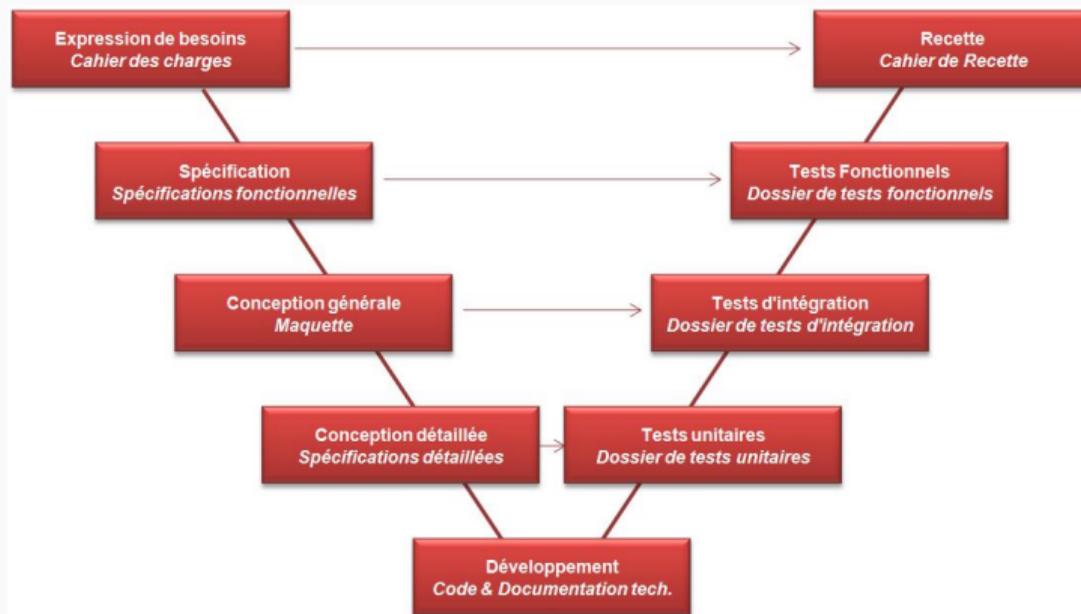
Quelques guidelines sur l'attitude à adopter

- > **Cherchez** les bugs
- > Définissez les sorties voulues **avant** de lancer le test
- > Observez méticuleusement les oracles
- > Idéalement : le développement et le test sont faits par deux personnes différentes

- > exercer un **maximum de comportements différents** du programme
- > **tests nominaux**
 - > cas de fonctionnements les plus fréquents
 - > forte probabilité de survenir
 - > peu de risque d'erreurs
- > **tests de robustesse**
 - > cas de fonctionnements les moins fréquents
 - > faible probabilité de survenir
 - > risque d'erreurs important

Très difficile à définir en pratique

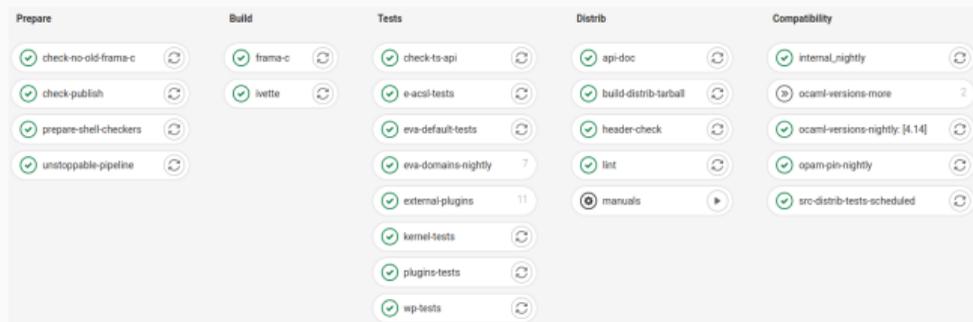
- > **test unitaire** : tester les différents composants en isolation,
- > **test d'intégration** : tester après composition des composants,
- > **test système/de conformité** : valider globalement le code
 - > inclut si nécessaire les tests de sécurité, de performance, ...
- > **test de validation/acceptation** : valider l'adéquation aux besoins du client
 - > similaire au test système, mais réalisé/vérifié par le **client**
- > **test de non-régression** : vérifier que les corrections/évolutions du code n'ont pas introduit de bugs



- > sans oublier la **maintenance** et les **tests de non-régression**
- > dans des processus plus mainstream, on répète des V courts

Tester tout le temps

- > chaque ajout de code est testé avant intégration
- > chaque intégration est re-validée a posteriori
- > des tests plus poussés peuvent être joués à intervalles réguliers



Bénéfices

- > pas de travail ajouté pour le développeur,
- > détection au plus tôt des problèmes

Méthodologie

- 1 > écrire un seul test qui décrit une partie du problème à résoudre,
- 2 > vérifier que le test échoue (le code qui permet sa réussite n'existe pas encore),
- 3 > écrire juste assez de code pour que le test réussisse,
- 4 > vérifier que le test passe, ainsi que les autres tests existants,
- 5 > remanier le code/le test sans altérer le comportement,
- 6 > revenir à 1 si la fonctionnalité n'est pas encore complète.

Bénéfices

- > Permet d'avoir rapidement de nombreux tests de bonne qualité
- > Construit la bonne qualité du logiciel pendant le développement

Tester chaque composant en isolation

- > généralement la granularité est au niveau de **la fonction**

Approches

- > top-down (depuis le point d'entrée) ou bottom-up (depuis les feuilles)
- > dans la réalité, souvent un mix des deux au cas par cas

Code incomplet

- > code appelé manquant : création d'une **doublure**
- > code appelant manquant : création d'un **lanceur**

> top-down

- ✗ Les *doublures* peuvent être complexes à écrire
- ✓ On teste beaucoup plus les composants de haut niveau

> bottom-up

- ✓ Les lanceurs sont généralement assez facile écrire
- ✗ On risque de moins tester les composants de haut niveau

- > Dummy : nécessaire au test (*e.g.* paramètre) mais pas utilisé
- > Stub : utilisée par le test, mais a une liste fixe de retours
- > Spy : utilisée par le test, observe cet usage (pour contrôle *a posteriori*)
- > Mock : utilisée par le test, observe cet usage (avec contrôle *immédiat*)
- > Fake : répond à la spécification du code remplacé, mais simplifié

Plus la doublure est avancée, plus elle est difficile à produire et fragilise le test

- > **Test exhaustif** : on teste toutes les entrées (impossible en pratique)
- > **Test probabiliste** : on teste à partir de données statistiques
- > **Test fonctionnel** : test boîte noire, à partir des spécifications, sans regard sur le code
- > **Test structurel** : test boîte blanche, en prenant en compte la structure du code

Test aléatoire : loi uniforme

- ✓ Test massif facile (si les oracles sont automatisés)
- ✓ Pas de biais statistique
- ✗ Très difficile de générer des comportements spécifiques

Test statistique : loi dépendant de la cible

- > Biais statistique important
 - ✓ On teste beaucoup les comportements communs
 - ✗ Mais qui sont généralement ceux qui ne posent pas de problèmes
- ✗ Loi statistique généralement dure à produire

Fuzz Testing

- > génération automatique de données aléatoires en entrée,
- > algorithmes plus ou moins malins pour explorer le comportement du programme,
- > très efficace pour détecter des plantages ou erreurs critiques,
- > particulièrement apprécié pour les questions de sécurité

Basé sur les entrées du composant testé

- ✓ aucune connaissance interne du composant nécessaire
- ✓ bien adapté pour vérifier la conformité par rapport à la spécification
- ✗ difficile de détecter des défauts précis de programmation
- ✗ difficile de produire une aide automatisée pour l'évaluation de ces tests

Bien adapté du niveau unitaire à système *

*En particulier, à haut niveau, le binaire est suffisant pour tester

Basé sur la structure du code

- ✗ nécessite le code
- ✓ le code est plus précis que la spécification
- ✗ le code est généralement gros par rapport à la spécification
- ✓ tests plus précis
- ✗ ? tests plus nombreux
- ✓ efficace pour détecter les défauts précis de programmation
- ✗ peu adapté pour détecter le non-respect d'une spécification

Principe

- > diviser le domaine d'entrée en un nombre fini de classes
- > une valeur à tester par classe uniquement

Principe

- > diviser le domaine d'entrée en un nombre fini de classes
- > une valeur à tester par classe uniquement

Exemple : valeur absolue – `int abs (int) ;`

- > trois classes d'entrée “naturelles” : < 0 , $= 0$, > 0 ,
- > on sélectionne une valeur dans chaque classe (e.g. -42, 0, 24).

Basée sur l'interface

- > Basée uniquement sur les types des données d'entrée
- > Assez facile à automatiser

Basée sur les fonctionnalités

- > Prend en compte les relations entre les variables d'entrée
- > Plus pertinent
- > Peu automatisable

```
bool search(int v, list<int> l);
```

Interface

- > $\{empty(I), \neg empty(I)\}$
- > $\{< 0, = 0, > 0\}$

Fonctionnalités

- > $\{empty(I), v \in I, \neg empty(I) \wedge e \notin I\}$

Différentes situations

- > API interne : pas de programmation défensive
 - > généralement peu d'intérêt aux tests invalides
- > API externe : programmation défensive (parfois aidée par le typage)
 - > tests invalides nécessaires
- > Interface utilisateur : programmation défensive
 - > tests invalides nécessaires

Différentes situations

- > API interne : pas de programmation défensive
 - > généralement peu d'intérêt aux tests invalides
- > API externe : programmation défensive (parfois aidée par le typage)
 - > tests invalides nécessaires
- > Interface utilisateur : programmation défensive
 - > tests invalides nécessaires

Conseil : **Attention à en faire !**

Différentes situations

- > API interne : pas de programmation défensive
 - > généralement peu d'intérêt aux tests invalides
- > API externe : programmation défensive (parfois aidée par le typage)
 - > tests invalides nécessaires
- > Interface utilisateur : programmation défensive
 - > tests invalides nécessaires

Conseil : **Attention à en faire !**

Conseil : Attention à ne pas **trop** en faire !

Tactique pour améliorer l'efficacité des jeux de données

Les erreurs se produisent souvent sur les cas limites, testons ces cas

- > s'intègre très naturellement aux tests basés sur les classes
- > génère plus de blocs, donc plus cher
- > on teste les bornes des classes d'équivalence

Exemples

- > sur une classe $[N..M]$, tester N , $N + 1$, $M - 1$, et M
- > ensemble vide, singleton
- > fichier vide, fichier de taille max, fichier juste trop gros,
- > ...

À un programme P , on associe un graphe de contrôle étiqueté $(\mathcal{S}, \mathcal{T}, \mathcal{E}, \mathcal{I}, \mathcal{F})$, où :

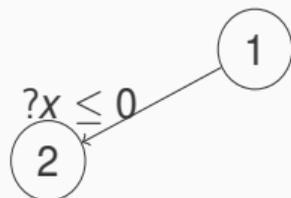
- > À chaque instruction de P on associe un nœud de \mathcal{S} .
- > $\mathcal{T} \subset (\mathcal{S} \times \mathcal{S})$ est l'ensemble des arcs
- > $\mathcal{E} : \mathcal{S} \mapsto \{x = e \mid ?e \mid \text{skip}\}$ associe une étiquette à chaque arc.
- > $\mathcal{I} \in \mathcal{S}$ représente l'état initial
- > $\mathcal{F} \in \mathcal{S}$ représente l'état final

```
if (x <= 0) {           // 1
    x = 1                // 2
} else {
    ;                    // 3
}
y = 1;                  // 4
while (! (x <= 0)) {   // 5
    y = y * x;          // 6
    x--;                // 7
}
```

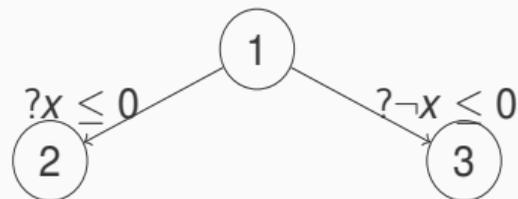
1

```
if (x <= 0) { // 1
    x = 1 // 2
} else {
    ; // 3
}
y = 1; // 4
while (! (x <= 0)) { // 5
    y = y * x; // 6
    x--; // 7
}
```

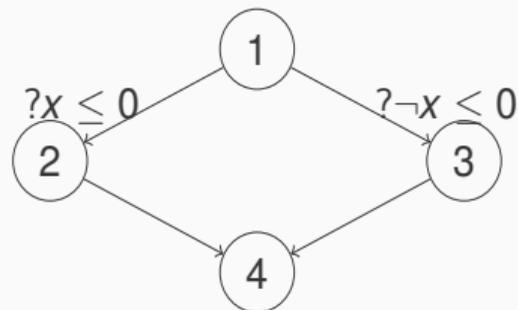
```
if (x <= 0) { // 1
    x = 1 // 2
} else {
    ; // 3
}
y = 1; // 4
while (! (x <= 0)) { // 5
    y = y * x; // 6
    x--; // 7
}
```



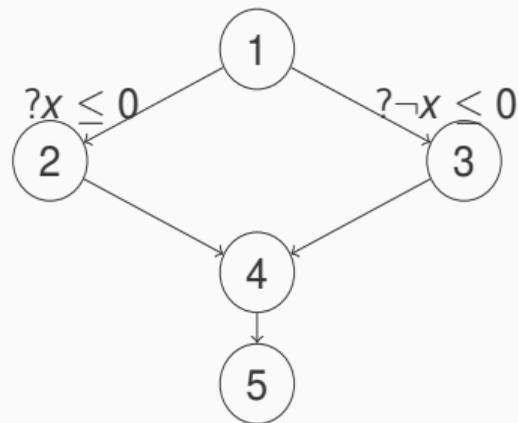
```
if (x <= 0) { // 1
    x = 1 // 2
} else {
    ; // 3
}
y = 1; // 4
while (! (x <= 0)) { // 5
    y = y * x; // 6
    x--; // 7
}
```



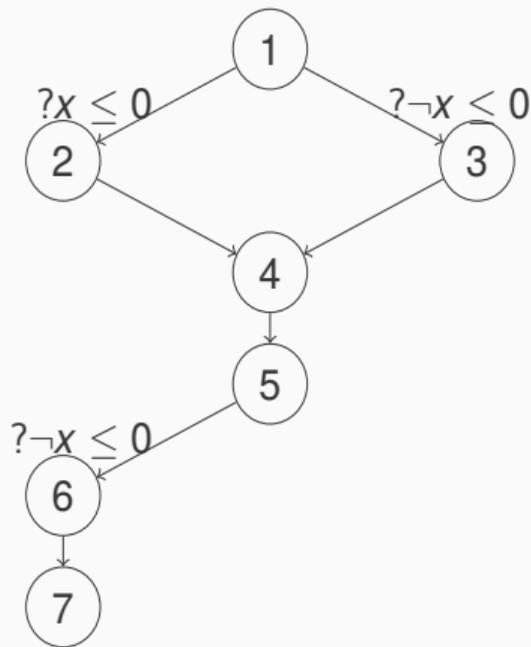
```
if (x <= 0) { // 1
    x = 1 // 2
} else {
    ; // 3
}
y = 1; // 4
while (! (x <= 0)) { // 5
    y = y * x; // 6
    x--; // 7
}
```



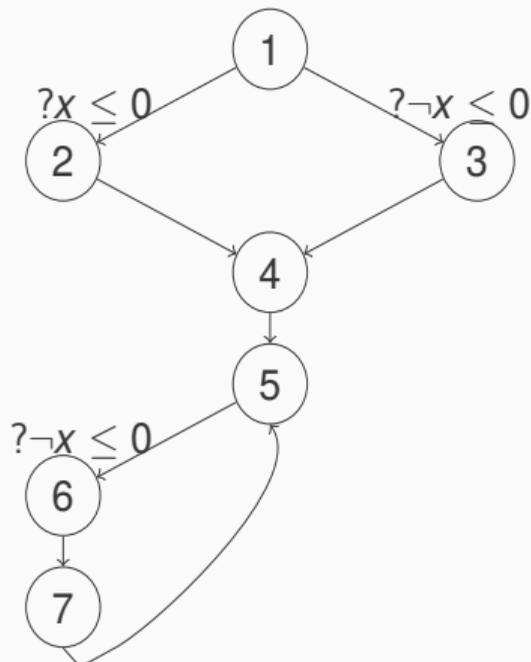
```
if (x <= 0) { // 1
    x = 1 // 2
} else {
    ; // 3
}
y = 1; // 4
while (! (x <= 0)) { // 5
    y = y * x; // 6
    x--; // 7
}
```



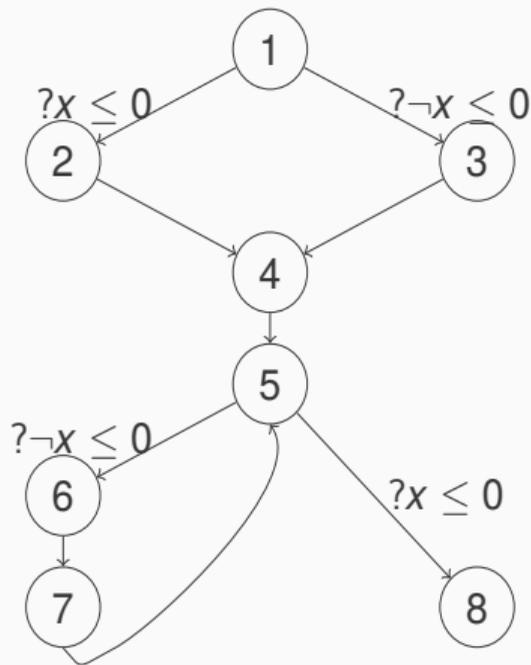
```
if (x <= 0) { // 1
    x = 1 // 2
} else {
    ; // 3
}
y = 1; // 4
while (! (x <= 0)) { // 5
    y = y * x; // 6
    x--; // 7
}
```



```
if (x <= 0) { // 1
    x = 1 // 2
} else {
    ; // 3
}
y = 1; // 4
while (! (x <= 0)) { // 5
    y = y * x; // 6
    x--; // 7
}
```



```
if (x <= 0) { // 1
    x = 1 // 2
} else {
    ; // 3
}
y = 1; // 4
while (! (x <= 0)) { // 5
    y = y * x; // 6
    x--; // 7
}
```

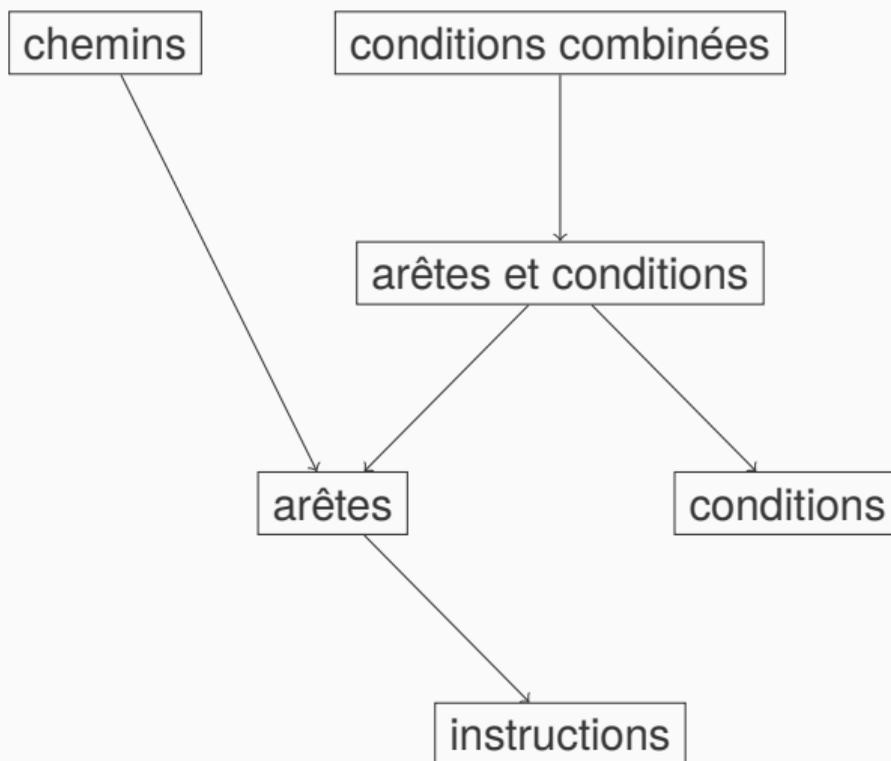


- > **Instructions** : chaque **instruction** exécutable (nœud), exceptée `skip`, du programme apparaît dans un cas de test.
- > **Arêtes** : chaque **arête** exécutable (exceptée `skip`) du graphe de contrôle apparaît dans un cas de test.
- > **Conditions** : chaque **prédicat atomique** du programme apparaît dans un cas de test, **évalué à vrai** et une autre **évalué à faux**.
- > **Arêtes et conditions** : couverture des **arêtes** + couverture des **conditions**
- > **Couverture des conditions combinées** : similaire à la couverture des conditions, mais **toutes les combinaisons de prédicats interprétés à vrai et faux** doivent être couvertes

Chaque chemin exécutable du graphe de contrôle apparaît dans un cas de test.

- > Boucles dont le nombre d'itérations est inconnu!
 - > Couverture des k -chemins : pour chaque boucle, tous les chemins contenant i répétitions de la boucle, $0 \leq i \leq k$.
 - > Couverture PLCS (Portion Linéaire de Code suivie d'un Saut) : un saut est l'entrée, la sortie ou l'arrivée d'un branchement.
- > Caractère **indécidable** des chemins/instructions/arêtes **exécutable** requis par certains critères
 - > la présence de code non exécutable est souvent le signe d'un **code mal écrit**, voire faux
 - > interdit dans les programmes critiques

HIÉRARCHIE DES CRITÈRES DE COUVERTURE STRUCTURELLE

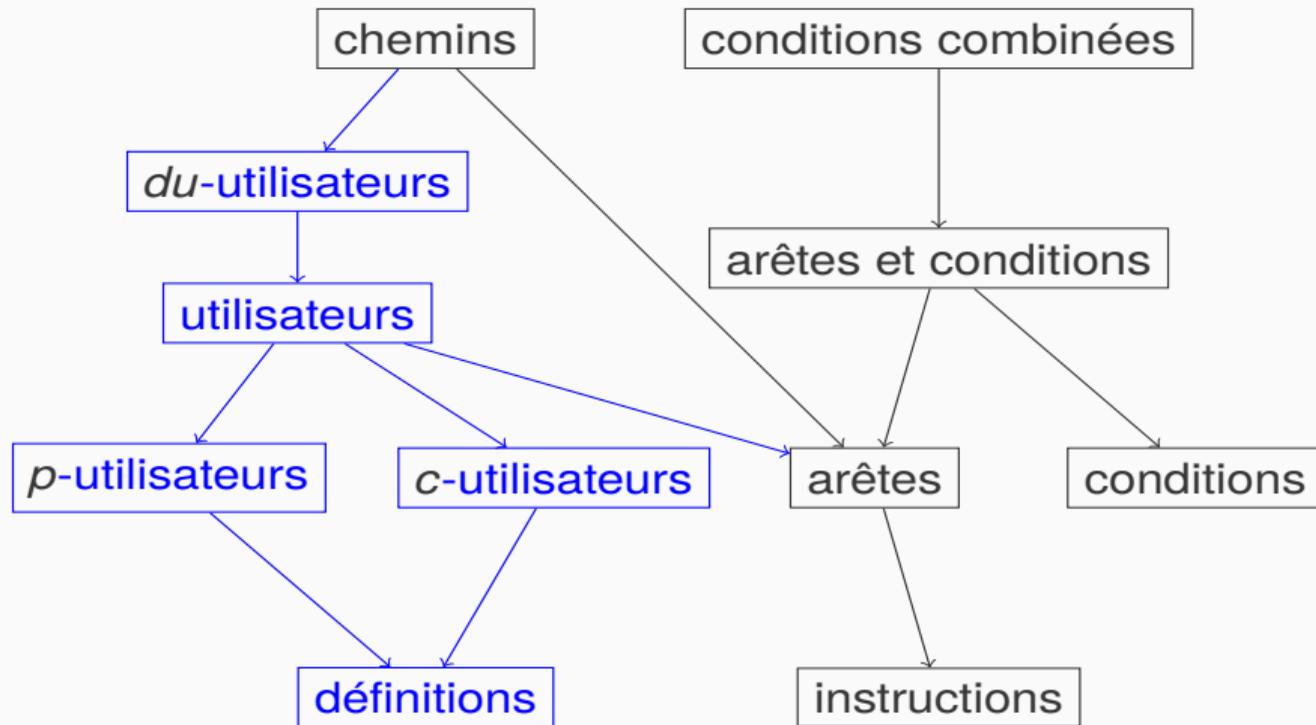


Les critères fondés sur le flot de contrôle ne sont pas assez fins

- > **variable définie** : lorsque sa valeur est modifiée
- > **variable utilisée** :
 - > **p-utilisée** : lorsque sa valeur est lue dans un prédicat de branchement
 - > **c-utilisée** : dans les autres cas (valeur lue pour un calcul)
- > une **instruction** i est **utilisatrice** d'une variable x par rapport à une instruction de définition j si :
 - > x est définie en j
 - > x est utilisée en i
 - > il existe un chemin sur lequel x n'est pas redéfinie entre i et j

- > **Couverture toutes les définitions** : pour chaque définition, un chemin utilisateur apparaît dans un cas de test.
- > **Couverture tous les p -utilisateurs** : pour chaque définition, chaque instruction p -utilisatrice apparaît dans un cas de test.
- > **Couverture tous les c -utilisateurs** : pour chaque définition, chaque instruction c -utilisatrice apparaît dans un cas de test.
- > **Couverture tous les utilisateurs** : p -utilisateurs + c -utilisateurs.
- > **Couverture toutes les du -utilisateurs** : pour chaque définition d , tous les chemins entre d et chacune de ses instructions utilisatrices sont couverts.

HIÉRARCHIE DES CRITÈRES DE COUVERTURE STRUCTURELLE



D'autres critères de couverture existent

Que déduire si la campagne de tests trouve **peu d'erreurs** ?

- > **optimiste** : le programme est très bon
- > **pessimiste** : les tests sont mauvais

Que déduire si la campagne de tests trouve **beaucoup d'erreurs**

- > **optimiste** : la majorité des erreurs ont été corrigées
- > **pessimiste** : le programme est de mauvaise qualité : il reste beaucoup d'erreurs

> À la fin du budget alloué ?

- > À la fin du budget alloué ? **Mauvaise idée** : autant ne rien faire

- > À la fin du budget alloué ? **Mauvaise idée** : autant ne rien faire
- > tous les tests passent ?

- > À la fin du budget alloué ? **Mauvaise idée** : autant ne rien faire
- > tous les tests passent ? **Mauvaise idée** : aucun intérêt de trouver des bugs

- > À la fin du budget alloué ? **Mauvaise idée** : autant ne rien faire
- > tous les tests passent ? **Mauvaise idée** : aucun intérêt de trouver des bugs
- > (presque) plus de bugs trouvés ?

- > À la fin du budget alloué ? **Mauvaise idée** : autant ne rien faire
- > tous les tests passent ? **Mauvaise idée** : aucun intérêt de trouver des bugs
- > (presque) plus de bugs trouvés ? **Fausse bonne idée** : effets de saturation

- > À la fin du budget alloué ? **Mauvaise idée** : autant ne rien faire
- > tous les tests passent ? **Mauvaise idée** : aucun intérêt de trouver des bugs
- > (presque) plus de bugs trouvés ? **Fausse bonne idée** : effets de saturation
- > quand le critère fixé est atteint ?

- > À la fin du budget alloué ? **Mauvaise idée** : autant ne rien faire
- > tous les tests passent ? **Mauvaise idée** : aucun intérêt de trouver des bugs
- > (presque) plus de bugs trouvés ? **Fausse bonne idée** : effets de saturation
- > quand le critère fixé est atteint ? **Bonne idée**, mais :
 - > attention à ne pas chercher à atteindre le critère plutôt qu'à trouver des bugs
 - > distinguer objectifs et moyens

- > nombre d'erreurs trouvées ne croît pas linéairement avec le temps
- > 2 phases :
 - > augmentation du nombre d'erreurs trouvées puis atteinte d'un palier : effet de seuil
 - > stagnation du nombre d'erreurs trouvées
- > conséquences :
 - > ne pas arrêter les tests en période d'augmentation du nombre d'erreurs trouvées
 - > ne pas arrêter les tests dès le premier palier

Proposition réaliste (?) de G. J. Myers

- > constat 1 : le but est de trouver des bugs
- > constat 2 : il y a toujours des bugs
- > conséquence : exprimer le critère d'arrêt en fonction du nombre de bugs trouvés

- > programme de 10 kLOC
- > estimation de 50 erreurs par kLOC, donc 500 erreurs
- > supposition de 200 erreurs de codage et 300 erreurs de conception
- > volonté de trouver 98% des premières et 95% des secondes
- > donc : trouver 196 erreurs de codage et 285 erreurs de conception
- > critère d'arrêt choisi :
 - > test unitaire : 130 erreurs trouvées (65%)
 - > test d'intégration : 240 trouvées (30% de 200 et 60% de 300) ou 4 mois écoulés
 - > test système : 111 erreurs trouvées ou 3 mois écoulés

“Testing can only reveal the presence of errors but never their absence.”

(E. W. Dijkstra, Notes on Structured Programming, 1972)

- > Correspond aux **besoins réels** de beaucoup d'industriels,
- > Activité “ancienne”
 - > **bien intégrée aux cycles de développements** actuels
 - > ne nécessite pas de modifier les processus industriels
 - > ne nécessite pas de modifier les équipes
 - > existence de personnes qualifiées sur le marché
- > activité simple
- > **le retour sur investissement** proportionnel aux efforts
- > **peut trouver des erreurs indépendamment de celles cherchées**

- > Glenford J. Myers *The Art of Software Testing*
J. Willey & Sons, 1979
- > Boris Beizer *Software Testing Techniques*
Van Nostrand Reinhold, 1990 (seconde édition)
- > S. Xanthakis, M. Maurice, A. de Amescua, O. Hourri, L. Griffet
Test & Contrôle des Logiciels
Edition EC2, 1994
- > S. Xanthakis, P. Régnier, C. Karapoulios *Le test des logiciels*
Edition Hermes Sciences, 1999
- > Aditya P. Mathur *Foundations of Software Testing*
Edition Pearson Education, 2008

Partie III

Spécification

Comment spécifier ?

Différents types de spécification

Différence fondamentale entre programme et spécification

- > la spécification est le “quoi”, ce que l'on veut du système,
- > le programme est le “comment”, ce que fait le système pour accomplir le “quoi”

La spécification est :

- > le point de départ pour le développement
- > le point de départ pour la vérification

Comment ?

- > documents de spécification
- > documentation du code

Avantages et inconvénients

- ✓ indispensables
- ✓ les moins difficiles à produire
- ✓ les plus facilement compréhensibles
- ✗ sources d'ambigüité, d'incohérence, d'incorrection, d'incomplétude,
- ✗ incompréhensibles par les outils

Nécessaire

Comment ?

- > diagrammes UML
- > OCL (Object Constraint Language), le langage logique d'UML

Avantages et inconvénients

- ✓ permettent de spécifier facilement le modèle d'une application
- ✓ permettent de générer automatiquement le code trivial
- ✓ existence de pas mal outils facilitant leurs déploiements
- ✗ mêmes problèmes que les langues naturelles, mais de manière plus limitée

Souvent suffisant

Comment ?

- fondée sur des formalismes mathématiques/logiques solides

Avantages et inconvénients

- ✓ augmentent la compréhension et la confiance dans le code spécifié
- ✓ possibilité de vérifier des propriétés à propos des spécifications
- ✓ possibilité de prouver qu'un programme satisfait sa spécification
- ✗ nécessite de connaître et comprendre les formalismes sous-jacents
- ✗ problèmes des spécifications erronées, mais faibles risques

Nécessaire pour un code critique

Définition

Une spécification est exécutable lorsque l'on peut produire un code qui permet sa vérification **pendant l'exécution**.

Le langage de spécification doit être adapté

- ✗ logique généralement moins expressive,
- ✓ possibilité de vérifier des propriétés à **runtime**.

Composants :

- > variables,
- > prédicats/rerelations à propos des variables,
- > connecteurs logiques (\neg , \wedge , \vee , \Rightarrow , \Leftrightarrow),
- > quantificateurs (\exists , \forall)

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \Rightarrow B$	$A \Leftrightarrow B$
F	F	V	F	F	V	V
F	V	V	F	V	V	F
V	F	F	F	V	F	F
V	V	F	V	V	V	V

$\forall x.P(x)$

- > “pour tout x , $P(x)$ est vraie”
- > ex : $\forall n : nat, n + 1 > n$

$\exists x.P(x)$

- > “il existe un x tel que $P(x)$ est vraie”
- > ex : $\exists n : nat, n > 0$

Erreur commune avec l'existentielle et l'implication !

$$\exists x, P(x) \Rightarrow Q(x)$$

Présentation

- > langage de spécification pour le langage C
- > basé sur la notion de **contrat**
- > permet de spécifier des **propriétés fonctionnelles** de programmes

Composants basiques

- > une logique du premier ordre typée,
- > expressions C pures,
- > types du $\mathbb{C} + \mathbb{Z} + \mathbb{R}$
- > prédicats et fonctions builtins

C et la mémoire

- > les pointeurs sont omniprésents en C,
- > et sont causes d'erreur,
- > besoin important d'une manière de modéliser cela.

Builtins ACSL

- > `\valid(ptr), \valid(ptr+(0..n))`
- > `\separated(ptr, other+(0..m))`
- > `\initialized(ptr)`
- > ...

- > **But** : spécification de fonctions
- > **Approche** : indiquer des assertions à propos des fonctions
 - > **Précondition** supposée vraie à l'entrée (garanti par l'appelant)
 - > **Postcondition** doit être vraie en sortie (garanti par la fonction)
- > quand la précondition est fausse, rien n'est garanti

But d'un contrat

Refléter la spécification informelle

Annotations au sein du code :

```
/*@  
  requires ... ;  
  ensures  ... ;  
*/  
void fonction( ... ){ ... }
```

- > introduction d'une annotation : /*@ (ou //@),
- > introduction d'une précondition : **requires**,
- > introduction d'une postcondition : **ensures**.

```
/*@  
  behavior <name> :  
    assumes   ... ;  
    requires  ... ;  
    ensures   ... ;  
  
    ...  
    disjoint behaviors ;  
    complete behaviors ;  
*/  
void function( ... ){ ... }
```

- > reflet du besoin de partitionner la spécification
- > permet justement de vérifier la complétude et la disjonction des cas

```
int main(void) {  
    int x = 42 ;  
    //@ assert x == 42 ;  
}
```

```
int abs(int x) {  
    if (x < 0) return -x;  
    return x ;  
}
```

```
/*@ ensures x < 0 ==> \result == -x ;  
   ensures x >= 0 ==> \result == x ;  
*/  
int abs(int x) {  
    if (x < 0) return -x;  
    return x ;  
}
```

```
/*@ behavior negative:
    assumes x < 0 ;
    ensures \result == -x ;
behavior positive:
    assumes x >= 0 ;
    ensures \result == x ;

    disjoint behaviors ;
    complete behaviors ;

*/
int abs(int x) {
    if (x < 0) return -x;
    return x ;
}
```

```
/*@ ensures \result == \abs(x) ; */  
int abs(int x) {  
    if (x < 0) return -x;  
    return x ;  
}
```

```
/*@ ensures \result == \abs(x) ; */  
int abs(int x) {  
    if (x < 0) return -x;  
    return x ;  
}
```

La spécification n'est peut-être pas parfaite ceci dit, on verra ...

```
size_t search(int v, int* arr, size_t l) {  
    for(size_t i = 0; i < l; i++)  
        if(arr[i] == v) return i;  
    return l;  
}
```

```
/*@ behavior in:  
    assumes \exists size_t o ; 0 <= o < l && arr[o] == v;  
    ensures 0 <= \result < l && arr[\result] == v;  
behavior notin:  
    assumes \forall size_t o ; 0 <= o < l ==> arr[o] != v;  
    ensures \result == 1;  
disjoint behaviors;  
complete behaviors;  
*/ // Complete ? Ou pas ? On verra  
size_t search(int v, int* arr, size_t l);
```

```
size_t bsearch(int* arr, size_t l, int value){
    if(l == 0) return 0 ;

    size_t low = 0 ;
    size_t up = l ;

    while(low < up){
        size_t mid = low + (up - low)/2 ;
        if      (arr[mid] > value) up = mid ;
        else if(arr[mid] < value) low = mid+1 ;
        else return mid ;
    }
    return l ;
}
```

```
/*@ requires
    \forall integer i, j ;
        0 <= i <= j < l ==> arr[i] <= arr[j] ;
    ...
*/
size_t bsearch(int* arr, size_t l, int value);
```

- > la spécification peut être longue par rapport au code
- > la spécification peut correspondre à plusieurs codes
- > quid du caractère exécutable d'ACSL ?

- > **Eiffel** : Analysis, Design and Programming Language 2nd edition. 2006.
`https://se.inf.ethz.ch/~meyer/ongoing/etl/EIFFEL-STANDARD.pdf`
- > **Java Modeling Language (JML)**. 2000.
`http://www.eecs.ucf.edu/~leavens/JML/index.shtml`
- > **Spec#**. 2004.
`http://research.microsoft.com/en-us/projects/specsharp`
- > **ANSI C Specification Language (ACSL)**.
`https://frama-c.com/html/acsl.html`
- > **Executable ANSI C Specification Language (E-ACSL)**.
`https://frama-c.com/fc-plugins/e-acsl.html`
- > **Spark-2014**.
`https://docs.adacore.com/spark2014-docs/html/lrm/`

Partie IV

Vérification à l'exécution

Spécification exécutable ?

Comment générer le code correspondant à une spécification ?

Comment transformer cette annotation ?

```
//@ assert a > INT_MIN;
```

Comment transformer cette annotation ?

```
//@ assert a > INT_MIN;
```

```
assert (a > INT_MIN);
```

Est-ce toujours si simple ?

Comment transformer cette annotation ?

```
//@ assert (a-b) / c > INT_MIN;
```

Comment transformer cette annotation ?

```
//@ assert (a-b) / c > INT_MIN;
```

```
assert ((a-b) / c > INT_MIN);
```

Comment transformer cette annotation ?

```
//@ assert (a-b) / c > INT_MIN;
```

```
assert ((a-b) / c > INT_MIN);
```

Non! Risque de division par 0, de débordement ...

Les transformations ne doivent pas introduire de bugs!

Combien de **comportements indéterminés** existent dans la norme ISO/C ?

Combien de **comportements indéterminés** existent dans la norme ISO/C ?

191

Combien de **comportements indéterminés** existent dans la norme ISO/C ?

191

Simple

Subtle



Modulo 0

Strict aliasing violation

Shift with negative exponent

Data-race

Instrumentation automatique des erreurs à l'exécution

Accès invalide

```
//@ assert \valid(p+i) ;  
int x = p[i] ;
```

Division par zéro

```
//@ assert d != 0 ;  
int x = 42 / d ;
```

Débordement arithmétique

```
//@ assert INT_MIN <= a+b ;  
//@ assert a+b <= INT_MAX ;  
int x = a + b ;
```

- > Initialisation,
- > **float** vers **int**,
- > *shift* hors bornes,
- > ...

```
assert ((long) a - b <= INT_MAX) ;  
//@ assert a - b <= INT_MAX ;  
assert ((long) a - b >= INT_MIN) ;  
//@ assert a - b >= INT_MIN ;  
assert (c != 0) ;  
//@ assert c != 0  
assert ((a-b) / c > INT_MIN) ;  
//@ assert (a-b) / c > INT_MIN;
```

Méthode

- > On génère une ligne,
- > On l'analyse avec Rte,
- > On génère (ou pas) une alarme,
- > Et ainsi de suite.

Comment transformer cette annotation ?

```
//@ assert \valid(p+(0..i)) ;
```

Comment transformer cette annotation ?

```
//@ assert \valid(p+(0..i)) ;
```

Pas de solution simple

- > il faut modéliser la mémoire à l'exécution,
- > contrôler que tous les accès sont corrects,
- > **coût potentiellement important à l'exécution**

Comment transformer cette annotation ?

```
//@ assert \forall integer i ; 0 <= i <= 42 ==> P(i);
```

Comment transformer cette annotation ?

```
//@ assert \forall integer i ; 0 <= i <= 42 ==> P(i);
```

```
for(int i = 0 ; i <= 42 ; ++i){  
    assert(translated_P(i));  
}
```

Comment transformer cette annotation ?

```
//@ assert \exists integer i ; 0 <= i <= 42 && P(i);
```

Comment transformer cette annotation ?

```
//@ assert \exists integer i ; 0 <= i <= 42 && P(i);
```

```
int i = 0 ;  
while (i <= 42) {  
    if (translated_P(i)) break;  
    i++;  
}  
assert (i <= 42);
```

Comment transformer cette annotation ?

```
//@ assert \forall integer i ; P(i);  
//@ assert \exists integer i ; P(i);
```

Comment transformer cette annotation ?

```
//@ assert \forall integer i ; P(i);  
//@ assert \exists integer i ; P(i);
```

On ne peut pas

- > On ne peut pas traduire toute la logique ACSL,
- > nécessité de borner les quantifications

Un langage et un greffon de Frama-C

Langage :

- > sous-ensemble exécutable d'ACSL
- > les quantifications doivent être bornées
- > toutes les constructions logiques ne sont pas supportées

Greffon :

- > transforme les annotations E-ACSL en code exécutable
- > de manière correcte
- > modélise la mémoire pour détecter les violations

```
int main(void) {  
    int a = 42 ;  
    //@ assert a == 42;  
}
```

On lance la commande :

```
frama-c e-acsl-t1.c -e-acsl -then-last -ocode res.c -print
```

Voir en particulier :

```
__gen_e_acsl_assert_data.kind = "Assertion";  
__gen_e_acsl_assert_data.pred_txt = "a \342\211\241 42";  
__gen_e_acsl_assert_data.file = "e-acsl-t1.c";  
__gen_e_acsl_assert_data.fct = "main";  
__gen_e_acsl_assert_data.line = 3;  
__e_acsl_assert(a == 42, & __gen_e_acsl_assert_data);
```

```
e-acsl-gcc.sh
```

- > permet la compilation et le *link* aux bibliothèques E-ACSL
- > peut aussi lancer Frama-C et faire l'instrumentation

```
e-acsl-gcc.sh main.c -c -O output
```

Génère :

- > `a.out.frama.c` le fichier instrumenté
- > `output` l'exécutable sans le monitoring à l'exécution
- > `output.e-acsl` l'exécutable avec le monitoring à l'exécution

Introduisons une erreur :

```
int main(void) {  
    int a = 41 ;  
    //@ assert a == 42;  
}
```

- > le fichier sans instrumentation s'exécute sans message
- > avec instrumentation, une erreur est levée

```
e-acsl-t2.c: In function 'main'  
e-acsl-t2.c:3: Error: Assertion failed:  
The failing predicate is:  
a == 42.  
With values:  
- a: 41
```

```
int main(void) {  
    int a[10];  
    for(int i = 0; i <= 10; i++) a[i] = 1 ;  
  
    /*@ assert \forall integer i ;  
           0 <= i < 10 ==> a[i] == 1; */  
}
```

```
int main(void) {  
    int a[10];  
    for(int i = 0; i <= 10; i++) a[i] = 1 ;  
  
    /*@ assert \forall integer i ;  
        0 <= i < 10 ==> a[i] == 1; */  
}
```

Remarques

- > on trouve bien une boucle,
- > on y trouve les contrôles à runtime,
- > **les potentielles RTEs du code d'origine ne sont pas instrumentées**
- > option `--rte=all`

On peut utiliser E-ACSL pour robustifier un outil.

Exemple :

- > monitorer avec E-ACSL **mais** :
 - > n'échouer sévèrement que sur les *undefined behavior*
 - > logger le reste

E-ACSL permet de changer l'action à réaliser sur l'assertion.

```
typedef struct eacsl_assert_data_t {  
    int blocking;  
    const char *kind;  
    const char *name; const char *pred_txt;  
    const char *file; int line; const char *fct;  
    eacsl_assert_data_value_t *values;  
} eacsl_assert_data_t;
```

```
void __e_acsl_assert(int pred, __e_acsl_assert_data_t * d) ;
```

```
e-acsl-gcc.sh -c file.c -O exe --external-assert=my_assert.c
```

- > E-ACSL transforme des annotations ACSL en code C,
 - > cela nécessite de restreindre le langage à son fragment **exécutable**
- > on l'utilise principalement pour monitorer du code,
- > les annotations peuvent provenir de l'utilisateur ou être générées
- > peut aller plus loin :
 - > analyse de consommation mémoire, *use-after-free*, ...
- > peut aider des outils de tests
 - > détecte des comportements difficiles à détecter
 - > permet d'exprimer directement des propriétés logiques
- > comme complément à des outils de transformation de code
 - > propriétés temporelles (Aoraï)
 - > flot d'information (SecureFlow)
 - > autre analyse "faite maison" ...

Partie V

Analyse statique correcte

Vérifier formellement l'absence de *runtime errors* ?

Interprétation abstraite ?

Sémantique opérationnelle

Une **sémantique opérationnelle** est définie par un **système de règles** indiquant comment **calculer** les expressions et les instructions du langage

- > **à grand pas** : elle décrit comment calculer le résultat final
- > **à petit pas** : elle décrit comment effectuer chaque étape de calcul.

Sémantique de traces

Une **sémantique de trace** décrit l'ensemble des traces d'exécution d'un programme.

Valeurs

L'ensemble Val des valeurs est l'ensemble des constantes entières.

Environnement d'évaluation

Un **environnement** (d'évaluation) ρ est une **fonction partielle de Var dans Val** . On note $\rho[x \mapsto z]$ l'environnement ρ' tel que, pour toute variable y :

$$\rho'(y) \triangleq \begin{cases} z & \text{si } x = y \\ \rho(y) & \text{sinon et si } y \in \text{dom}(\rho). \end{cases}$$

On note Env_ϵ l'ensemble des environnements : $Env_\epsilon \triangleq Var \multimap Val$.

Jugement d'évaluation

On note $\rho \vdash e \Downarrow_\epsilon z$ (resp. $\rho \vdash i \Downarrow \rho'$) l'évaluation de l'expression e (resp. l'instruction i) dans l'environnement ρ conduisant à la valeur $z \in Val$ (resp. l'environnement ρ').

$$\frac{x \in \text{dom}(\rho) \quad \rho(x) = z}{\rho \vdash x \Downarrow_{\epsilon} z}$$

$$\frac{}{\rho \vdash z \Downarrow_{\epsilon} z}$$

$$\frac{\rho \vdash e \Downarrow_{\epsilon} z \quad z' = -z}{\rho \vdash -e \Downarrow_{\epsilon} z'}$$

$$\frac{\rho \vdash e_1 \Downarrow_{\epsilon} z_1 \quad \rho \vdash e_2 \Downarrow_{\epsilon} z_2 \quad z' = z_1 + z_2}{\rho \vdash e_1 + e_2 \Downarrow_{\epsilon} z'}$$

$$\frac{\rho \vdash e_1 \Downarrow_{\epsilon} z_1 \quad \rho \vdash e_2 \Downarrow_{\epsilon} z_2 \quad z' = z_1 / z_2 \quad z_2 \neq 0}{\rho \vdash e_1 / e_2 \Downarrow_{\epsilon} z'}$$

$$\frac{\rho \vdash e \Downarrow_{\epsilon} 0}{\rho \vdash !e \Downarrow_{\epsilon} 1}$$

$$\frac{\rho \vdash e \Downarrow_{\epsilon} z \quad z \neq 0}{\rho \vdash !e \Downarrow_{\epsilon} 0}$$

$$\frac{\rho \vdash e_1 \Downarrow_{\epsilon} z_1 \quad \rho \vdash e_2 \Downarrow_{\epsilon} z_2 \quad z_1 = z_2}{\rho \vdash e_1 = e_2 \Downarrow_{\epsilon} 1}$$

$$\frac{\rho \vdash e_1 \Downarrow_{\epsilon} z_1 \quad \rho \vdash e_2 \Downarrow_{\epsilon} z_2 \quad z_1 \neq z_2}{\rho \vdash e_1 = e_2 \Downarrow_{\epsilon} 0}$$

$$\frac{\rho \vdash e_1 \Downarrow_{\epsilon} z_1 \quad \rho \vdash e_2 \Downarrow_{\epsilon} z_2 \quad z_1 \leq z_2}{\rho \vdash e_1 \leq e_2 \Downarrow_{\epsilon} 1}$$

$$\frac{\rho \vdash e_1 \Downarrow_{\epsilon} z_1 \quad \rho \vdash e_2 \Downarrow_{\epsilon} z_2 \quad z_1 > z_2}{\rho \vdash e_1 \leq e_2 \Downarrow_{\epsilon} 0}$$

$$\frac{}{\rho \vdash \text{Skip} \Downarrow \rho} \quad \frac{\rho \vdash \mathbf{e} \Downarrow_{\epsilon} z}{\rho \vdash x := \mathbf{e} \Downarrow \rho[x \mapsto z]}$$

$$\frac{\rho \vdash \mathbf{e} \Downarrow_{\epsilon} 0 \quad i_2 \Downarrow \rho_2}{\rho \vdash \text{if } \mathbf{e} \text{ then } i_1 \text{ else } i_2 \text{ fi} \Downarrow \rho_2}$$

$$\frac{\rho \vdash \mathbf{e} \Downarrow_{\epsilon} z \quad z \neq 0 \quad i_1 \Downarrow \rho_1}{\rho \vdash \text{if } \mathbf{e} \text{ then } i_1 \text{ else } i_2 \text{ fi} \Downarrow \rho_1}$$

$$\frac{\rho \vdash i_1 \Downarrow \rho_1 \quad \rho_1 \vdash i_2 \Downarrow \rho_2}{\rho \vdash i_1; i_2 \Downarrow \rho_2}$$

$$\frac{\rho \vdash \text{if } \mathbf{e} \text{ then } i; \text{while } \mathbf{e} \text{ do } i \text{ end else Skip fi} \Downarrow \rho'}{\rho \vdash \text{while } \mathbf{e} \text{ do } i \text{ end} \Downarrow \rho'}$$

$$\begin{array}{c}
 \frac{}{\overline{\emptyset \vdash 2 \Downarrow_{\epsilon} 2}} \\
 \hline
 \emptyset \vdash x := 2 \Downarrow \rho_1
 \end{array}
 \qquad
 \frac{
 \begin{array}{c}
 x \in \text{dom}(\rho_1) \\
 \rho_1(x) = 2 \\
 \hline
 \rho_1 \vdash x \Downarrow_{\epsilon} 2
 \end{array}
 \quad
 \frac{
 \overline{\rho_1 \vdash 3 \Downarrow_{\epsilon} 3} \quad 6 = 2 \times 3
 }{\rho_1 \vdash x * 3 \Downarrow_{\epsilon} \rho_2}
 }{\rho_1 \vdash x := x * 3 \Downarrow \rho_2}$$

$$\emptyset \vdash x := 2; x := x * 3 \Downarrow \rho_2$$

Notations :

- > $\rho_1 = \{x \mapsto 2\}$
- > $\rho_2 = \{x \mapsto 6\}$

Déterminisme

Cette sémantique est **déterministe** : au plus une règle est applicable à chaque fois.

Programmes erronés et ne terminant pas

- > Un programme contenant une erreur (e.g. division par zéro) n'a pas d'état résultant
- > Un programme ne terminant pas non plus
- > **Ils sont indistinguables**

Avantages et inconvénients

- ✓ proche d'une implémentation déclarative (OCaml, Prolog)
- ✗ erreurs et non-terminaison ne sont pas distinguables

À un programme P , on associe un graphe de contrôle étiqueté $(\mathcal{S}, \mathcal{T}, \mathcal{E}, \mathcal{I}, \mathcal{F})$, où :

- > À chaque instruction de P on associe un nœud de \mathcal{S} .
- > $\mathcal{T} \subset (\mathcal{S} \times \mathcal{S})$ est l'ensemble des arcs
- > $\mathcal{E} : \mathcal{S} \mapsto \{x := e \mid ?e \mid \text{skip}\}$ associe une étiquette à chaque arc.
- > $\mathcal{I} \in \mathcal{S}$ représente l'état initial
- > $\mathcal{F} \in \mathcal{S}$ représente l'état final

À un programme P , on associe un graphe de contrôle étiqueté $(\mathcal{S}, \mathcal{T}, \mathcal{E}, \mathcal{I}, \mathcal{F})$, où :

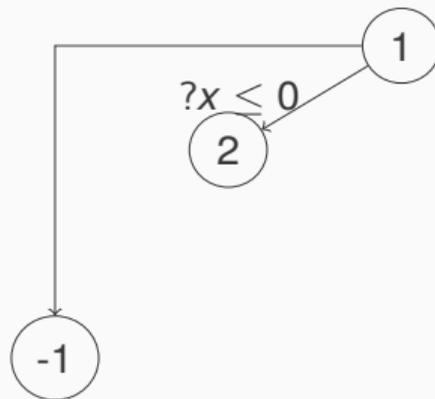
- À chaque instruction de P on associe deux nœuds de \mathcal{S} qui représentent les états avant et après l'instruction.
- $\mathcal{T} \subset (\mathcal{S} \times \mathcal{S})$ est l'ensemble des arcs
- $\mathcal{E} : \mathcal{S} \mapsto \{x := e \mid ?e \mid \text{skip}\}$ associe une étiquette à chaque arc.
- $\mathcal{I} \in \mathcal{S}$ représente l'état initial
- $\mathcal{F} \in \mathcal{S}$ représente l'état final
- $\epsilon \in \mathcal{S}$ représente l'état d'erreur

```
if (x <= 0) { // 1
  x = 1 // 2
} else {
  ; // 3
}
y = 1; // 4
while (! (x <= 0)) { // 5
  y = y * x; // 6
  x--; // 7
}
```

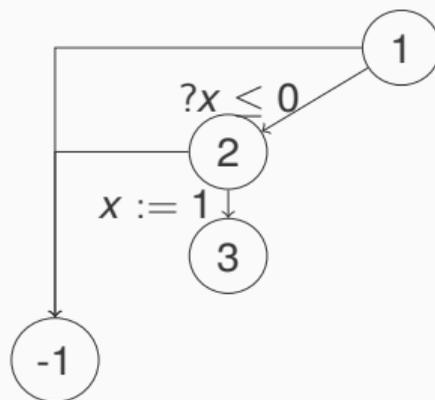
1

```
if (x <= 0) { // 1
  x = 1 // 2
} else {
  ; // 3
}
y = 1; // 4
while (! (x <= 0)) { // 5
  y = y * x; // 6
  x--; // 7
}
```

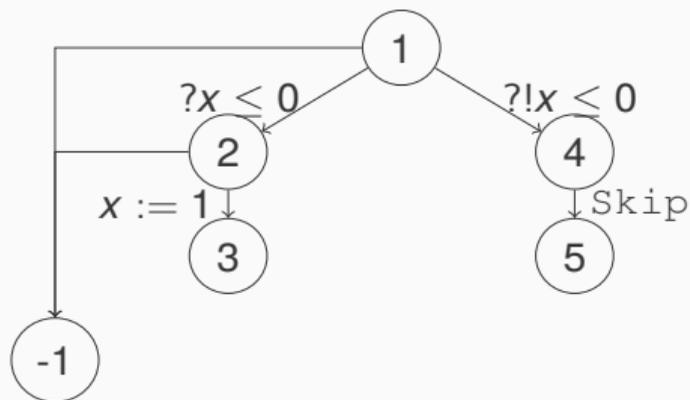
```
if (x <= 0) {           // 1
    x = 1                // 2
} else {
    ;                    // 3
}
y = 1;                  // 4
while (! (x <= 0)) {   // 5
    y = y * x;          // 6
    x--;                // 7
}
```



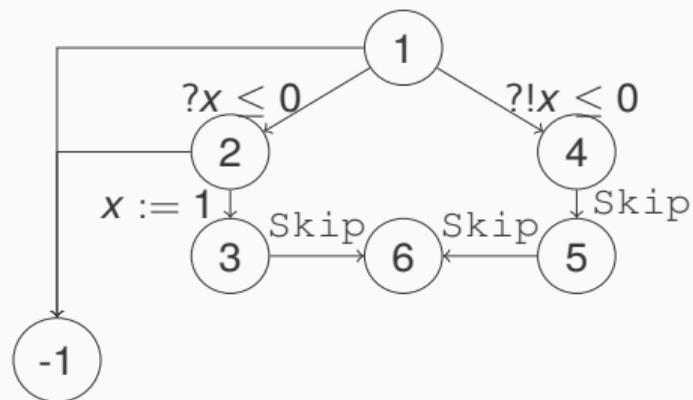
```
if (x <= 0) { // 1
  x = 1 // 2
} else {
  ; // 3
}
y = 1; // 4
while (! (x <= 0)) { // 5
  y = y * x; // 6
  x--; // 7
}
```



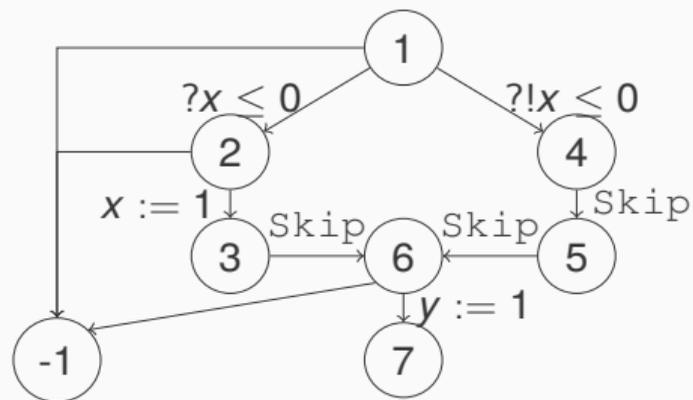
```
if (x <= 0) { // 1
  x = 1 // 2
} else {
  ; // 3
}
y = 1; // 4
while (! (x <= 0)) { // 5
  y = y * x; // 6
  x--; // 7
}
```



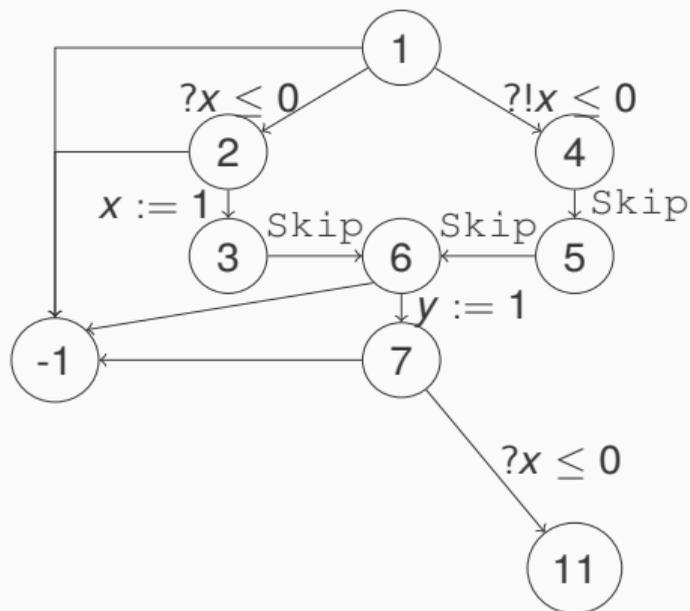
```
if (x <= 0) { // 1
  x = 1 // 2
} else {
  ; // 3
}
y = 1; // 4
while (! (x <= 0)) { // 5
  y = y * x; // 6
  x--; // 7
}
```



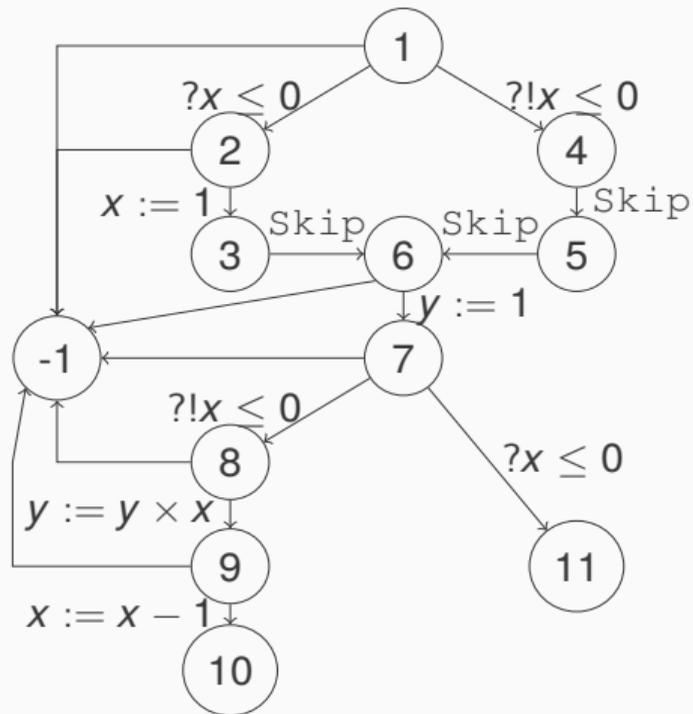
```
if (x <= 0) { // 1
  x = 1 // 2
} else {
  ; // 3
}
y = 1; // 4
while (! (x <= 0)) { // 5
  y = y * x; // 6
  x--; // 7
}
```



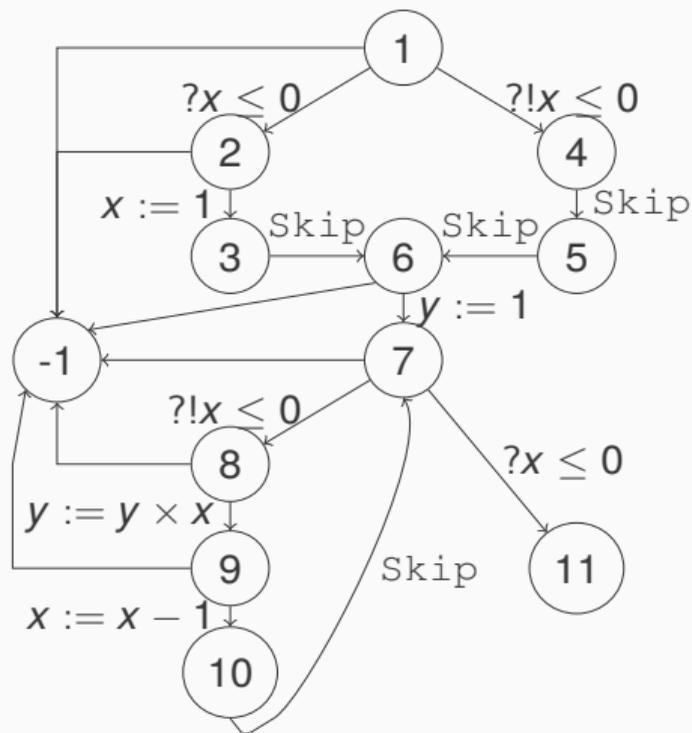
```
if (x <= 0) { // 1
  x = 1 // 2
} else {
  ; // 3
}
y = 1; // 4
while (! (x <= 0)) { // 5
  y = y * x; // 6
  x--; // 7
}
```



```
if (x <= 0) { // 1
  x = 1 // 2
} else {
  ; // 3
}
y = 1; // 4
while (! (x <= 0)) { // 5
  y = y * x; // 6
  x--; // 7
}
```



```
if (x <= 0) { // 1
  x = 1 // 2
} else {
  ; // 3
}
y = 1; // 4
while (! (x <= 0)) { // 5
  y = y * x; // 6
  x--; // 7
}
```



SKIP

$$\frac{}{\rho^b \rightarrow^b_{\text{skip}} \rho^b}$$

AFFECTATION

$$\frac{\rho^b \vdash e \Downarrow_{\epsilon} v}{\rho^b \rightarrow^b_{x:=e} \rho^b[x \mapsto v]}$$

CONDITION

$$\frac{\rho^b \vdash e \Downarrow_{\epsilon} v \quad v \neq 0}{\rho^b \rightarrow^b_{?e} \rho^b}$$

ERREUR

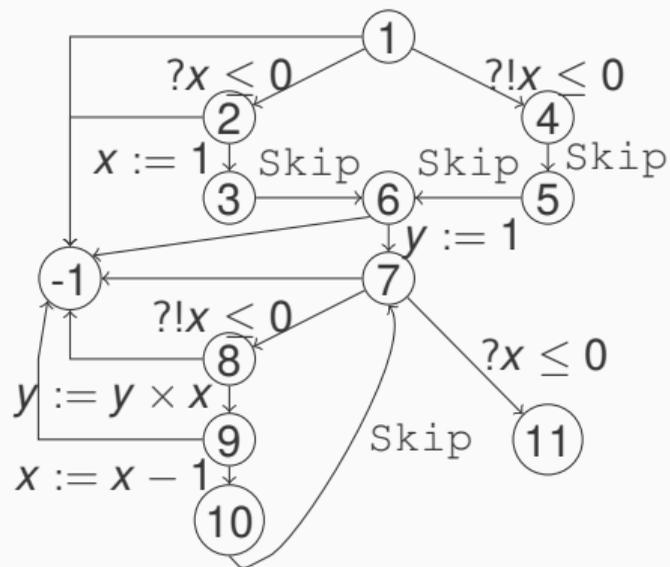
$$\frac{\forall v, \neg \rho^b \vdash e \Downarrow_{\epsilon} v}{\rho^b \rightarrow^b_{\text{Err}(e)} \rho^b}$$

Définition

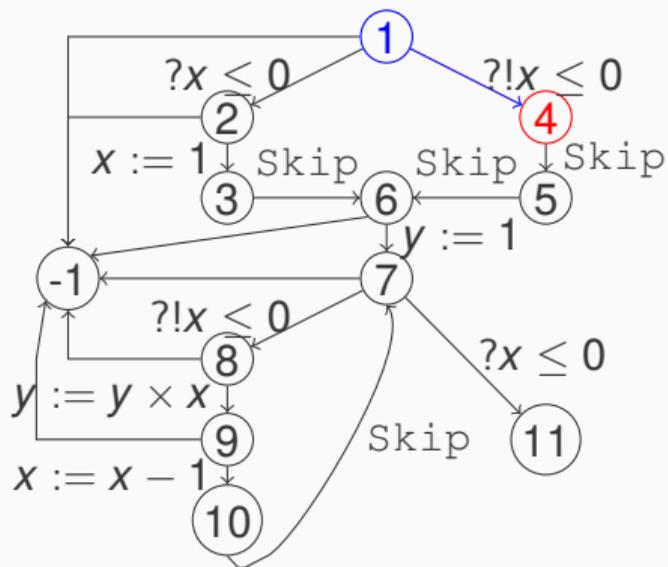
Un **état concret** du programme P est un couple $(s, \rho^b) \in \mathcal{S} \times Env_\epsilon$.

L'ensemble des **transitions** de P est une relation sur les états, \Rightarrow^b_P défini par

$$(s_1, \rho_1) \Rightarrow^b_P (s_2, \rho_2) \Leftrightarrow (s_1, s_2) \in \mathcal{T} \wedge \rho_1 \rightarrow^b_{\epsilon(s_1, s_2)} \rho_2$$



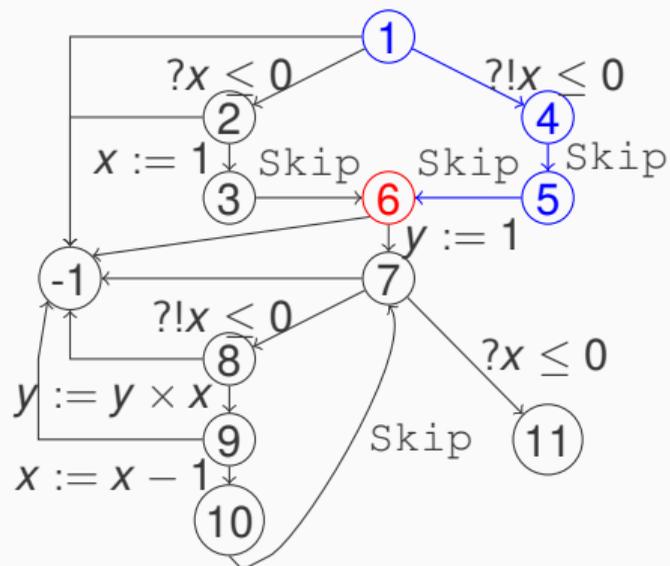
$(s_1, \{x \mapsto 1\})$



$(s_1, \{x \mapsto 1\})$

$(s_4, \{x \mapsto 1\})$

$(s_5, \{x \mapsto 1\})$

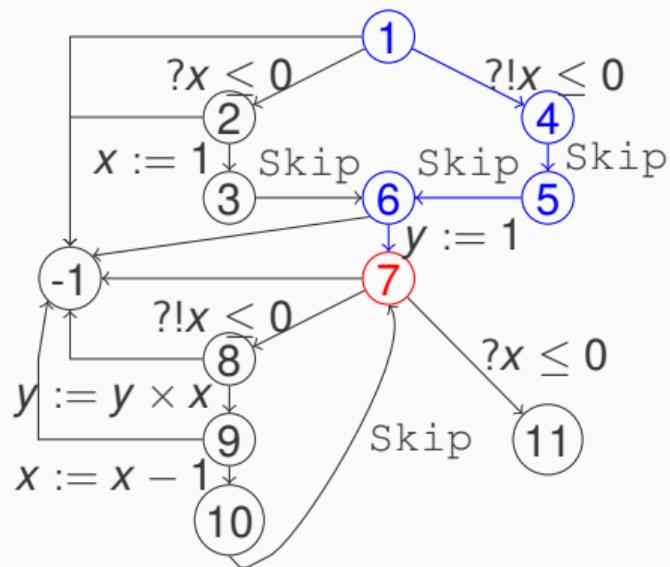


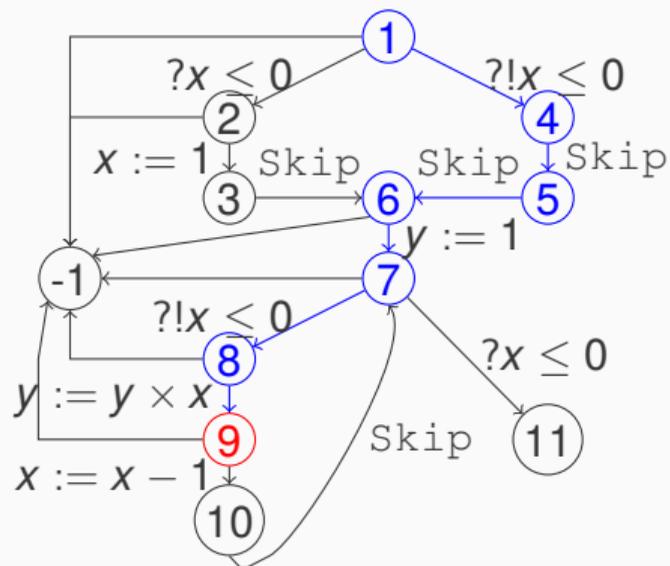
$(s_1, \{x \mapsto 1\})$

$(s_4, \{x \mapsto 1\})$

$(s_5, \{x \mapsto 1\})$

$(s_6, \{x \mapsto 1\})$





$(s_1, \{x \mapsto 1\})$

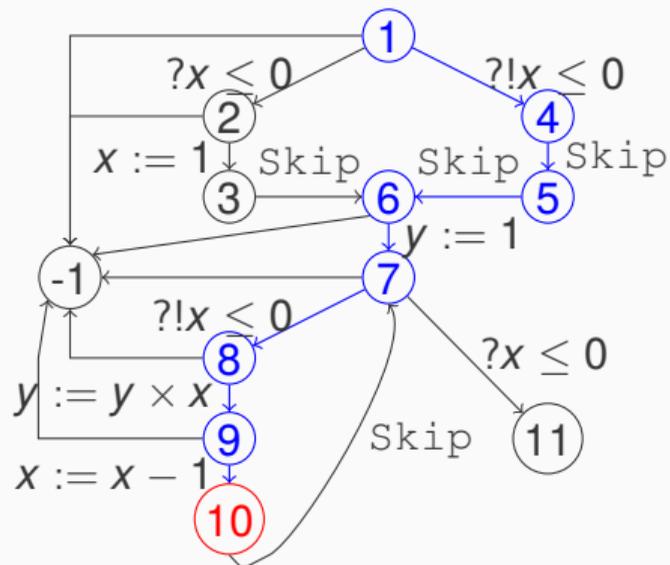
$(s_4, \{x \mapsto 1\})$

$(s_5, \{x \mapsto 1\})$

$(s_6, \{x \mapsto 1\})$

$(s_7, \{x \mapsto 1, y \mapsto 1\})$

$(s_8, \{x \mapsto 1, y \mapsto 1\})$



$(s_1, \{x \mapsto 1\})$

$(s_4, \{x \mapsto 1\})$

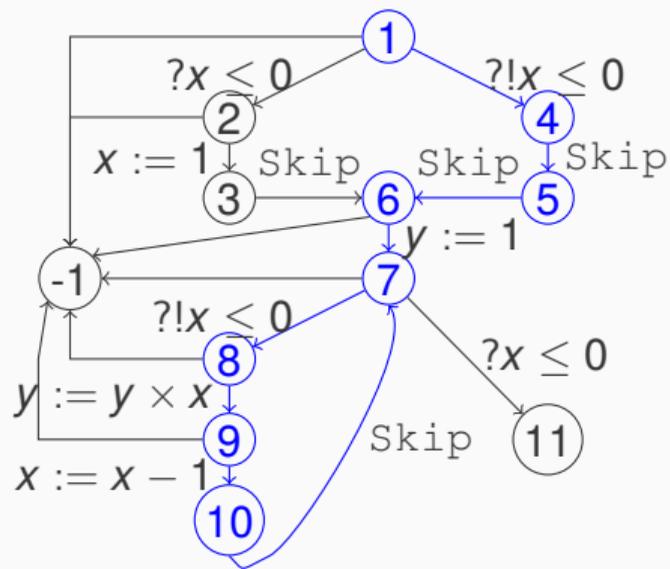
$(s_5, \{x \mapsto 1\})$

$(s_6, \{x \mapsto 1\})$

$(s_7, \{x \mapsto 1, y \mapsto 1\})$

$(s_8, \{x \mapsto 1, y \mapsto 1\})$

$(s_9, \{x \mapsto 1, y \mapsto 1\})$



$(s_1, \{x \mapsto 1\})$

$(s_4, \{x \mapsto 1\})$

$(s_5, \{x \mapsto 1\})$

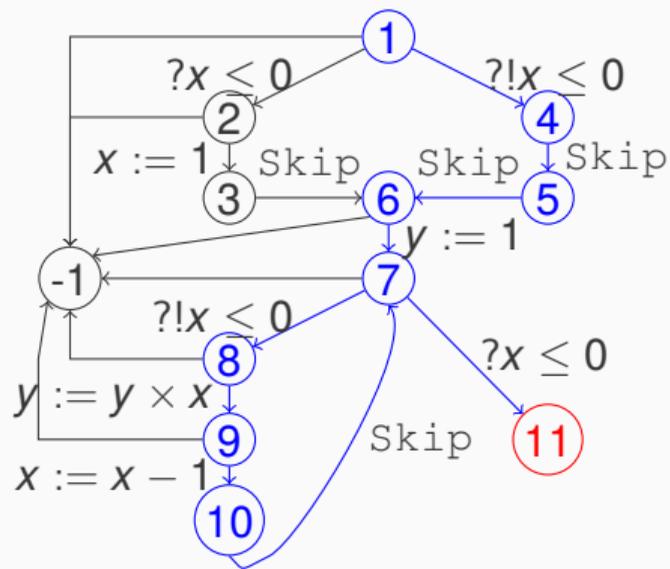
$(s_6, \{x \mapsto 1\})$

$(s_7, \{x \mapsto 1, y \mapsto 1\})$

$(s_8, \{x \mapsto 1, y \mapsto 1\})$

$(s_9, \{x \mapsto 1, y \mapsto 1\})$

$(s_{10}, \{x \mapsto 0, y \mapsto 1\})$



$(s_1, \{x \mapsto 1\})$ $(s_4, \{x \mapsto 1\})$

$(s_5, \{x \mapsto 1\})$ $(s_6, \{x \mapsto 1\})$

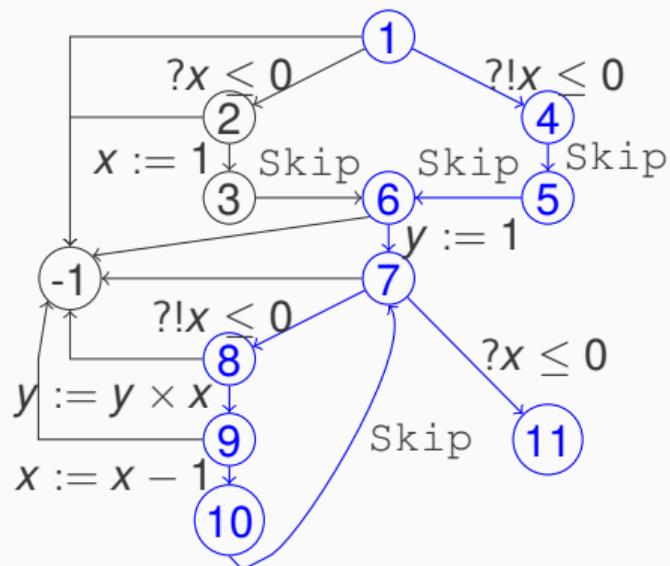
$(s_7, \{x \mapsto 1, y \mapsto 1\})$

$(s_8, \{x \mapsto 1, y \mapsto 1\})$

$(s_9, \{x \mapsto 1, y \mapsto 1\})$

$(s_{10}, \{x \mapsto 0, y \mapsto 1\})$

$(s_7, \{x \mapsto 0, y \mapsto 1\})$



$(s_1, \{x \mapsto 1\})$ $(s_4, \{x \mapsto 1\})$

$(s_5, \{x \mapsto 1\})$ $(s_6, \{x \mapsto 1\})$

$(s_7, \{x \mapsto 1, y \mapsto 1\})$

$(s_8, \{x \mapsto 1, y \mapsto 1\})$

$(s_9, \{x \mapsto 1, y \mapsto 1\})$

$(s_{10}, \{x \mapsto 0, y \mapsto 1\})$

$(s_7, \{x \mapsto 0, y \mapsto 1\})$

$(s_{11}, \{x \mapsto 0, y \mapsto 1\})$

Théorème d'équivalence

Soit $(s_0, \rho_0^b), \dots, (s_n, \rho_n^b)$ une trace finie de P .

- > Si $s_n = \mathcal{F}$, $\rho_0^b \vdash P \Downarrow \rho_n^b$
- > Si $s_n = \epsilon$, $\forall \rho^b, \neg \rho_0^b \vdash P \Downarrow \rho^b$

Réciproquement :

- > Si $\rho_0^b \vdash P \Downarrow \rho_n^b$, alors il existe une trace $(s_0, \rho^b), \dots, (s_n, \rho'^b)$
- > Sinon (i.e. pas de dérivation)
 - > soit il existe une trace $(s_0, \rho^b), \dots, (\epsilon, \rho'^b)$,
 - > soit il existe une trace infinie $(s_0, \rho^b), \dots$

- > L'ensemble des traces est **infini**
- > Existence de traces **infinies**
- > **Approximation** de la sémantique de traces se concentrant sur des propriétés particulières
 - > **Analyse flot de données**
 - > Correction ?
 - > Terminaison ?

Fondations

- > **Gordon D. Plotkin** *A Structural Approach to Operational Semantics*
Tech. Rep. DAIMI FN-19, Computer Science Department, Aarhus University. 1981.
- > **Gilles Kahn** *Natural Semantics*
4th Annual Symposium on Theoretical Aspects of Computer Science.
Springer-Verlag. London. 1987.

Ouvrages de référence :

- > **Glynn Winskel** *The formal semantics of programming language : an introduction*
MIT Press 1993.
- > **Carl A. Gunter** *Semantics of Programming Language : Structures and Techniques*
Foundations of Computing. MIT Press 1992.

Ordre partiel

Un **ordre partiel** \leq sur un ensemble X est une relation binaire

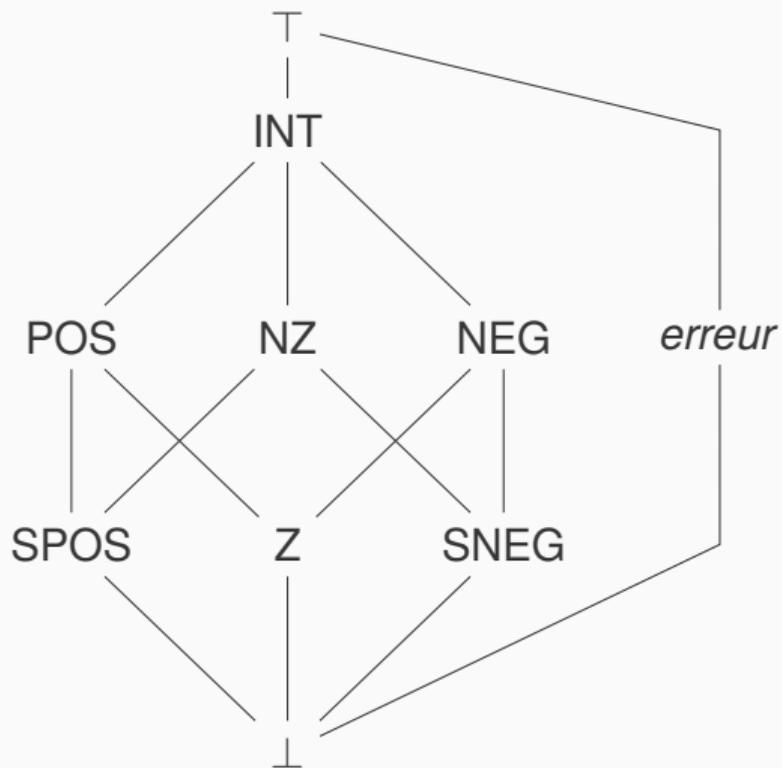
- > **réflexive** : $\forall x \in X, x \leq x$
- > **transitive** : $\forall x, y, z \in X, (x \leq y \wedge y \leq z) \implies x \leq z$
- > **antisymétrique** : $\forall x, y \in X, (x \leq y \wedge y \leq x) \implies x = y$

(E, \sqsubseteq) est un **treillis complet** si et seulement si

- > \sqsubseteq est un ordre partiel
- > $\forall F \subset E, \{x | \forall y \in F, x \sqsubseteq y\}$ possède un plus grand élément, $\sqcap F$
- > $\forall F \subset E, \{x | \forall y \in F, y \sqsubseteq x\}$ possède un plus petit élément, $\sqcup F$
- > $\sqcup E$ est noté \perp , $\sqcap E$ est noté \top

Chaînes

- > Une **chaîne** est une séquence (x_i) ordonnée : $x_1 \sqsubseteq x_2 \sqsubseteq \dots x_n \sqsubseteq \dots$ (chaîne ascendante) ou $x_1 \supseteq x_2 \supseteq \dots x_n \supseteq \dots$ (chaîne descendante)
- > Un treillis complet vérifie la **condition de chaîne ascendante (descendante)** si et seulement si toute chaîne ascendante (descendante) est **stationnaire** :
 $\exists N, \forall i, j \geq N, x_i = x_j$



Théorème du point fixe de Tarski (1953)

Toute fonction **monotone** f sur un treillis complet (E, \sqsubseteq) admet un **plus petit point fixe**, $\text{lfp}(f) = f(\text{lfp}(f))$. De plus :

$$\text{lfp}(f) = \sqcap \{x \mid f(x) \sqsubseteq x\}$$

Itération de Kleene

$$x_0 = \perp$$

$$x_{n+1} = f(x_n)$$

- > (x_i) est croissante
- > Si elle converge, elle atteint $\text{lfp}(f)$.
- > C'est le cas si E vérifie la condition de chaîne ascendante.

Vue informelle

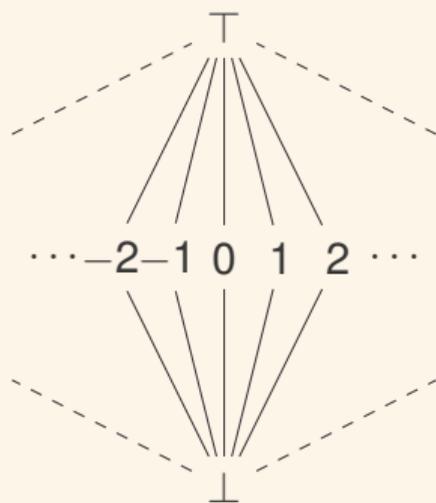
- > on propage des informations le long du graphe de contrôle (**fonction de transition**)
- > lorsqu'un sommet admet plusieurs prédécesseurs (**sommet de jonction**), on considère la réunion des informations propagées
- > on termine lorsque la propagation des informations ne fait plus "augmenter" le résultat

Questions

- > Ce procédé termine-t-il toujours ?
- > Ce procédé est-il correct ?

Treillis de base

On part du treillis \mathcal{C} suivant :

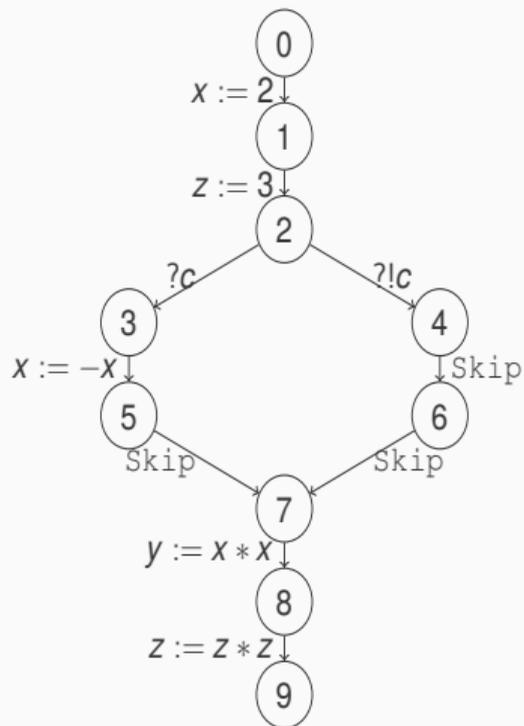


Cadre d'analyse

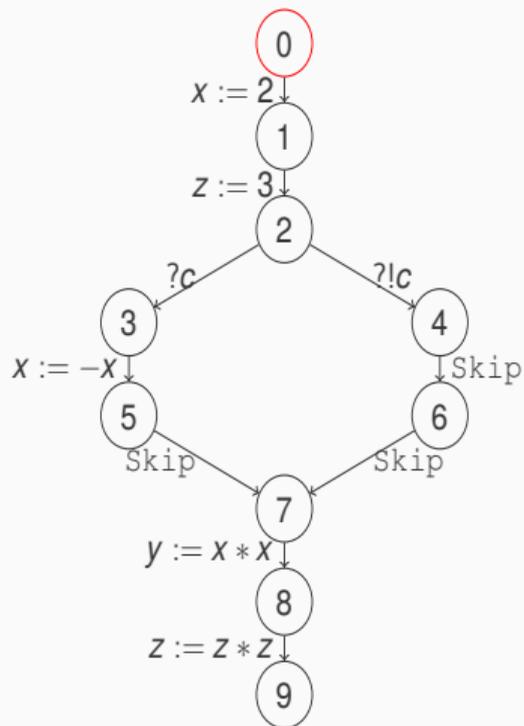
- > $E = Var \rightarrow \mathcal{C}$
- > $i = \lambda x. \top$
- > $f_{x:=e}(\sigma^\sharp) = \sigma^\sharp[x \leftarrow \llbracket e \rrbracket_\sharp \sigma^\sharp]$
- > $f_{\text{skip}}(\sigma^\sharp) = \sigma^\sharp$
- > $f_{?e}(\sigma^\sharp) = \begin{cases} \perp & \text{si } \llbracket e \rrbracket_\sharp \sigma^\sharp = 0 \\ \sigma^\sharp & \text{sinon} \end{cases}$

Propriétés

- > **Correction** : Si à la fin de l'analyse, on a $\sigma^\sharp_i(x) = v$, toute trace passant par i associe v à x .
- > **Terminaison** : L'analyse termine toujours

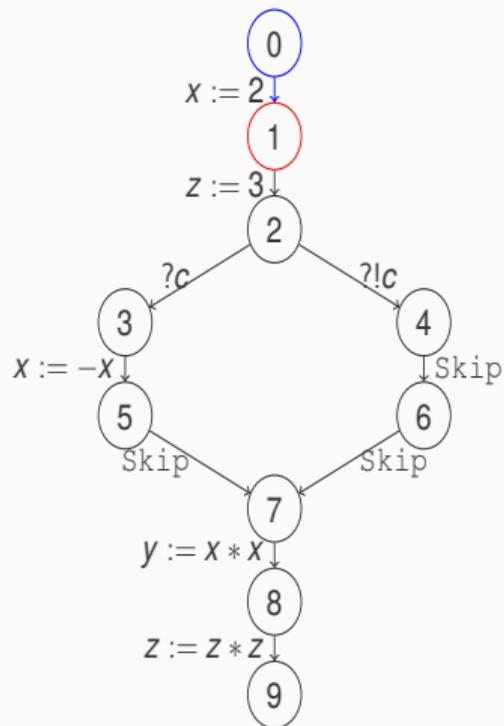


nœud	c	x	y	z
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				



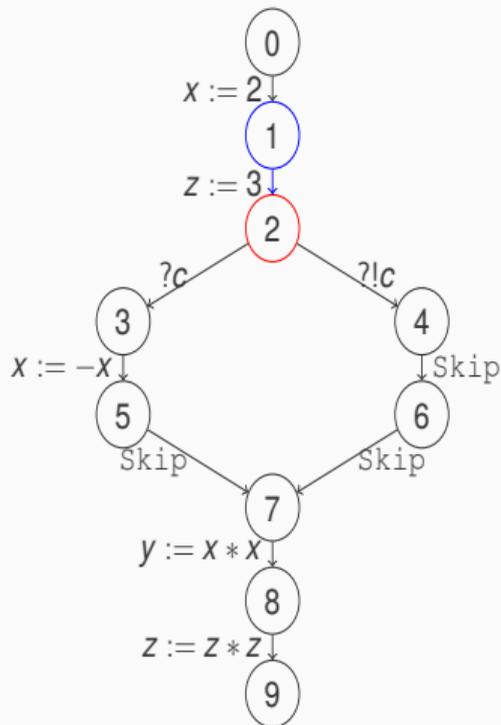
T = "pas d'information"

nœud	c	x	y	z
0	T	T	T	T
1				
2				
3				
4				
5				
6				
7				
8				
9				



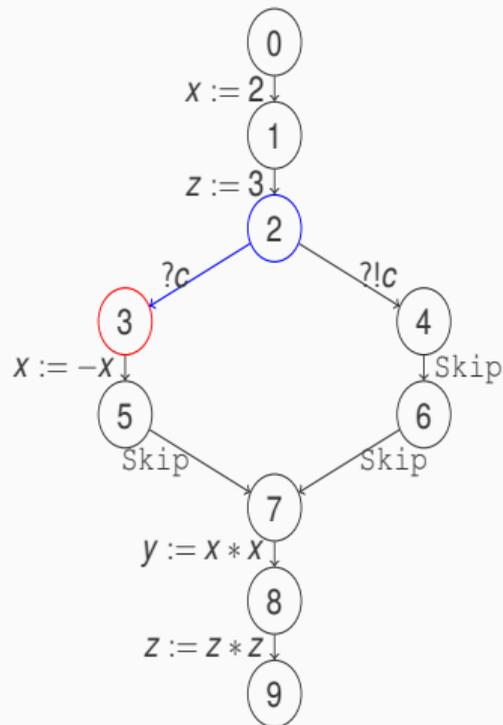
\top = "pas d'information"

nœud	c	x	y	z
0	\top	\top	\top	\top
1	\top	2	\top	\top
2				
3				
4				
5				
6				
7				
8				
9				



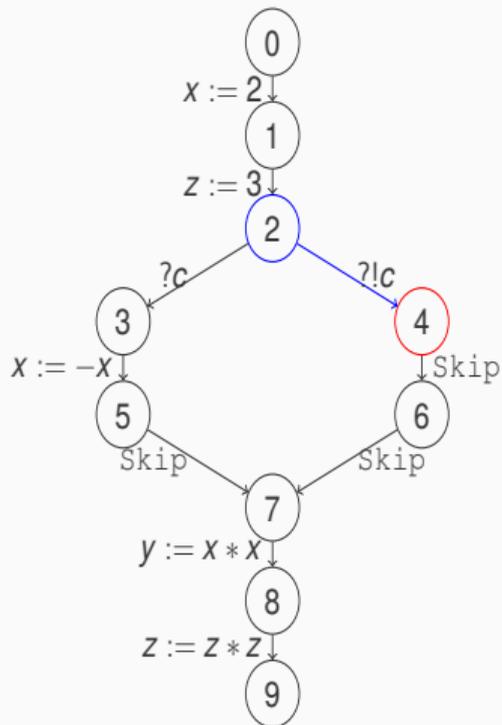
T = "pas d'information"

noeud	c	x	y	z
0	T	T	T	T
1	T	2	T	T
2	T	2	T	3
3				
4				
5				
6				
7				
8				
9				



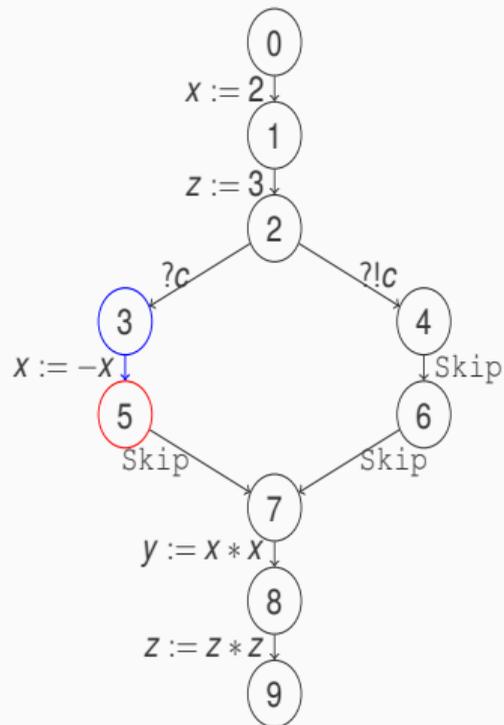
\top = "pas d'information"

noeud	c	x	y	z
0	\top	\top	\top	\top
1	\top	2	\top	\top
2	\top	2	\top	3
3	\top	2	\top	3
4				
5				
6				
7				
8				
9				



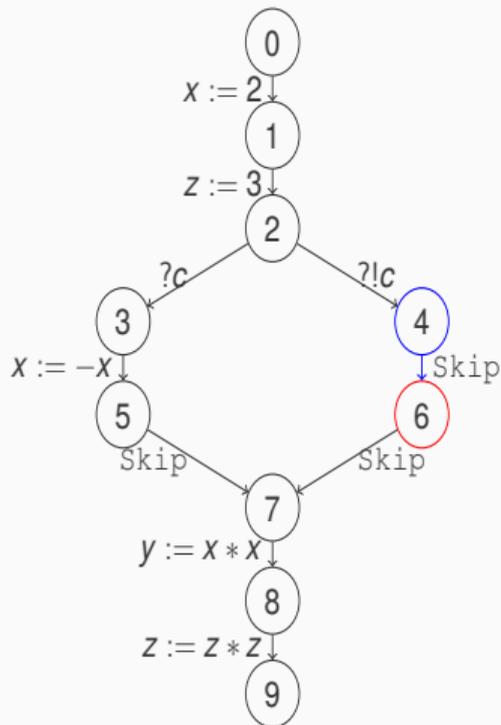
T = "pas d'information"

noeud	c	x	y	z
0	T	T	T	T
1	T	2	T	T
2	T	2	T	3
3	T	2	T	3
4	0	2	T	3
5				
6				
7				
8				
9				



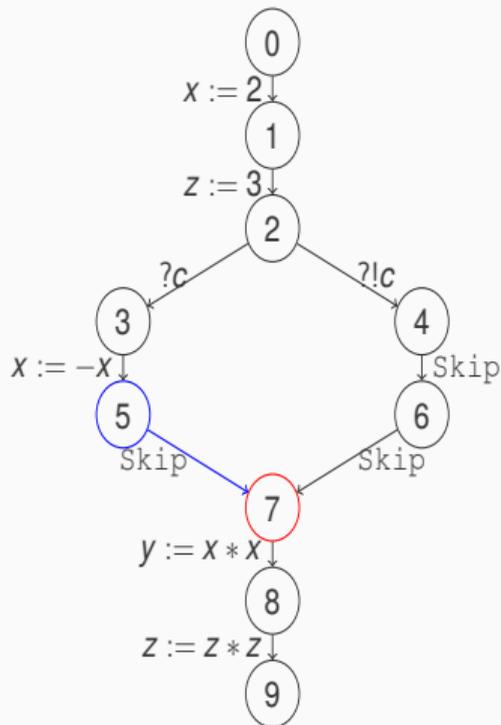
\top = "pas d'information"

nœud	c	x	y	z
0	\top	\top	\top	\top
1	\top	2	\top	\top
2	\top	2	\top	3
3	\top	2	\top	3
4	0	2	\top	3
5	\top	-2	\top	3
6				
7				
8				
9				



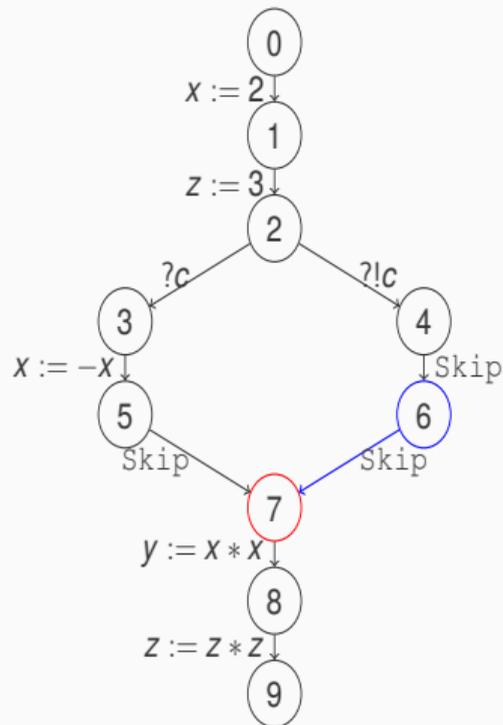
\top = "pas d'information"

nœud	c	x	y	z
0	\top	\top	\top	\top
1	\top	2	\top	\top
2	\top	2	\top	3
3	\top	2	\top	3
4	0	2	\top	3
5	\top	-2	\top	3
6	0	2	\top	3
7				
8				
9				



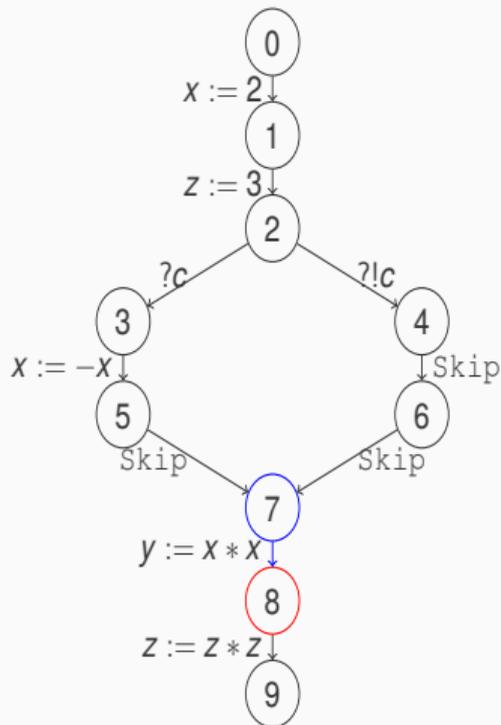
\top = "pas d'information"

noeud	c	x	y	z
0	\top	\top	\top	\top
1	\top	2	\top	\top
2	\top	2	\top	3
3	\top	2	\top	3
4	0	2	\top	3
5	\top	-2	\top	3
6	0	2	\top	3
7	\top	-2	\top	3
8				
9				



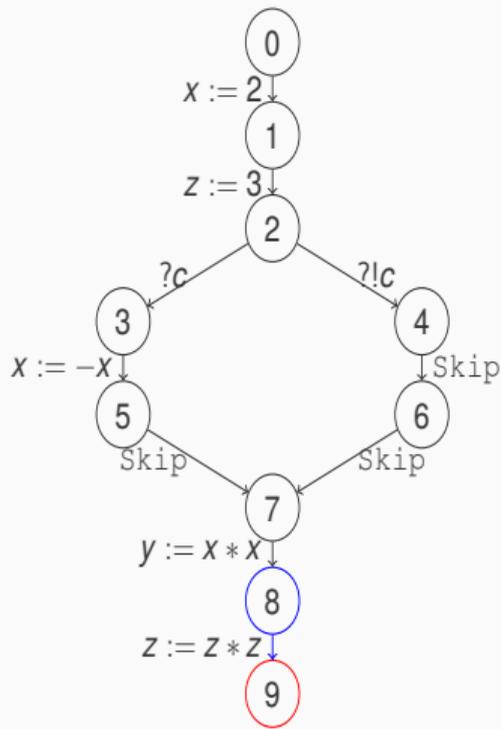
\top = "pas d'information"

noeud	c	x	y	z
0	\top	\top	\top	\top
1	\top	2	\top	\top
2	\top	2	\top	3
3	\top	2	\top	3
4	0	2	\top	3
5	\top	-2	\top	3
6	0	2	\top	3
7	\top	\top	\top	3
8				
9				



\top = "pas d'information"

noeud	c	x	y	z
0	\top	\top	\top	\top
1	\top	2	\top	\top
2	\top	2	\top	3
3	\top	2	\top	3
4	0	2	\top	3
5	\top	-2	\top	3
6	0	2	\top	3
7	\top	\top	\top	3
8	\top	\top	\top	3
9				

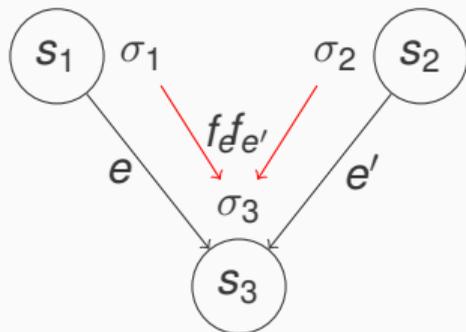


\top = "pas d'information"

nœud	c	x	y	z
0	\top	\top	\top	\top
1	\top	2	\top	\top
2	\top	2	\top	3
3	\top	2	\top	3
4	0	2	\top	3
5	\top	-2	\top	3
6	0	2	\top	3
7	\top	\top	\top	3
8	\top	\top	\top	3
9	\top	\top	\top	9

Lorsqu'un sommet a plusieurs prédécesseurs (ou successeurs pour une analyse arrière), il faut intégrer les informations qui en proviennent :

- > **Sous-approximation** : $\sigma(s) = \prod_{s' \in \text{Pred}(s)} \{f_{e(s',s)}(\sigma(s'))\}$ (on prend l'information qui est vraie sur **tous** les chemins).
- > **Sur-approximation** : $\sigma(s) = \sqcup_{s' \in \text{Pred}(s)} \{f_{e(s',s)}(\sigma(s'))\}$ (on prend l'information qui est vraie sur **au moins un** chemin).



Définition

À chaque point de programme s d'un programme P on associe l'ensemble :

$$\mathcal{C}_P(s) = \left\{ \rho^b \in Env^b \mid \begin{array}{l} \exists (s_0, \rho_0), (s_1, \rho_1), \dots, (s_i, \rho_i), \dots \in \mathbb{T}(P) \\ \exists k \ s = s_k \text{ et } \rho^b = \rho_k \end{array} \right\}$$

$$\mathcal{C}_P \in Env_P^b = \{f^b : \mathcal{S} \rightarrow \wp(Env^b)\}$$

Note

Env_P^b peut être muni d'un ordre partiel :

$$\forall f_1^b, f_2^b \in Env_P^b \quad f_1^b \subseteq f_2^b \quad \text{ssi} \quad \forall s \in \mathcal{S}, f_1^b(s) \subseteq f_2^b(s)$$

On cherche à calculer la **sémantique collectrice** de P , $Coll_P$, comme une analyse flot de données.

- > On se donne un ensemble d'états initiaux arbitraires $R_{\mathcal{I}}$
- > Les fonctions de transition sont de la forme :

$$\mathcal{P}_e^b = \lambda R. \{ \rho \mid \exists \rho' \in R, \rho' \rightarrow_e^b \rho \}$$

- > On part de $\lambda s \in \mathcal{S}. \begin{cases} R_{\mathcal{I}} & \text{si } s = \mathcal{I} \\ \emptyset & \text{sinon} \end{cases}$ (qui appartient à Env_P^b)

On cherche à calculer la **sémantique collectrice** de P , $Coll_P$, comme une analyse flot de données.

- > On se donne un ensemble d'états initiaux arbitraires $R_{\mathcal{I}}$
- > Les fonctions de transition sont de la forme :

$$\mathcal{P}_e^b = \lambda R. \{\rho \mid \exists \rho' \in R, \rho' \rightarrow_e^b \rho\}$$

- > On part de $\lambda s \in \mathcal{S}. \begin{cases} R_{\mathcal{I}} & \text{si } s = \mathcal{I} \\ \emptyset & \text{sinon} \end{cases}$ (qui appartient à Env_P^b)

Lemme

\mathcal{P}_e^b est une fonction monotone

On cherche à calculer la **sémantique collectrice** de P , $Coll_P$, comme une analyse flot de données.

- > On se donne un ensemble d'états initiaux arbitraires $R_{\mathcal{I}}$
- > Les fonctions de transition sont de la forme :

$$\mathcal{P}_e^b = \lambda R. \{ \rho \mid \exists \rho' \in R, \rho' \rightarrow_e^b \rho \}$$

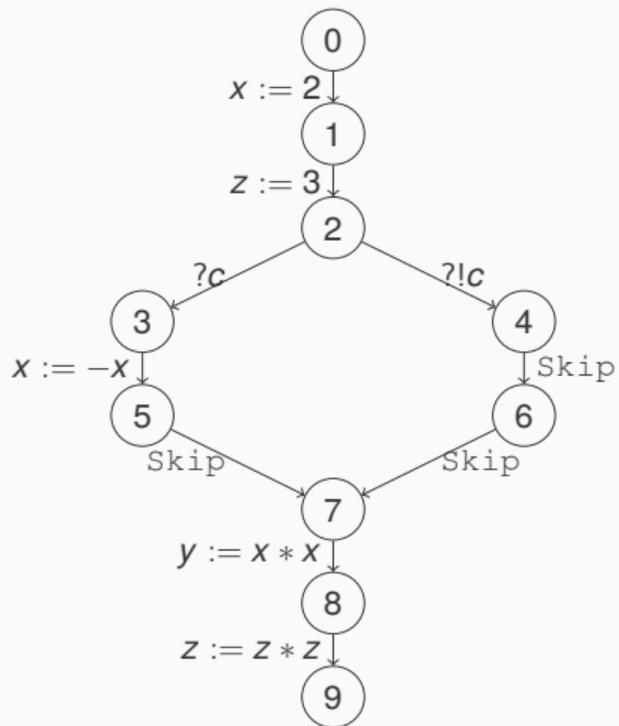
- > On part de $\lambda s \in \mathcal{S}. \begin{cases} R_{\mathcal{I}} & \text{si } s = \mathcal{I} \\ \emptyset & \text{sinon} \end{cases}$ (qui appartient à Env_P^b)

Lemme

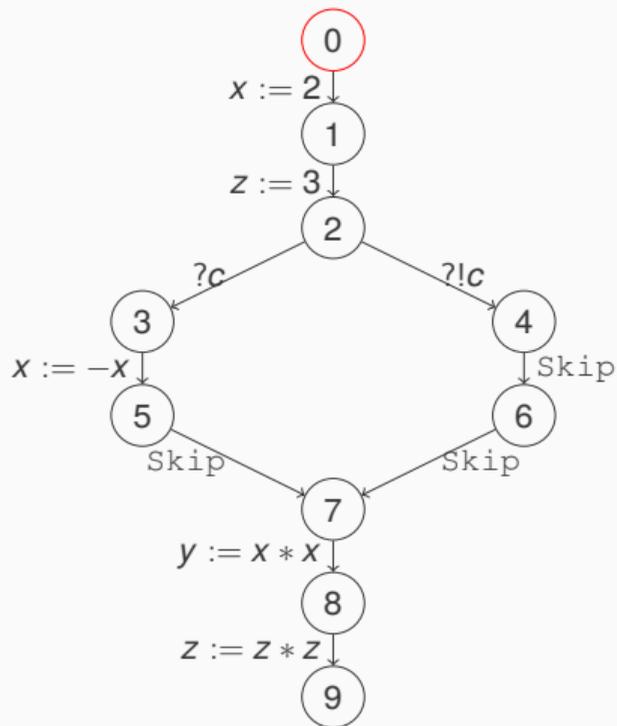
\mathcal{P}_e^b est une fonction monotone

Remarque

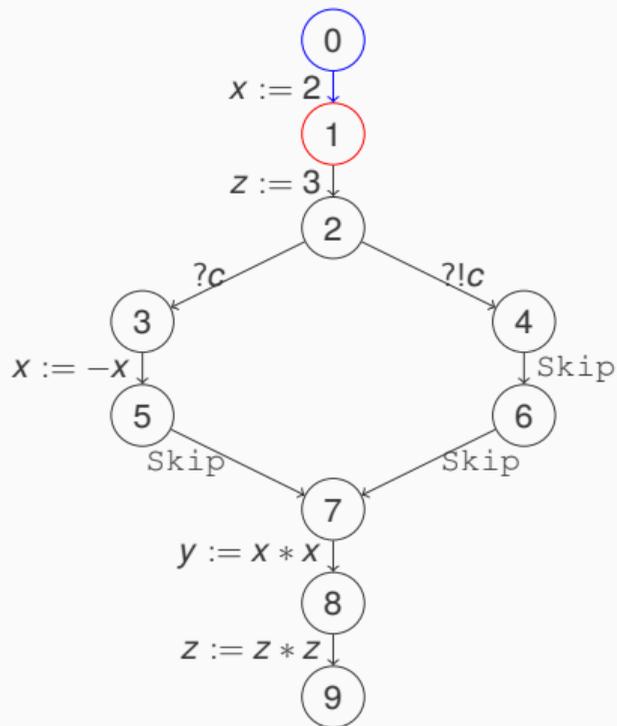
Par contre $\wp(Env_P^b)$ ne vérifie pas la condition de chaîne ascendante.



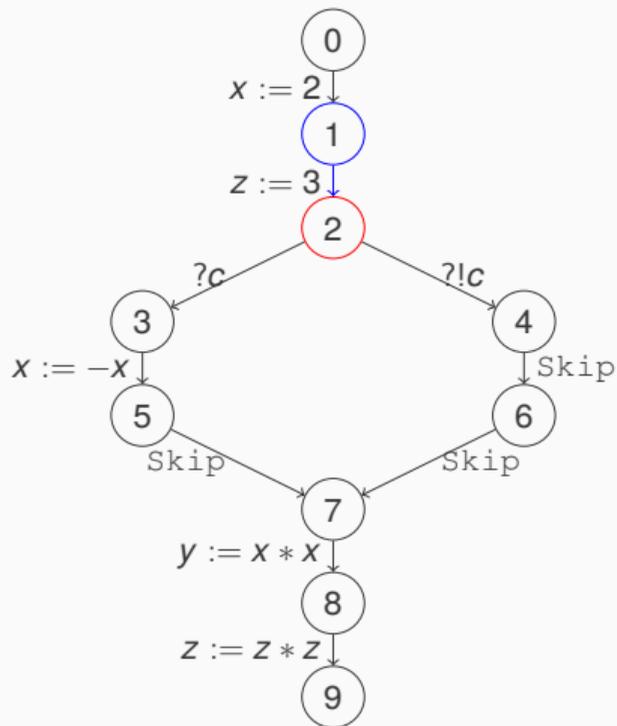
nœuds	environnements
0	
1	
2	
3	
4	
5	
6	



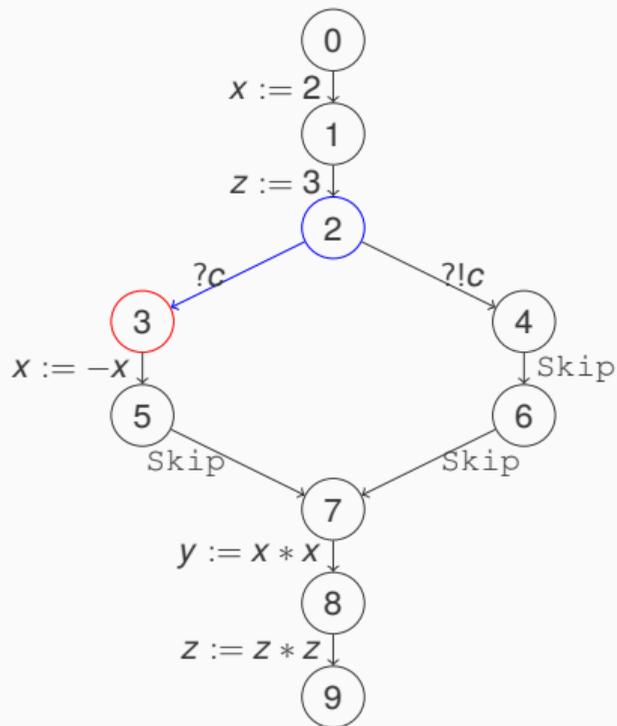
nœuds	environnements
0	$\{\lambda x.n_x \mid n_x \in \mathbb{Z}\}$
1	
2	
3	
4	
5	
6	



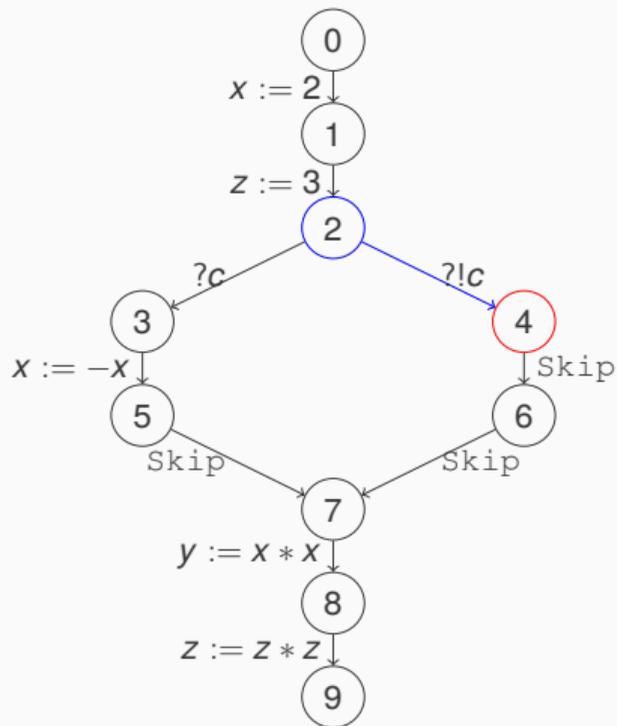
nœuds	environnements
0	$\{\lambda x.n_x \mid n_x \in \mathbb{Z}\}$
1	$\{x \leftarrow 2; y \leftarrow n_y;$ $z \leftarrow n_z; c \leftarrow n_c\}$
2	
3	
4	
5	
6	



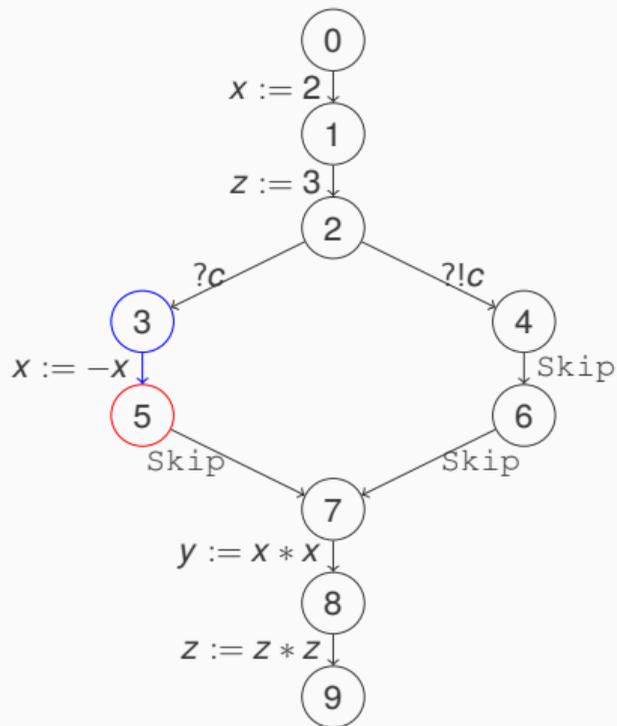
nœuds	environnements
0	$\{\lambda x.n_x \mid n_x \in \mathbb{Z}\}$
1	$\{x \leftarrow 2; y \leftarrow n_y;$ $z \leftarrow n_z; c \leftarrow n_c\}$
2	$\{x \leftarrow 2; y \leftarrow n_y;$ $z \leftarrow 3; c \leftarrow n_c\}$
3	
4	
5	
6	



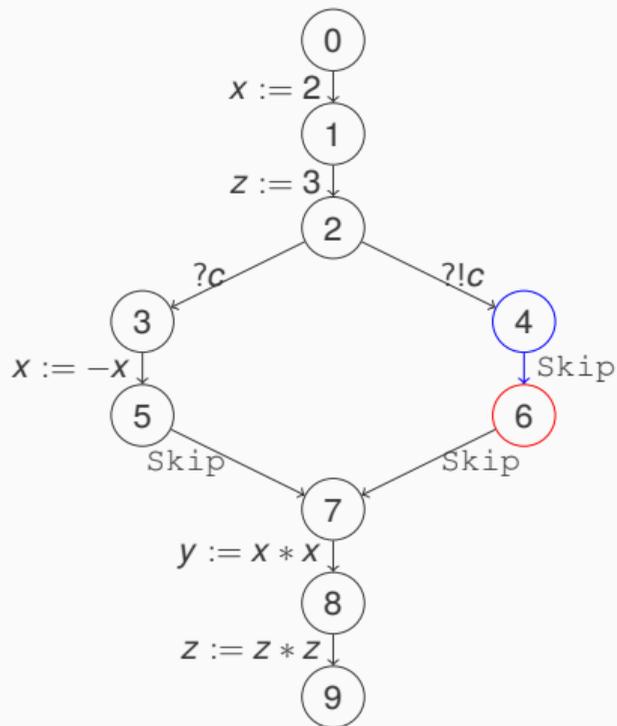
nœuds	environnements
0	$\{\lambda x.n_x \mid n_x \in \mathbb{Z}\}$
1	$\{x \leftarrow 2; y \leftarrow n_y;$ $z \leftarrow n_z; c \leftarrow n_c\}$
2	$\{x \leftarrow 2; y \leftarrow n_y;$ $z \leftarrow 3; c \leftarrow n_c\}$
3	$\{x \leftarrow 2; y \leftarrow n_y;$ $z \leftarrow 3; c \leftarrow n_c\}; n_c \neq 0$
4	
5	
6	



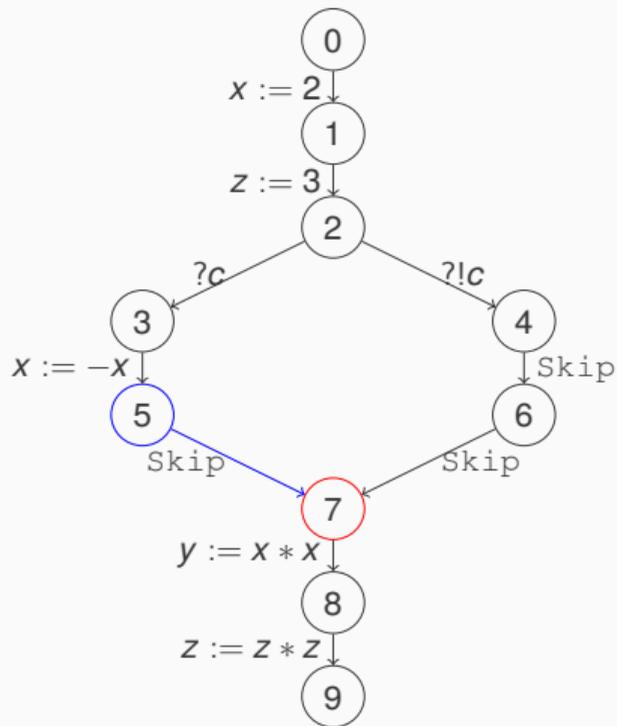
nœuds	environnements
0	$\{\lambda x.n_x \mid n_x \in \mathbb{Z}\}$
1	$\{x \leftarrow 2; y \leftarrow n_y;$ $z \leftarrow n_z; c \leftarrow n_c\}$
2	$\{x \leftarrow 2; y \leftarrow n_y;$ $z \leftarrow 3; c \leftarrow n_c\}$
3	$\{x \leftarrow 2; y \leftarrow n_y;$ $z \leftarrow 3; c \leftarrow n_c\}; n_c \neq 0$
4	$\{x \leftarrow 2; y \leftarrow n_y;$ $z \leftarrow 3; c \leftarrow 0\}$
5	
6	



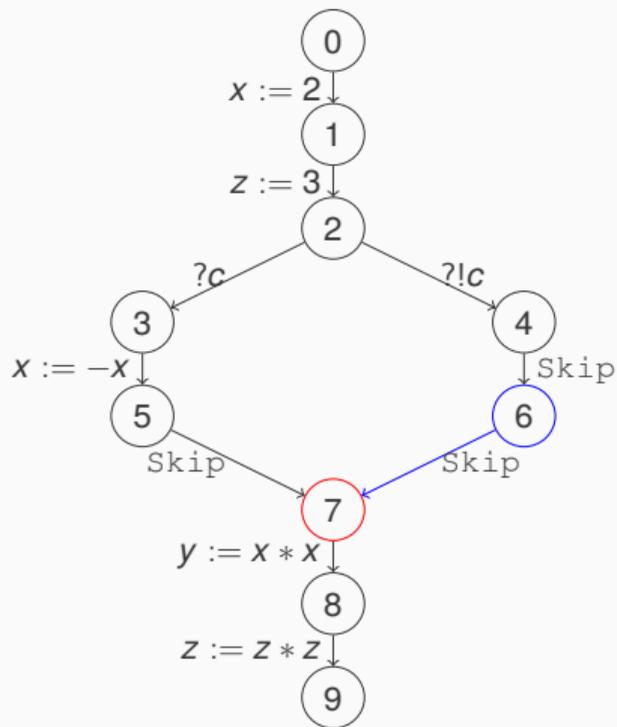
nœuds	environnements
0	$\{\lambda x.n_x \mid n_x \in \mathbb{Z}\}$
1	$\{x \leftarrow 2; y \leftarrow n_y;$ $z \leftarrow n_z; c \leftarrow n_c\}$
2	$\{x \leftarrow 2; y \leftarrow n_y;$ $z \leftarrow 3; c \leftarrow n_c\}$
3	$\{x \leftarrow 2; y \leftarrow n_y;$ $z \leftarrow 3; c \leftarrow n_c\}; n_c \neq 0$
4	$\{x \leftarrow 2; y \leftarrow n_y;$ $z \leftarrow 3; c \leftarrow 0\}$
5	$\{x \leftarrow -2; y \leftarrow n_y;$ $z \leftarrow 3; c \leftarrow n_c\}; n_c \neq 0$
6	



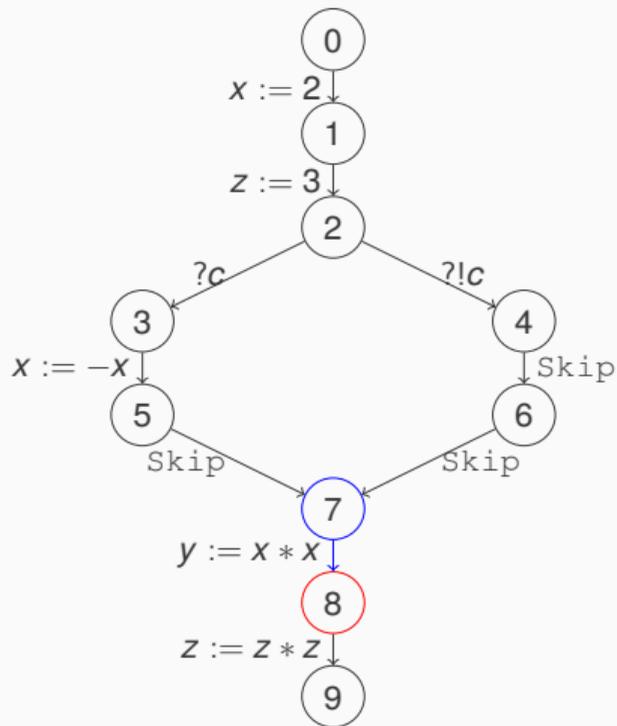
nœuds	environnements
0	$\{\lambda x.n_x \mid n_x \in \mathbb{Z}\}$
1	$\{x \leftarrow 2; y \leftarrow n_y;$ $z \leftarrow n_z; c \leftarrow n_c\}$
2	$\{x \leftarrow 2; y \leftarrow n_y;$ $z \leftarrow 3; c \leftarrow n_c\}$
3	$\{x \leftarrow 2; y \leftarrow n_y;$ $z \leftarrow 3; c \leftarrow n_c\}; n_c \neq 0$
4	$\{x \leftarrow 2; y \leftarrow n_y;$ $z \leftarrow 3; c \leftarrow 0\}$
5	$\{x \leftarrow -2; y \leftarrow n_y;$ $z \leftarrow 3; c \leftarrow n_c\}; n_c \neq 0$
6	$\{x \leftarrow 2; y \leftarrow n_y;$ $z \leftarrow 3; c \leftarrow 0\}$



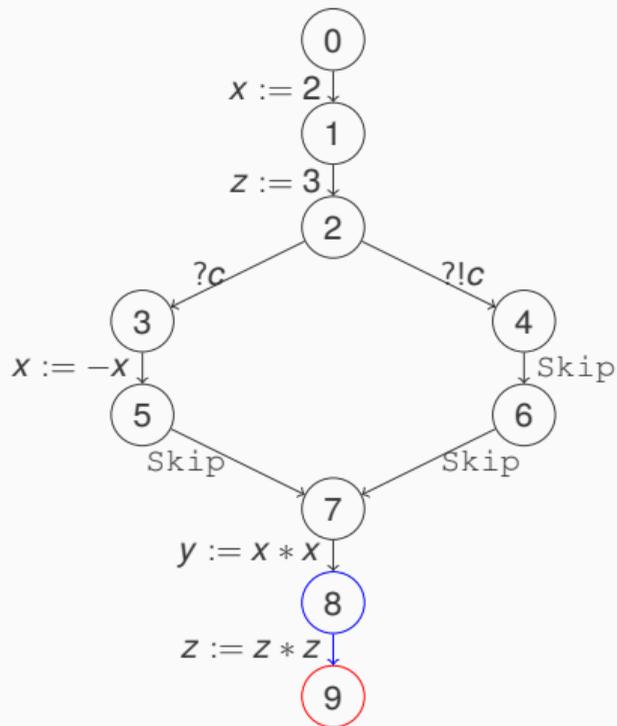
nœuds	environnements
7	$\{x \leftarrow -2; y \leftarrow n_y;$ $z \leftarrow 3; c \leftarrow n_c\}; n_c \neq 0$
8	
9	



nœuds	environnements
7	$\{x \leftarrow -2; y \leftarrow n_y;$ $z \leftarrow 3; c \leftarrow n_c\}; n_c \neq 0$ $\cup \{x \leftarrow 2; y \leftarrow n_y$ $z \leftarrow 3; c \leftarrow 0\}$
8	
9	



nœuds	environnements
7	$\{x \leftarrow -2; y \leftarrow n_y;$ $z \leftarrow 3; c \leftarrow n_c\}; n_c \neq 0$ $\cup \{x \leftarrow 2; y \leftarrow n_y$ $z \leftarrow 3; c \leftarrow 0\}$
8	$\{x \leftarrow -2; y \leftarrow 4;$ $z \leftarrow 3; c \leftarrow n_c\}; n_c \neq 0$ $\cup \{x \leftarrow 2; y \leftarrow 4$ $z \leftarrow 3; c \leftarrow 0\}$
9	



nœuds	environnements
7	$\{x \leftarrow -2; y \leftarrow n_y;$ $z \leftarrow 3; c \leftarrow n_c\}; n_c \neq 0$ $\cup \{x \leftarrow 2; y \leftarrow n_y$ $z \leftarrow 3; c \leftarrow 0\}$
8	$\{x \leftarrow -2; y \leftarrow 4;$ $z \leftarrow 3; c \leftarrow n_c\}; n_c \neq 0$ $\cup \{x \leftarrow 2; y \leftarrow 4$ $z \leftarrow 3; c \leftarrow 0\}$
9	$\{x \leftarrow -2; y \leftarrow 4;$ $z \leftarrow 9; c \leftarrow n_c\}; n_c \neq 0$ $\cup \{x \leftarrow 2; y \leftarrow 4$ $z \leftarrow 9; c \leftarrow 0\}$

Théorème

Tout état associé à un point de programme s dans une trace de P est dans $Coll_P(s)$:

$$C_P = Coll_P$$

Approximation

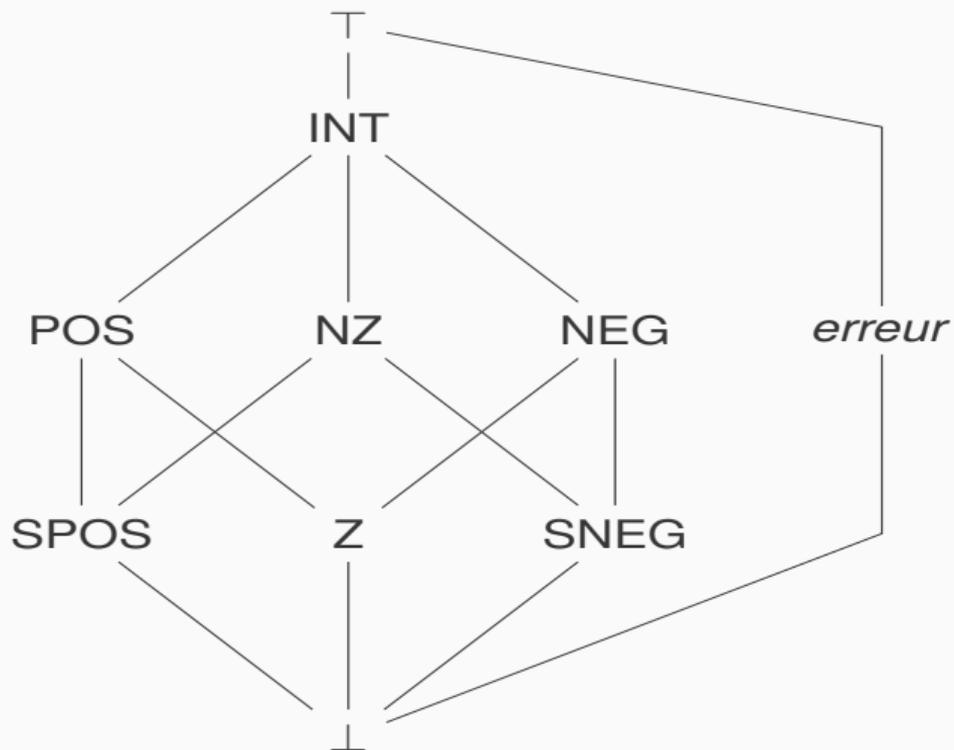
On perd la relation entre les états : Si s_1 et s_2 sont deux états reliés par un arc, avec $\rho_1 \in Coll_P(s_1)$, $\rho_2 \in Coll_P(s_2)$, on n'a pas forcément de trace contenant (s_1, ρ_1) , (s_2, ρ_2)

Terminaison

Calculer $Coll_P$ par itération ne termine pas toujours. Nous voulons des abstractions \mathcal{A}

- > **correctes** : telles que $Coll_P(s) \subset \mathcal{A}(s)$
- > qui **terminent** toujours, quel que soit P

- > définir un **domaine abstrait**
- > définir les fonctions de correspondance entre domaines concret et abstrait
 - > **abstraction** α
 - > **concrétisation** γ
- > preuve que (α, γ) est une **connexion de Galois**
- > déduction "automatique" des autres éléments abstraits
 - > environnements abstraits
 - > opérateurs abstraits
 - > fonction de transition abstraite
- > analyse flot de données avec ces opérations abstraites
- > garantie "automatique" de **correction** et de **terminaison**



On va considérer des environnements **abstraits**, $\mathbb{E}^\# : \text{Var} \rightarrow \text{Signes}$. Chacun représente un ensemble d'environnements concrets :

$$\begin{aligned}\gamma(\top) &= \mathbb{Z} \cup \{\text{erreur}\} \\ \gamma(\text{INT}) &= \mathbb{Z} \\ \gamma(\text{POS}) &= \{v \in \mathbb{Z} \mid v \geq 0\} \\ \gamma(\text{NEG}) &= \{v \in \mathbb{Z} \mid v \leq 0\} \\ \gamma(\text{SPOS}) &= \{v \in \mathbb{Z} \mid v > 0\} \\ \gamma(\text{SNEG}) &= \{v \in \mathbb{Z} \mid v < 0\} \\ \gamma(\text{Z}) &= \{0\} \\ \gamma(\text{NZ}) &= \{v \in \mathbb{Z} \mid v \neq 0\} \\ \gamma(\perp) &= \emptyset \\ \gamma(\rho^\#) &= \{\rho^b \mid \forall x, \rho^b(x) \in \gamma(\rho^\#(x))\}\end{aligned}$$

Les fonctions de transitions sont assez similaires aux opérations concrètes, sauf qu'on évalue les expressions dans Signes

$+\#$	T	\perp	INT	<i>erreur</i>	POS	NEG	NZ	SPOS	Z	SNEG
T										
\perp										
INT										
<i>erreur</i>										
POS										
NEG										
NZ										
SPOS										
Z										
SNEG										

Les fonctions de transitions sont assez similaires aux opérations concrètes, sauf qu'on évalue les expressions dans Signes

$+\#$	T	\perp	INT	<i>erreur</i>	POS	NEG	NZ	SPOS	Z	SNEG
T	T	\perp	T	<i>erreur</i>	T	T	T	T	T	T
\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp
INT	T	\perp	INT	<i>erreur</i>	INT	INT	INT	INT	INT	INT
<i>erreur</i>	<i>erreur</i>	\perp	<i>erreur</i>							
POS	T	\perp	INT	<i>erreur</i>	POS	INT	INT	SPOS	POS	INT
NEG	T	\perp	INT	<i>erreur</i>	INT	NEG	INT	INT	NEG	SNEG
NZ	T	\perp	INT	<i>erreur</i>	INT	INT	INT	INT	INT	INT
SPOS	T	\perp	INT	<i>erreur</i>	SPOS	INT	INT	SPOS	SPOS	INT
Z	T	\perp	INT	<i>erreur</i>	POS	NEG	Int	SPOS	Z	SNEG
SNEG	T	\perp	INT	<i>erreur</i>	INT	SNEG	INT	INT	SNEG	SNEG

SKIP

$$\frac{}{\rho^\sharp \rightarrow_{\text{skip}}^\sharp \rho^\sharp}$$

AFFECTATION

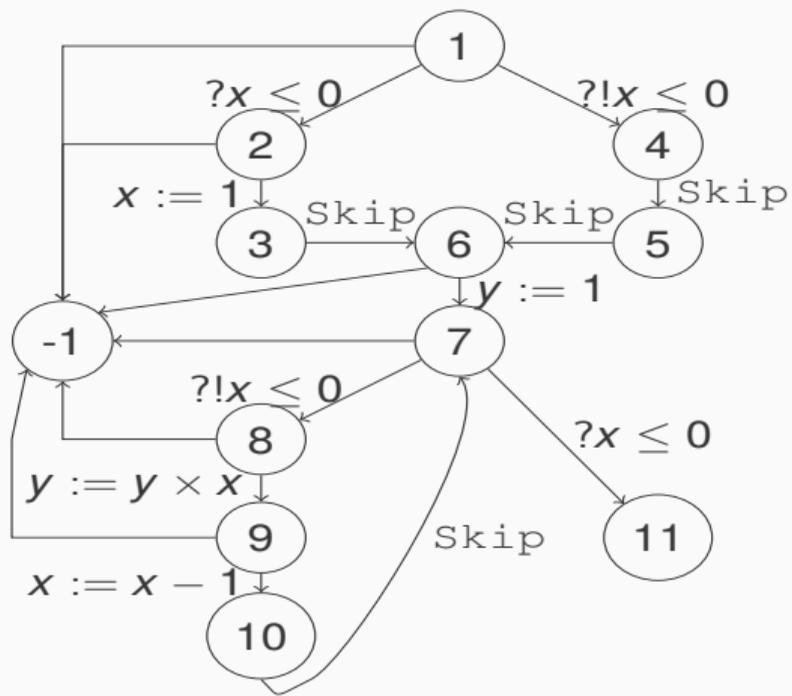
$$\frac{\rho^\sharp \vdash e \Downarrow_\epsilon^\sharp v \quad v \notin \{\perp \mid \text{erreur}\}}{\rho^\sharp \rightarrow_{x:=e}^\sharp \rho^\sharp[x \mapsto v]}$$

CONDITION

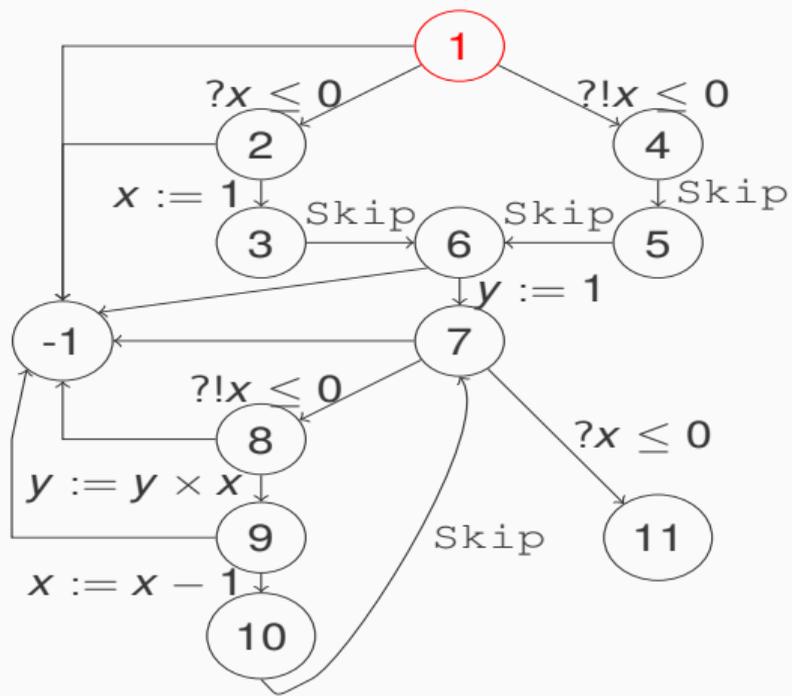
$$\frac{\rho^\sharp \vdash e \Downarrow_\epsilon^\sharp v \quad v \notin \{Z \mid \perp \mid \text{erreur}\}}{\rho^\sharp \rightarrow_{?e}^\sharp [\rho^\sharp]}$$

ERREUR

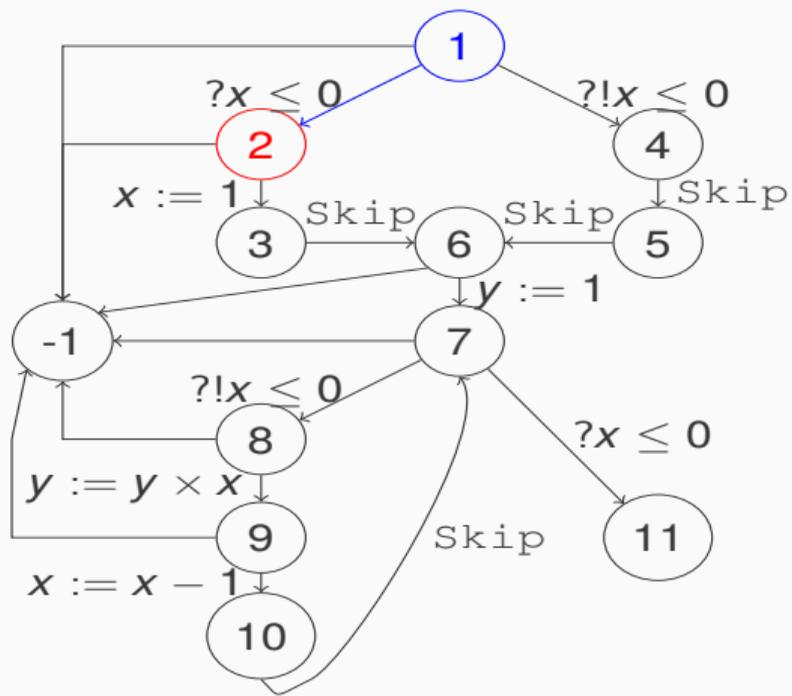
$$\frac{\rho^\sharp \vdash e \Downarrow_\epsilon^\sharp v \quad v \in \{T \mid \text{erreur}\}}{\rho^\sharp \rightarrow_{\text{Err}(e)}^\sharp \rho^\sharp}$$



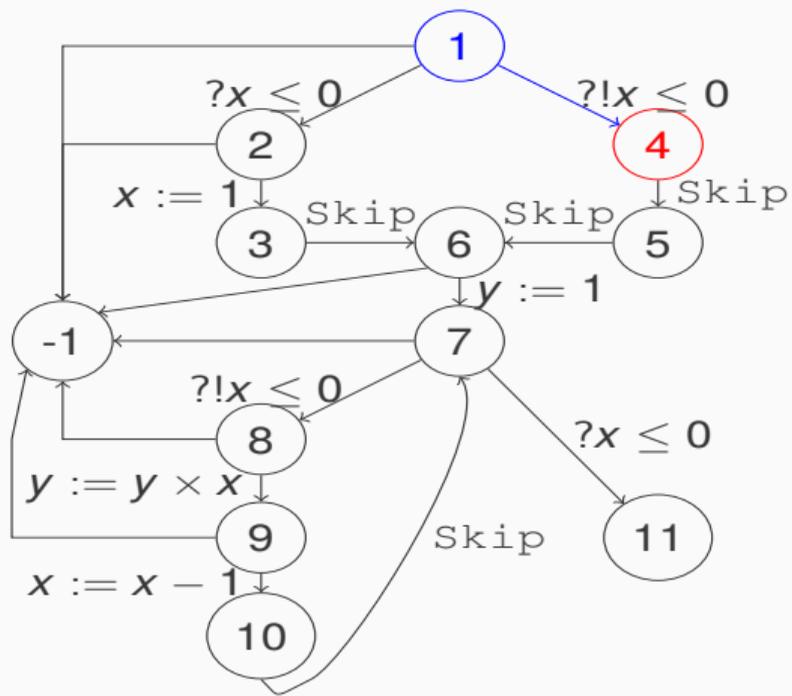
nœud	x	y
S ₁		
S ₂		
S ₃		
S ₄		
S ₅		
S ₆		
S ₇		
S ₈		
S ₉		
S ₁₀		
S ₁₁		
S ₋₁		



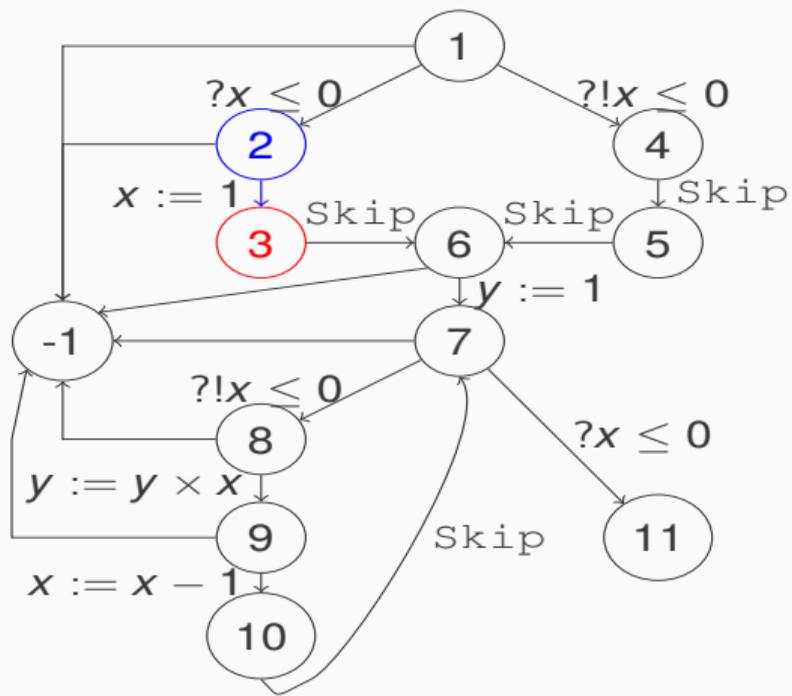
nœud	x	y
S₁	INT	INT
S ₂		
S ₃		
S ₄		
S ₅		
S ₆		
S ₇		
S ₈		
S ₉		
S ₁₀		
S ₁₁		
S ₋₁		



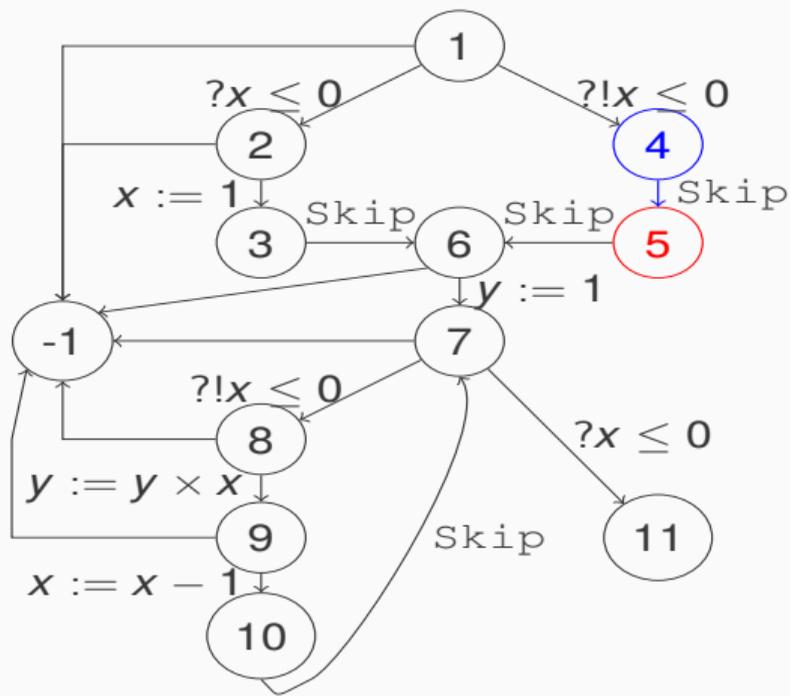
nœud	x	y
s_1	INT	INT
s_2	NEG	INT
s_3		
s_4		
s_5		
s_6		
s_7		
s_8		
s_9		
s_{10}		
s_{11}		
s_{-1}		



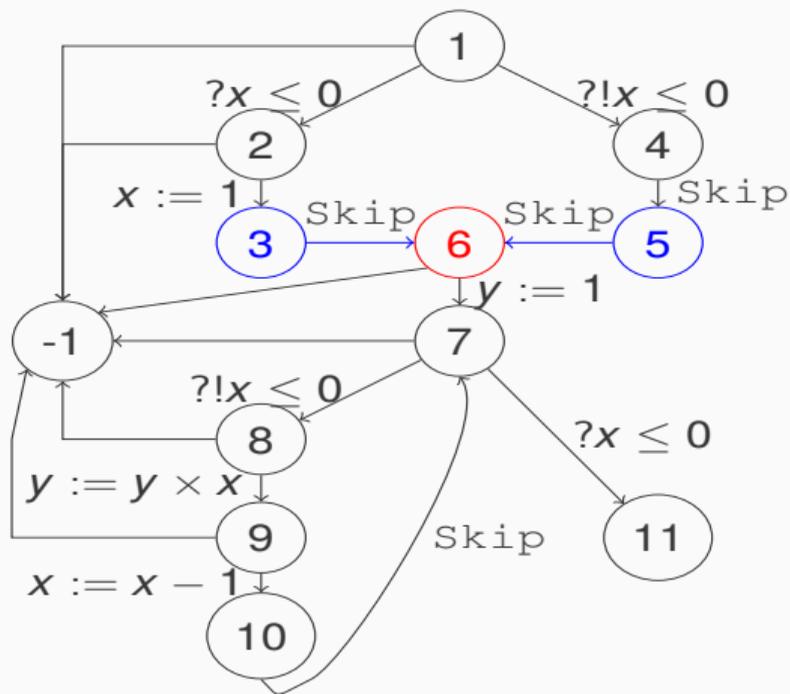
nœud	x	y
s₁	INT	INT
s ₂	NEG	INT
s ₃		
s₄	SPOS	INT
s ₅		
s ₆		
s ₇		
s ₈		
s ₉		
s ₁₀		
s ₁₁		
s ₋₁		



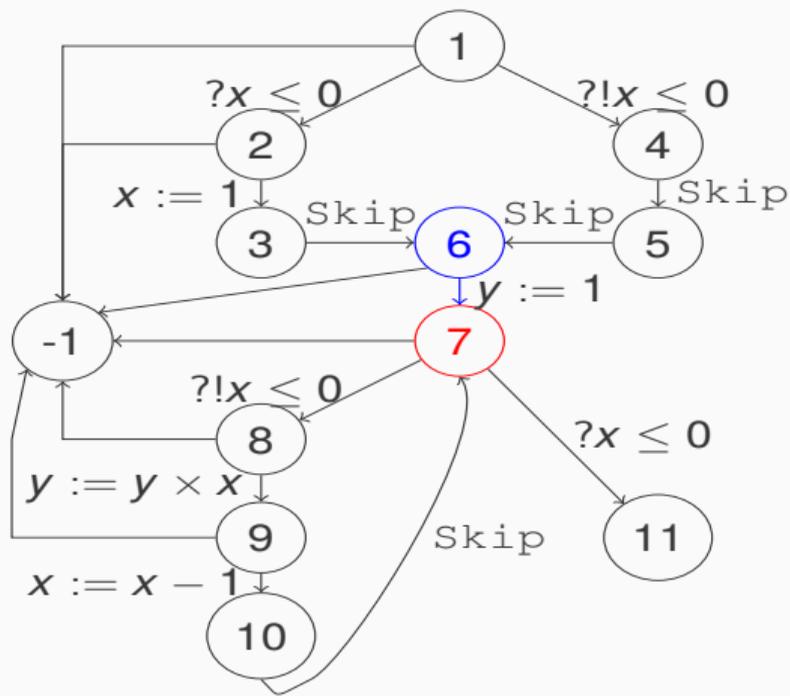
nœud	x	y
s ₁	INT	INT
s ₂	NEG	INT
s ₃	SPOS	INT
s ₄	SPOS	INT
s ₅		
s ₆		
s ₇		
s ₈		
s ₉		
s ₁₀		
s ₁₁		
s ₋₁		



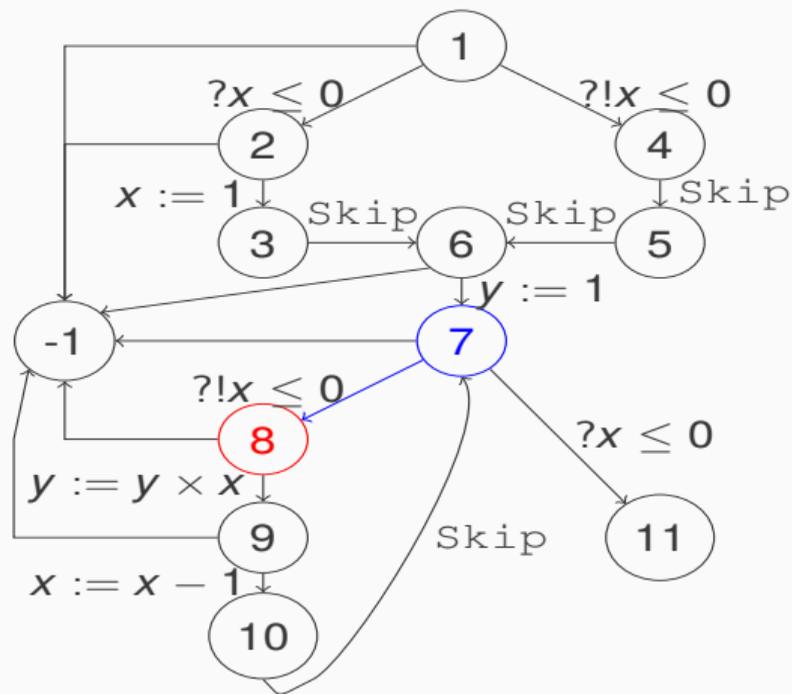
nœud	x	y
s ₁	INT	INT
s ₂	NEG	INT
s ₃	SPOS	INT
s ₄	SPOS	INT
s ₅	SPOS	INT
s ₆		
s ₇		
s ₈		
s ₉		
s ₁₀		
s ₁₁		
s ₋₁		



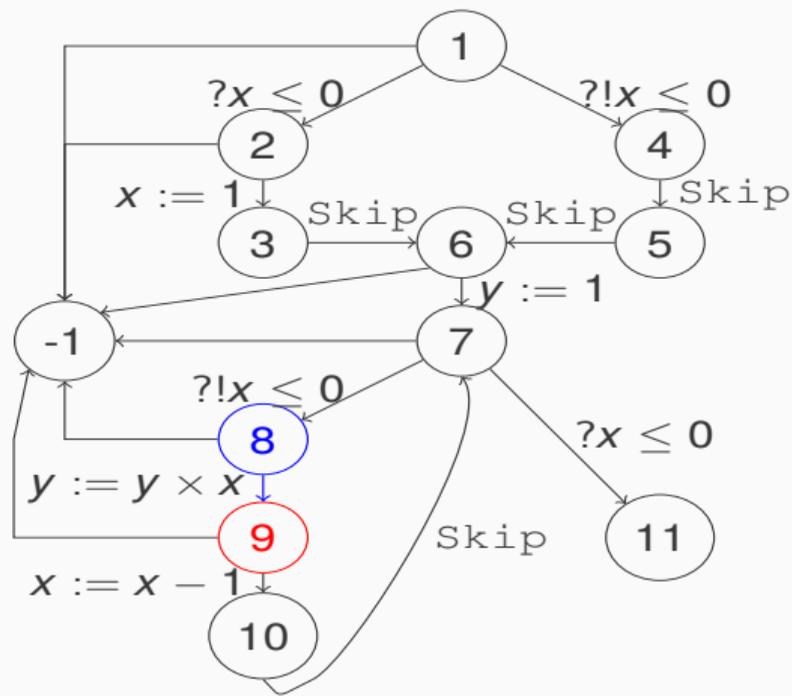
nœud	x	y
s ₁	INT	INT
s ₂	NEG	INT
s ₃	SPOS	INT
s ₄	SPOS	INT
s ₅	SPOS	INT
s ₆	SPOS	INT
s ₇		
s ₈		
s ₉		
s ₁₀		
s ₁₁		
s ₋₁		



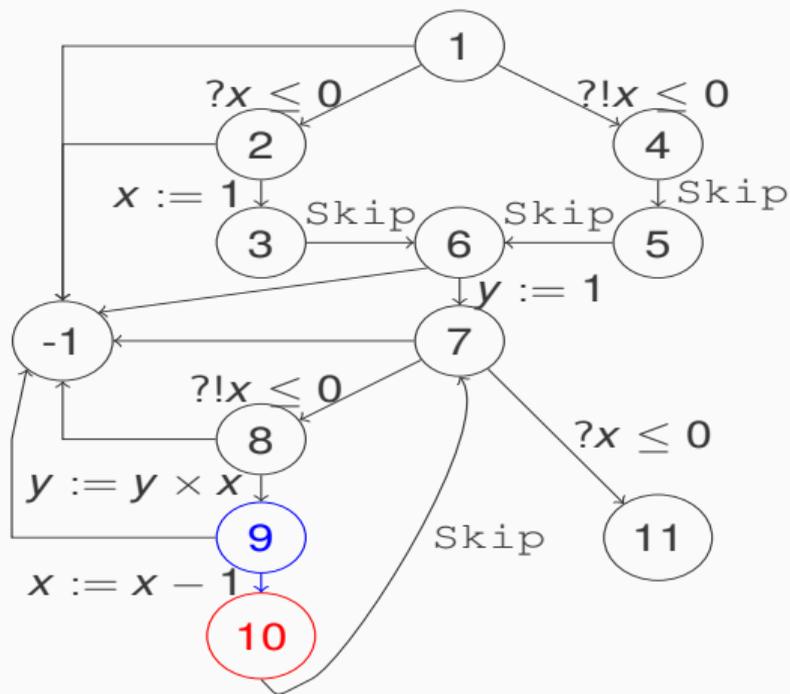
nœud	x	y
s ₁	INT	INT
s ₂	NEG	INT
s ₃	SPOS	INT
s ₄	SPOS	INT
s ₅	SPOS	INT
s ₆	SPOS	INT
s ₇	SPOS	SPOS
s ₈		
s ₉		
s ₁₀		
s ₁₁		
s ₋₁		



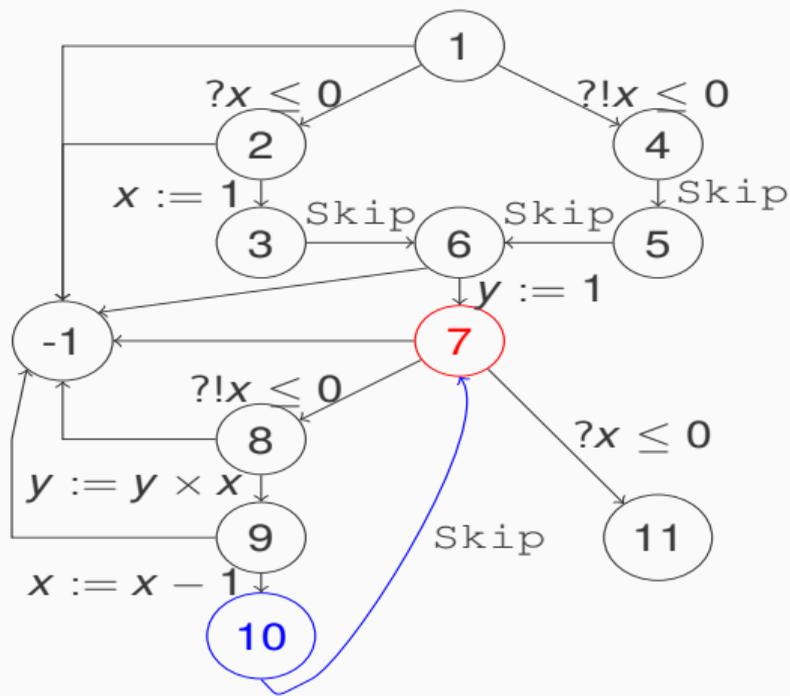
nœud	x	y
s ₁	INT	INT
s ₂	NEG	INT
s ₃	SPOS	INT
s ₄	SPOS	INT
s ₅	SPOS	INT
s ₆	SPOS	INT
s₇	SPOS	SPOS
s₈	SPOS	SPOS
s ₉		
s ₁₀		
s ₁₁		
s ₋₁		



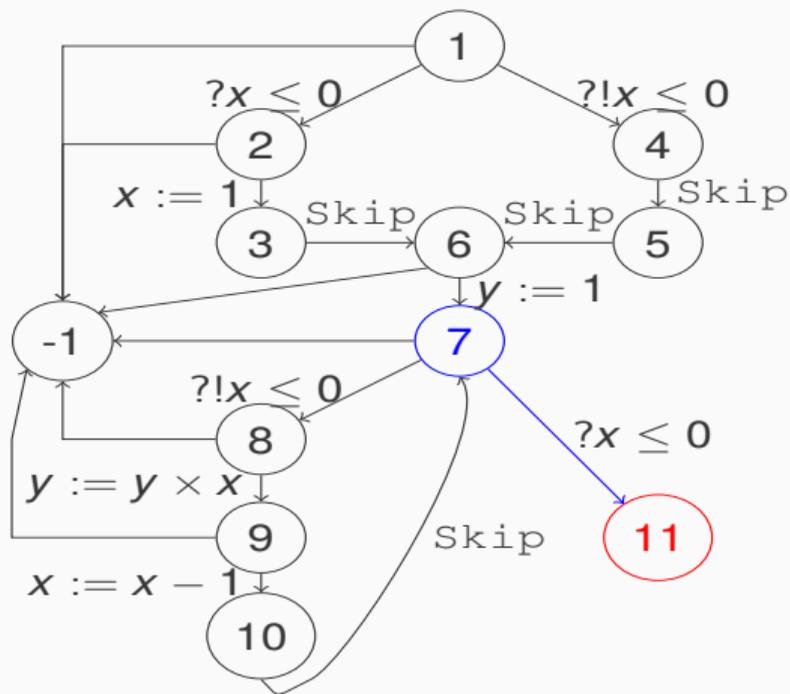
nœud	x	y
s ₁	INT	INT
s ₂	NEG	INT
s ₃	SPOS	INT
s ₄	SPOS	INT
s ₅	SPOS	INT
s ₆	SPOS	INT
s ₇	SPOS	SPOS
s ₈	SPOS	SPOS
s ₉	SPOS	SPOS
s ₁₀		
s ₁₁		
s ₋₁		



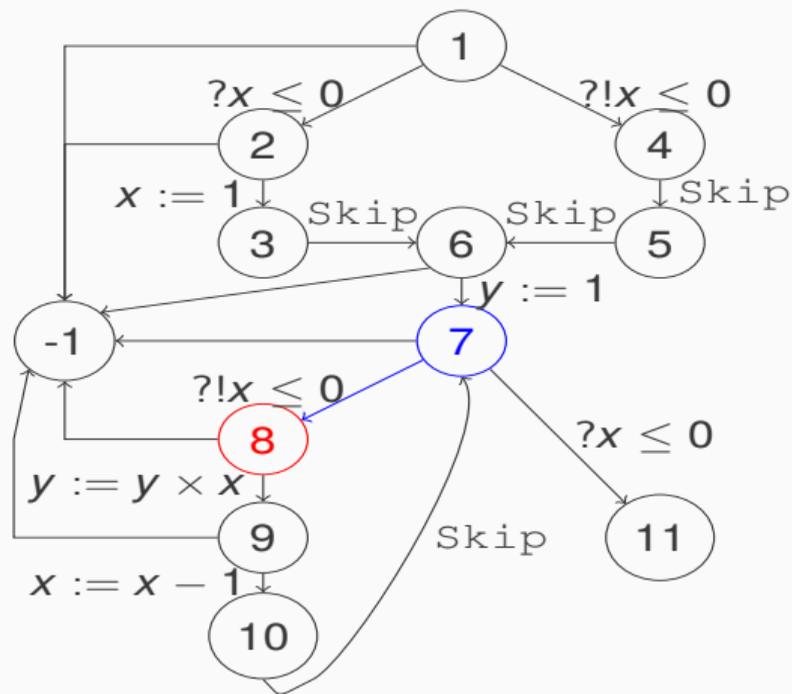
nœud	x	y
s ₁	INT	INT
s ₂	NEG	INT
s ₃	SPOS	INT
s ₄	SPOS	INT
s ₅	SPOS	INT
s ₆	SPOS	INT
s ₇	SPOS	SPOS
s ₈	SPOS	SPOS
s ₉	SPOS	SPOS
s ₁₀	INT	SPOS
s ₁₁		
s ₋₁		



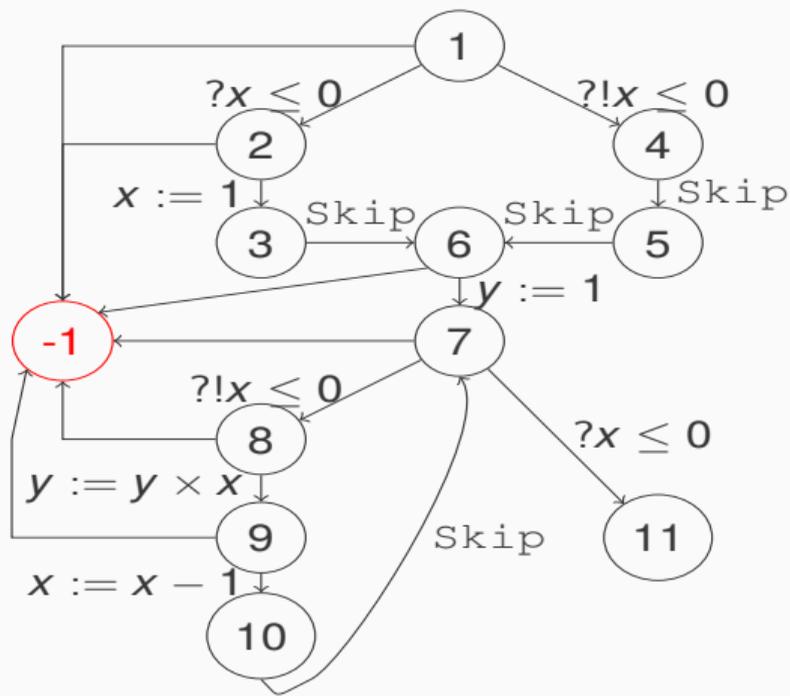
nœud	x	y
s ₁	INT	INT
s ₂	NEG	INT
s ₃	SPOS	INT
s ₄	SPOS	INT
s ₅	SPOS	INT
s ₆	SPOS	INT
s₇	INT	SPOS
s ₈	SPOS	SPOS
s ₉	SPOS	SPOS
s₁₀	INT	SPOS
s ₁₁		
s ₋₁		



nœud	x	y
s ₁	INT	INT
s ₂	NEG	INT
s ₃	SPOS	INT
s ₄	SPOS	INT
s ₅	SPOS	INT
s ₆	SPOS	INT
s₇	INT	SPOS
s ₈	SPOS	SPOS
s ₉	SPOS	SPOS
s ₁₀	INT	SPOS
s₁₁	NEG	SPOS
s ₋₁		



nœud	x	y
s ₁	INT	INT
s ₂	NEG	INT
s ₃	SPOS	INT
s ₄	SPOS	INT
s ₅	SPOS	INT
s ₆	SPOS	INT
s ₇	INT	SPOS
s ₈	SPOS	SPOS
s ₉	SPOS	SPOS
s ₁₀	INT	SPOS
s ₁₁	NEG	SPOS
s ₋₁		

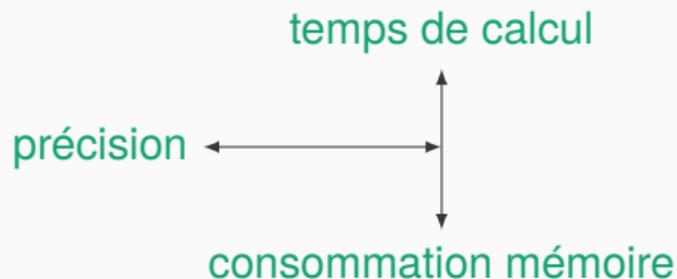


nœud	x	y
s ₁	INT	INT
s ₂	NEG	INT
s ₃	SPOS	INT
s ₄	SPOS	INT
s ₅	SPOS	INT
s ₆	SPOS	INT
s ₇	INT	SPOS
s ₈	SPOS	SPOS
s ₉	SPOS	SPOS
s ₁₀	INT	SPOS
s ₁₁	NEG	SPOS
s ₋₁	⊥	⊥

- > [Terminaison] L'analyse des signes termine toujours.
- > [Correction] Soit s un sommet du graphe de contrôle de P et x une variable.
 $Coll_P(s)(x) \in \gamma(\rho^\sharp(s)(x))$ où $\rho^\sharp(s)$ est l'environnement associé à s par l'analyse.
- > La correction peut se montrer "à la main", en raisonnant par induction sur P , mais il existe une manière systématique de construire l'analyse qui garantit que c'est correct.

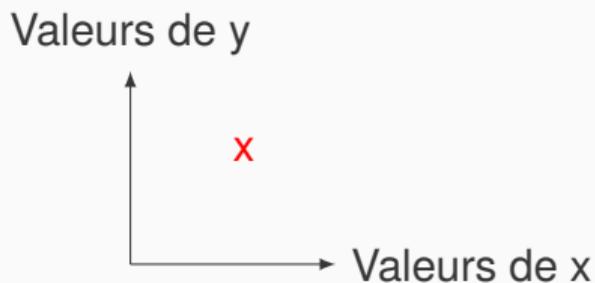
En pratique, bien choisir son domaine abstrait est fondamental

- > doit être suffisamment **précis**
- > en particulier, doit permettre d'énoncer la propriété souhaitée
- > doit être calculable pour un coût **temps/mémoire raisonnable**

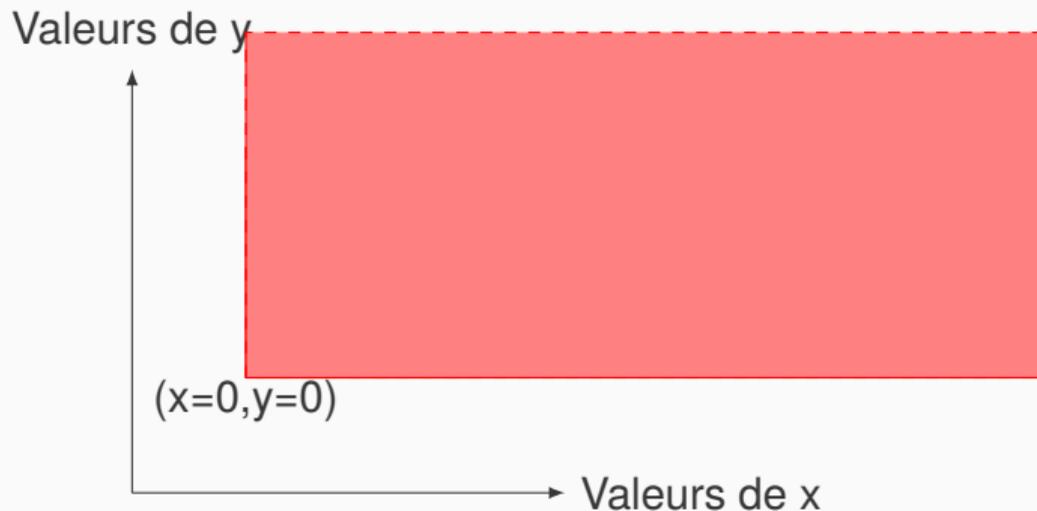


- > **domaines non relationnels** : aucune relation conservée entre les éléments du domaine \Rightarrow peu précis, mais pas cher
- > **domaines relationnels** : relations entre éléments du domaine \Rightarrow plus précis, mais plus cher

- > $x = z$ ($z \in \mathbb{Z}$)
- > domaine non relationnel
- > si la valeur exacte n'est pas connue, perte de toute information

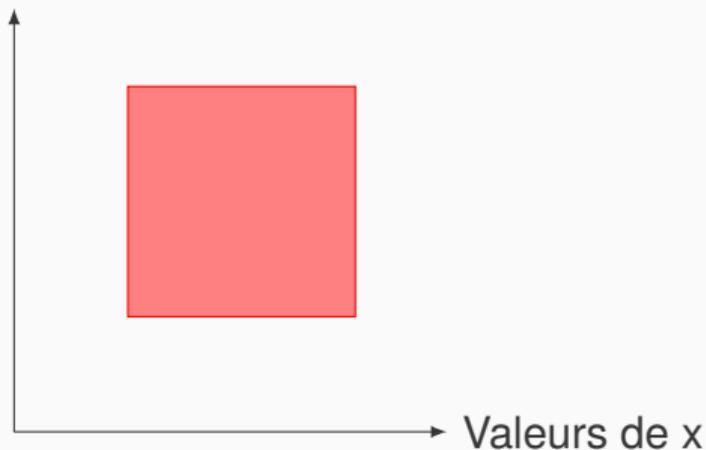


- > $x \text{ op } 0$ $\text{op} \in \{\geq, >, \leq, <, =, \neq\}$
- > domaine non relationnel,
- > conservation de la polarité des valeurs possibles



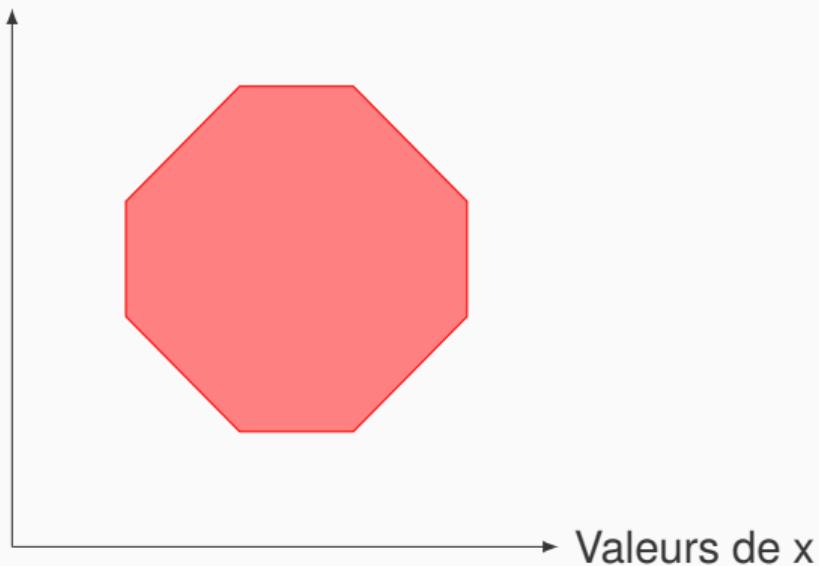
- > $x \in [i_0, i_1]$
- > domaine non relationnel,
- > conservation d'un intervalle regroupant toutes les valeurs possibles

Valeurs de y



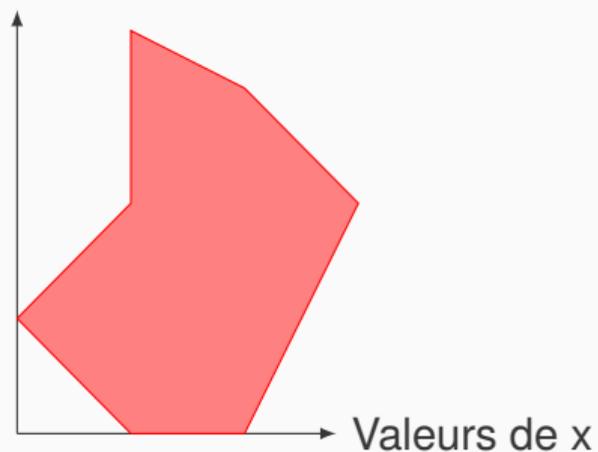
- > $\pm x \pm y \leq c$
- > domaine relationnel
- > conservation de relations linéaires simples entre éléments

Valeurs de y

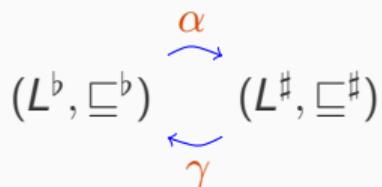


- > $kx + ly \leq c$
- > domaine relationnel
- > relations linéaires complexes entre deux éléments

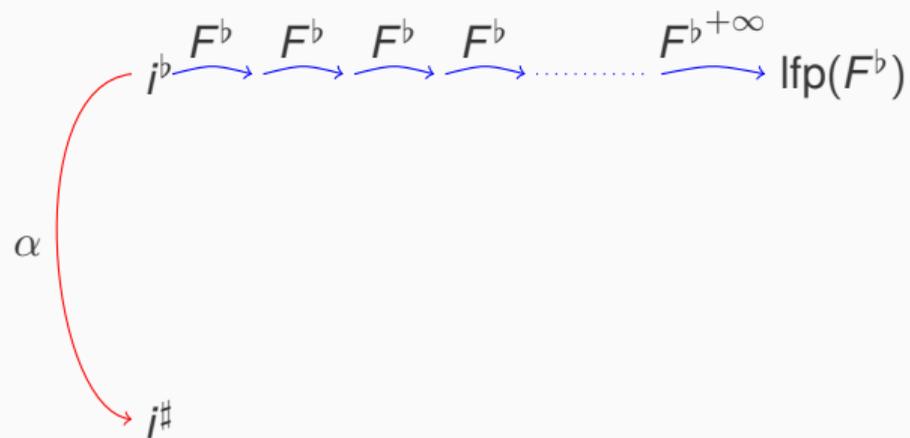
Valeurs de y

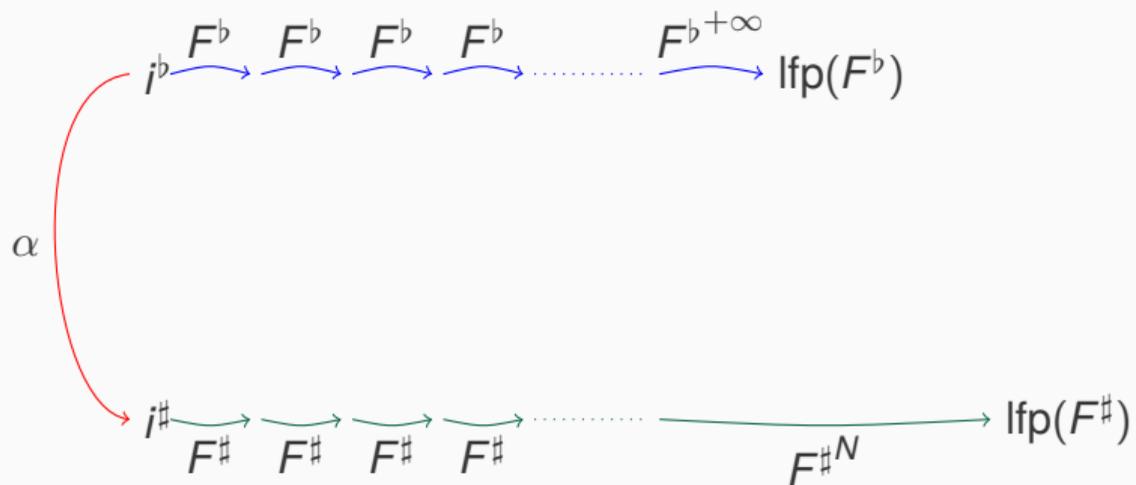


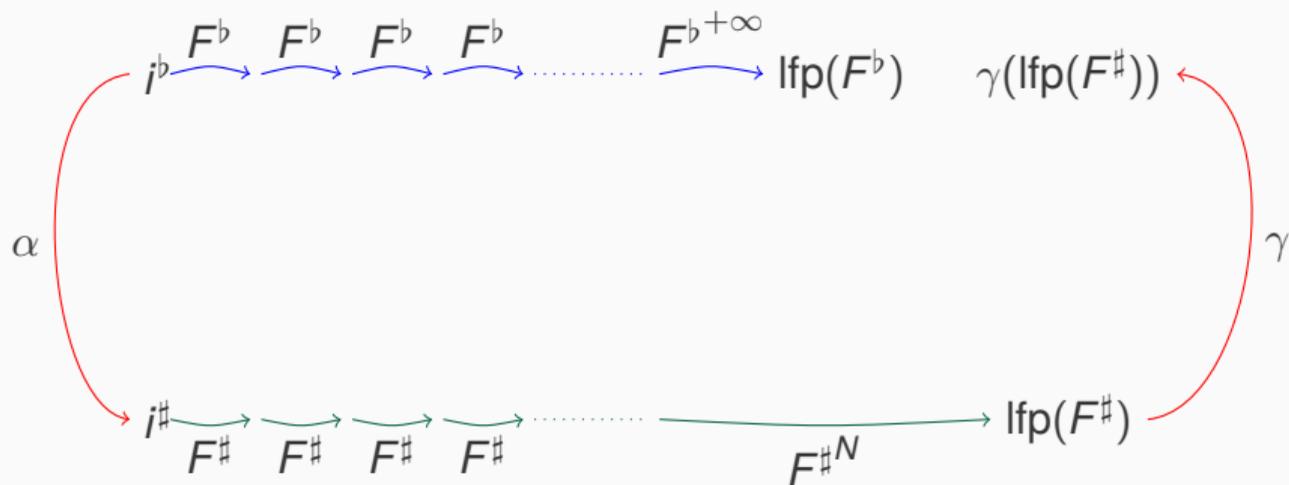
- > treillis concret (L^b, \sqsubseteq^b) pour l'analyse concrète
- > treillis abstrait $(L^\sharp, \sqsubseteq^\sharp)$ pour l'analyse abstraite
- > **relation d'ordre** \sqsubseteq^\sharp : "être moins précis que"
- > opérateur d'**abstraction** α : transforme un environnement concret en un abstrait
- > opérateur de **concrétisation** γ : transforme un environnement abstrait en un concret

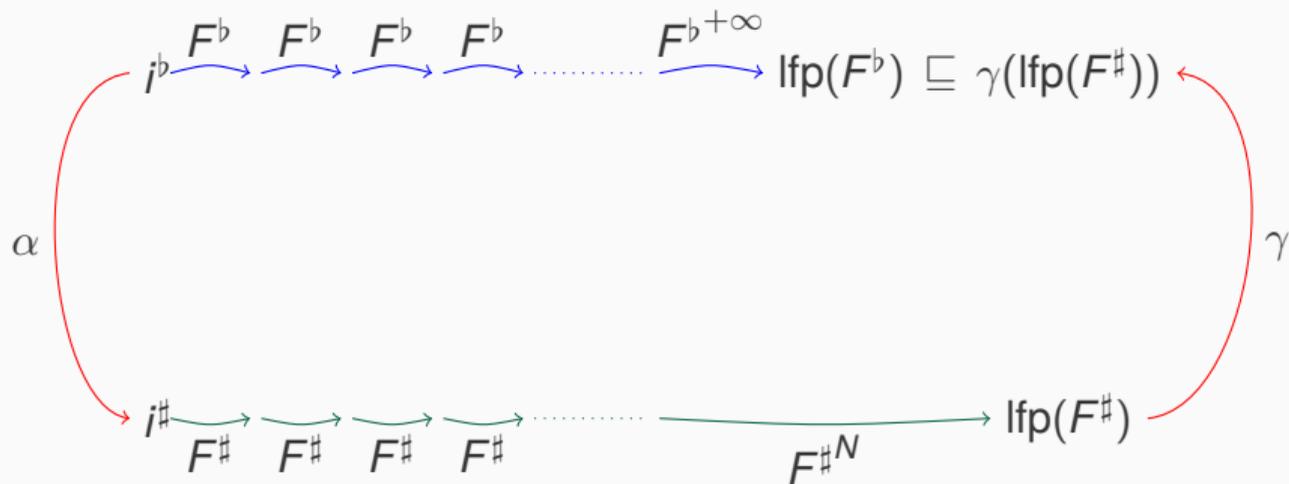










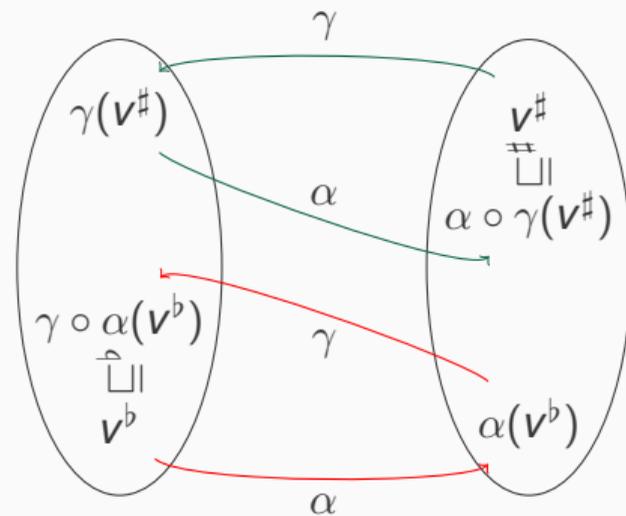


- > Conditions sur α et γ
- > Comment choisir $F^\#$?

Lemme

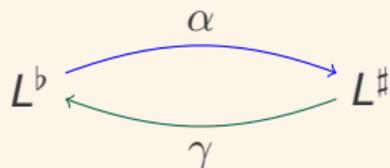
(α, γ) est une **connexion de Galois** si et seulement si les trois conditions suivantes sont satisfaites :

- 1 \rangle α et γ sont **monotones**.
- 2 \rangle $\forall v^b \in L^b, v^b \sqsubseteq^b (\gamma \circ \alpha)(v^b)$.
- 3 \rangle $\forall v^\# \in L^\#, (\alpha \circ \gamma)(v^\#) \sqsubseteq^\# v^\#$



Théorème (Cousot)

Soit une connexion de Galois entre treillis complet L^b et L^\sharp :



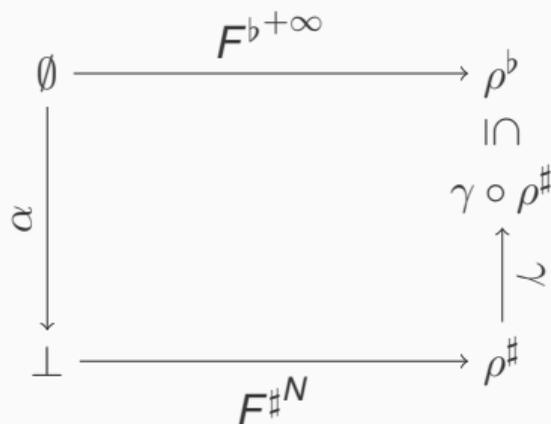
Si $f : L^b \rightarrow L^b$ est une fonction monotone alors :

$$\text{lfp}(f) \sqsubseteq^b \gamma(\text{lfp}(\alpha \circ f \circ \gamma)).$$

Monotonie

Soit $(L, \sqsubseteq, \sqcup, \sqcap)$ un treillis complet, et f et g deux **fonctions monotones** de D dans D .

si pour tout $x \in L$, $f(x) \sqsubseteq g(x)$, alors $\text{lfp}(f) \sqsubseteq \text{lfp}(g)$.



Le théorème de Cousot nous donne un bon candidat pour $F^{\#}$:

$$F^{\#} = \alpha \circ F^b \circ \gamma$$

- > **correct** “automatiquement” par construction,
- > mais **pas forcément calculable** (F^b elle-même ne l’est pas)
- > **approximation plus grossière** possible grâce au théorème de monotonie

α : $\mathcal{P}(\mathbb{Z}) \cup \{\text{erreur}\} \rightarrow \text{Signes}$

$$\alpha(\{0\}) = \mathbb{Z}$$

$$\alpha(Z) = \text{NZ} \quad \text{si } \emptyset \subset Z \subseteq \mathbb{Z}^*$$

$$\alpha(Z) = \text{POS} \quad \text{si } \emptyset \subset Z \subseteq \mathbb{Z}^+ = \mathbb{N}$$

$$\alpha(Z) = \text{NEG} \quad \text{si } \emptyset \subset Z \subseteq \mathbb{Z}^-$$

$$\alpha(Z) = \text{SPOS} \quad \text{si } \emptyset \subset Z \subseteq \mathbb{Z}^{+*} = \mathbb{N}^*$$

$$\alpha(Z) = \text{SNEG} \quad \text{si } \emptyset \subset Z \subseteq \mathbb{Z}^{-*}$$

$$\alpha(Z) = \text{INT} \quad \text{si erreur} \notin Z \supset \emptyset$$

$$\alpha(\{\text{erreur}\}) = \text{Erreur}$$

$$\alpha(\emptyset) = \perp$$

$$\alpha(Z) = \top \quad \text{si erreur} \in Z \text{ et } \emptyset \subset Z \setminus \{\text{erreur}\} \subseteq \mathbb{Z}$$

γ : *Signes* $\rightarrow \mathcal{P}(\mathbb{Z}) \cup \{\text{erreur}\}$

$$\gamma(\top) = \mathbb{Z} \cup \{\text{erreur}\}$$

$$\gamma(\text{INT}) = \mathbb{Z}$$

$$\gamma(\text{POS}) = \{v \in \mathbb{Z} \mid v \geq 0\}$$

$$\gamma(\text{NEG}) = \{v \in \mathbb{Z} \mid v \leq 0\}$$

$$\gamma(\text{SPOS}) = \{v \in \mathbb{Z} \mid v > 0\}$$

$$\gamma(\text{SNEG}) = \{v \in \mathbb{Z} \mid v < 0\}$$

$$\gamma(\text{Z}) = \{0\}$$

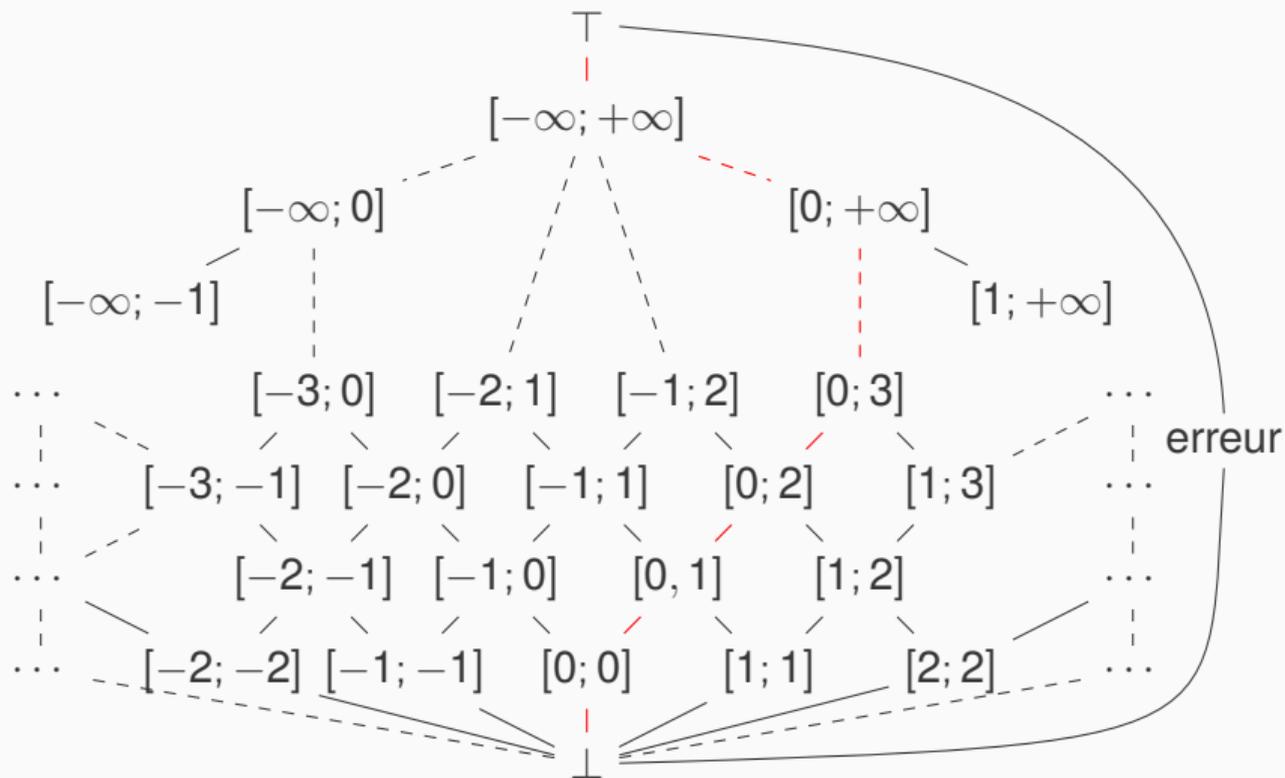
$$\gamma(\text{NZ}) = \{v \in \mathbb{Z} \mid v \neq 0\}$$

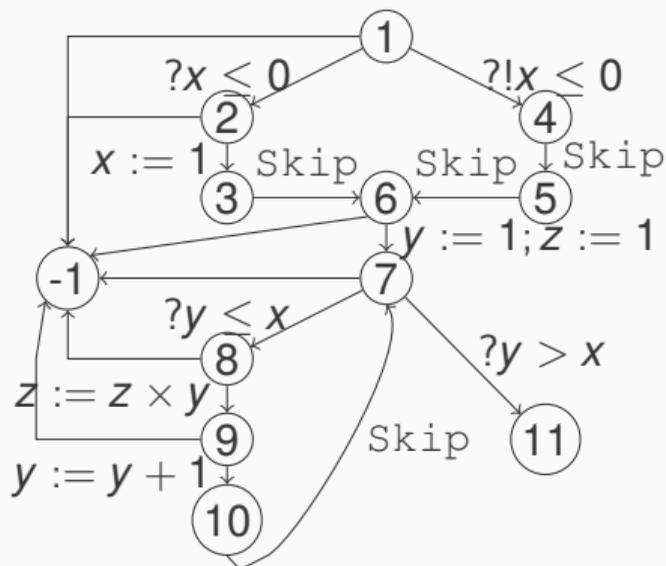
$$\gamma(\perp) = \emptyset$$

Lemme

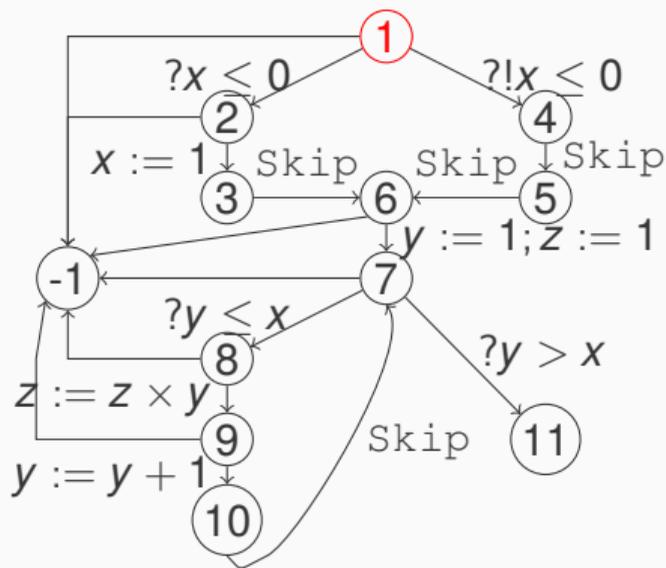
(α, γ) est une insertion de Galois.

- > on peut définir une **interprétation abstraite de la sémantique collectrice**
- > **correcte** par construction :-)
- > termine dans le cas du treillis des signes (treillis fini)
- > **ne termine pas** en présence de boucle si le treillis ne vérifie pas la condition de chaîne croissante :-)

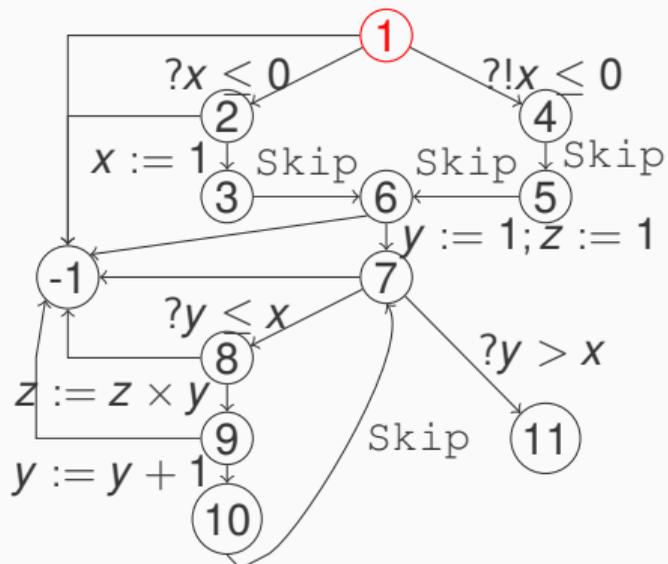




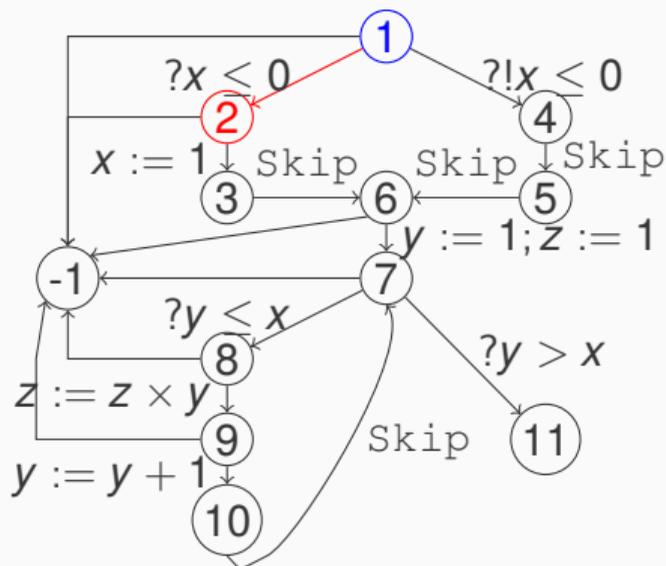
- $R(s_1) = \perp \#$
- $R(s_2) = \perp \#$
- $R(s_3) = \perp \#$
- $R(s_4) = \perp \#$
- $R(s_5) = \perp \#$
- $R(s_6) = \perp \#$
- $R(s_7) = \perp \#$
- $R(s_8) = \perp \#$
- $R(s_9) = \perp \#$
- $R(s_{10}) = \perp \#$
- $R(s_{11}) = \perp \#$
- $R(s_{-1}) = \perp \#$



- $R(s_1) = \perp \#$
- $R(s_2) = \perp \#$
- $R(s_3) = \perp \#$
- $R(s_4) = \perp \#$
- $R(s_5) = \perp \#$
- $R(s_6) = \perp \#$
- $R(s_7) = \perp \#$
- $R(s_8) = \perp \#$
- $R(s_9) = \perp \#$
- $R(s_{10}) = \perp \#$
- $R(s_{11}) = \perp \#$
- $R(s_{-1}) = \perp \#$



$R(s_1) = x \leftarrow [-\infty; +\infty]$
 $R(s_2) = \perp \#$
 $R(s_3) = \perp \#$
 $R(s_4) = \perp \#$
 $R(s_5) = \perp \#$
 $R(s_6) = \perp \#$
 $R(s_7) = \perp \#$
 $R(s_8) = \perp \#$
 $R(s_9) = \perp \#$
 $R(s_{10}) = \perp \#$
 $R(s_{11}) = \perp \#$
 $R(s_{-1}) = \perp \#$



$$R(s_1) = x \leftarrow [-\infty; +\infty]$$

$$R(s_2) = \perp \#$$

$$R(s_3) = \perp \#$$

$$R(s_4) = \perp \#$$

$$R(s_5) = \perp \#$$

$$R(s_6) = \perp \#$$

$$R(s_7) = \perp \#$$

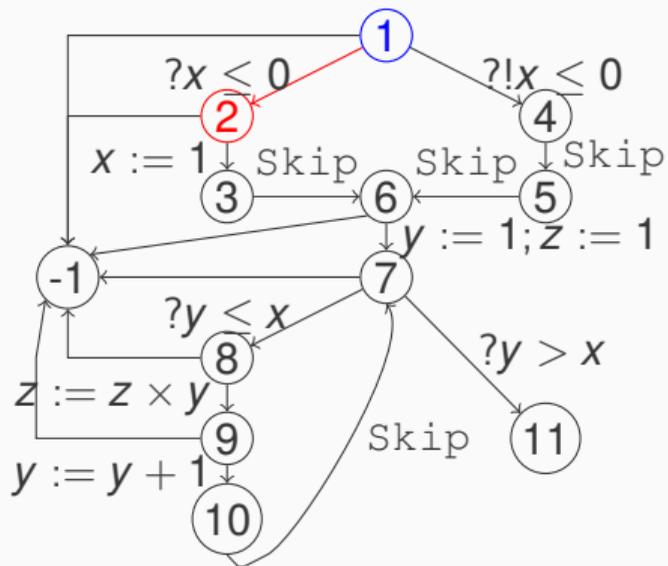
$$R(s_8) = \perp \#$$

$$R(s_9) = \perp \#$$

$$R(s_{10}) = \perp \#$$

$$R(s_{11}) = \perp \#$$

$$R(s_{-1}) = \perp \#$$



$$R(s_1) = x \leftarrow [-\infty; +\infty]$$

$$R(s_2) = x \leftarrow [-\infty; 0]$$

$$R(s_3) = \perp \#$$

$$R(s_4) = \perp \#$$

$$R(s_5) = \perp \#$$

$$R(s_6) = \perp \#$$

$$R(s_7) = \perp \#$$

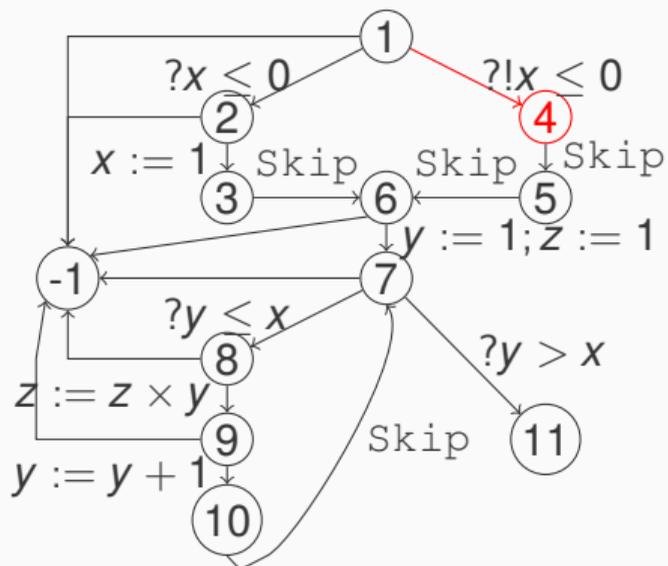
$$R(s_8) = \perp \#$$

$$R(s_9) = \perp \#$$

$$R(s_{10}) = \perp \#$$

$$R(s_{11}) = \perp \#$$

$$R(s_{-1}) = \perp \#$$



$$R(s_1) = x \leftarrow [-\infty; +\infty]$$

$$R(s_2) = x \leftarrow [-\infty; 0]$$

$$R(s_3) = \perp \#$$

$$R(s_4) = \perp \#$$

$$R(s_5) = \perp \#$$

$$R(s_6) = \perp \#$$

$$R(s_7) = \perp \#$$

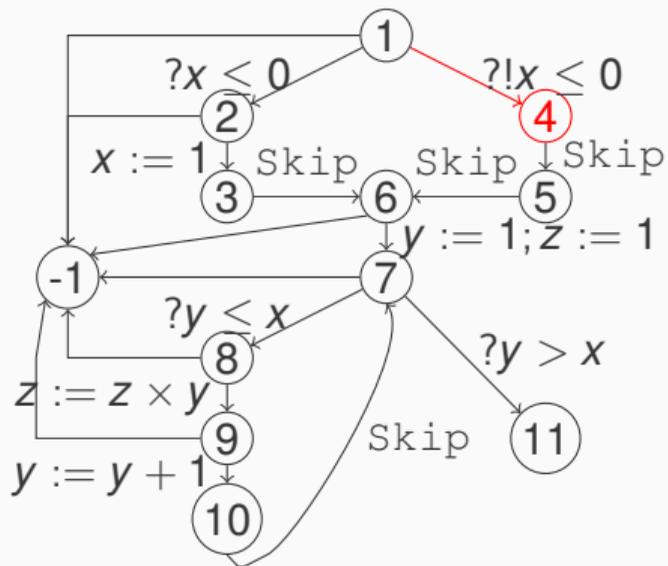
$$R(s_8) = \perp \#$$

$$R(s_9) = \perp \#$$

$$R(s_{10}) = \perp \#$$

$$R(s_{11}) = \perp \#$$

$$R(s_{-1}) = \perp \#$$



$$R(s_1) = x \leftarrow [-\infty; +\infty]$$

$$R(s_2) = x \leftarrow [-\infty; 0]$$

$$R(s_3) = \perp \#$$

$$R(s_4) = x \leftarrow [1; +\infty]$$

$$R(s_5) = \perp \#$$

$$R(s_6) = \perp \#$$

$$R(s_7) = \perp \#$$

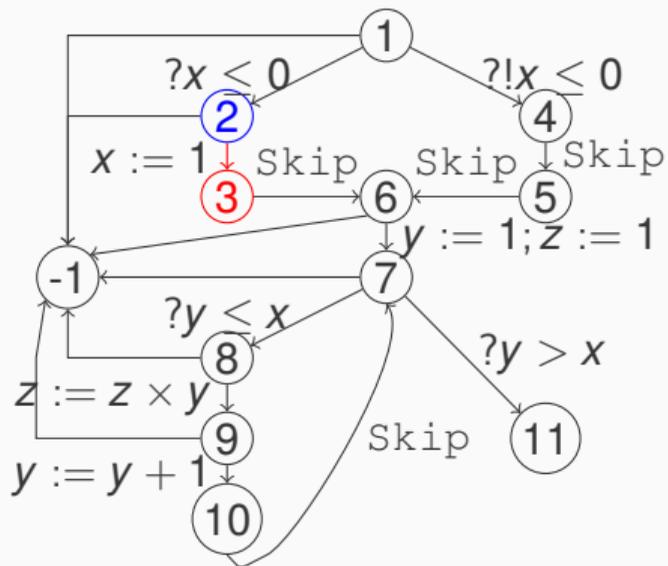
$$R(s_8) = \perp \#$$

$$R(s_9) = \perp \#$$

$$R(s_{10}) = \perp \#$$

$$R(s_{11}) = \perp \#$$

$$R(s_{-1}) = \perp \#$$



$$R(s_1) = x \leftarrow [-\infty; +\infty]$$

$$R(s_2) = x \leftarrow [-\infty; 0]$$

$$R(s_3) = \perp \#$$

$$R(s_4) = x \leftarrow [1; +\infty]$$

$$R(s_5) = \perp \#$$

$$R(s_6) = \perp \#$$

$$R(s_7) = \perp \#$$

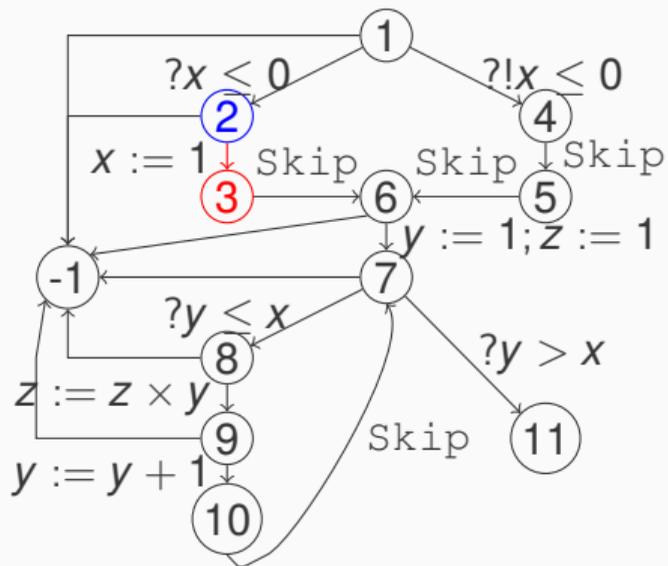
$$R(s_8) = \perp \#$$

$$R(s_9) = \perp \#$$

$$R(s_{10}) = \perp \#$$

$$R(s_{11}) = \perp \#$$

$$R(s_{-1}) = \perp \#$$



$$R(s_1) = x \leftarrow [-\infty; +\infty]$$

$$R(s_2) = x \leftarrow [-\infty; 0]$$

$$R(s_3) = x \leftarrow [1; 1]$$

$$R(s_4) = x \leftarrow [1; +\infty]$$

$$R(s_5) = \perp \#$$

$$R(s_6) = \perp \#$$

$$R(s_7) = \perp \#$$

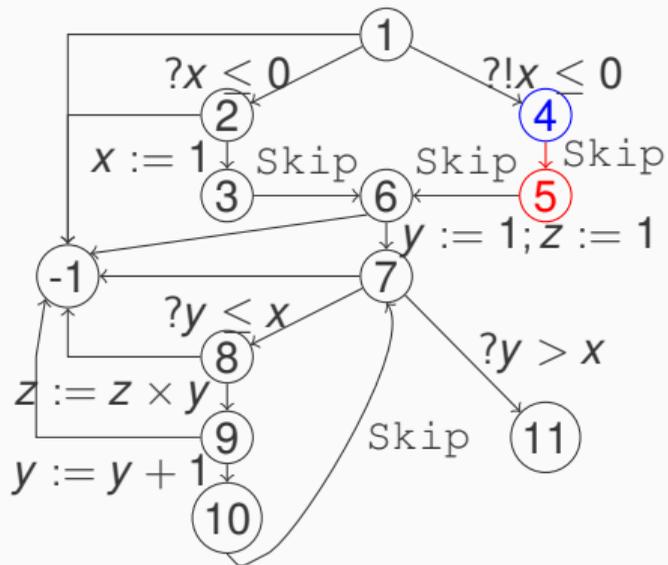
$$R(s_8) = \perp \#$$

$$R(s_9) = \perp \#$$

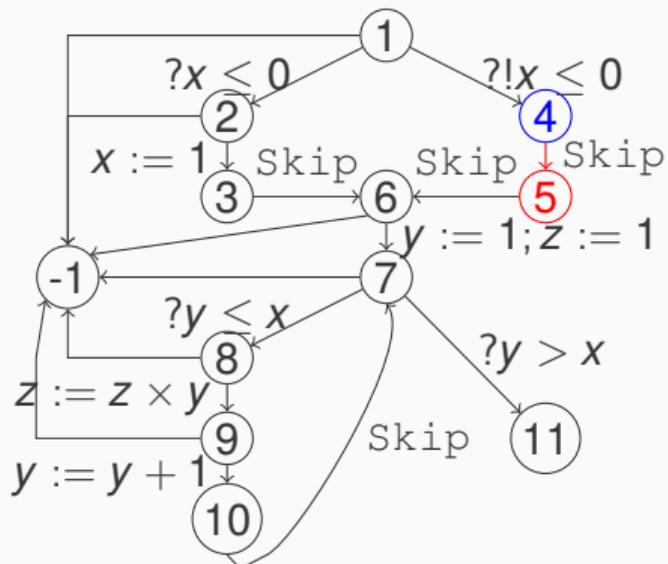
$$R(s_{10}) = \perp \#$$

$$R(s_{11}) = \perp \#$$

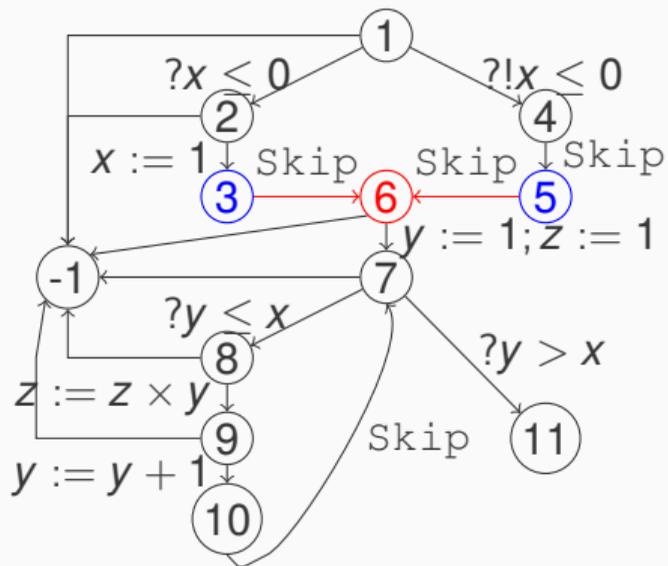
$$R(s_{-1}) = \perp \#$$



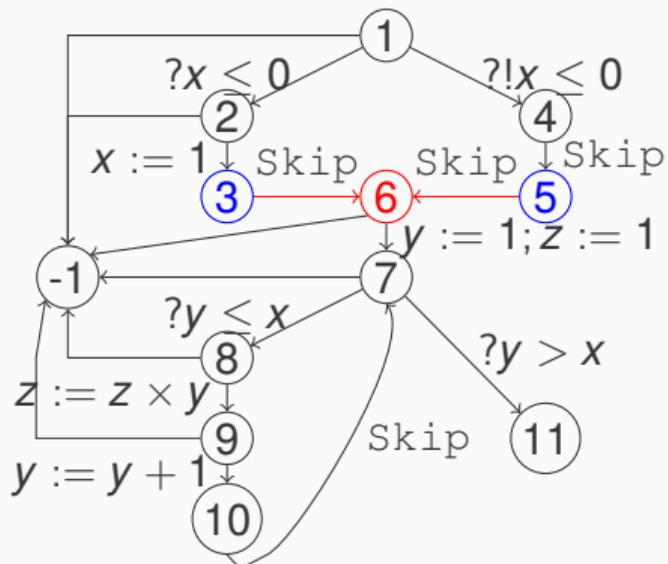
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = \perp \#$
- $R(s_6) = \perp \#$
- $R(s_7) = \perp \#$
- $R(s_8) = \perp \#$
- $R(s_9) = \perp \#$
- $R(s_{10}) = \perp \#$
- $R(s_{11}) = \perp \#$
- $R(s_{-1}) = \perp \#$



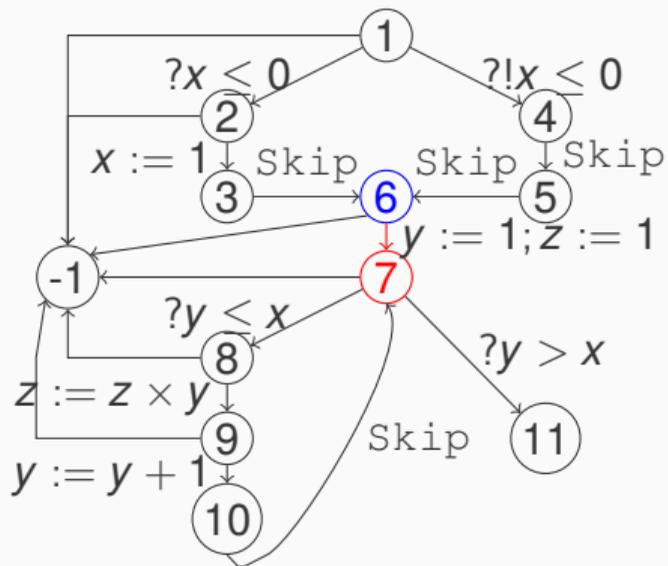
$R(s_1) = x \leftarrow [-\infty; +\infty]$
 $R(s_2) = x \leftarrow [-\infty; 0]$
 $R(s_3) = x \leftarrow [1; 1]$
 $R(s_4) = x \leftarrow [1; +\infty]$
 $R(s_5) = x \leftarrow [1; +\infty]$
 $R(s_6) = \perp^\#$
 $R(s_7) = \perp^\#$
 $R(s_8) = \perp^\#$
 $R(s_9) = \perp^\#$
 $R(s_{10}) = \perp^\#$
 $R(s_{11}) = \perp^\#$
 $R(s_{-1}) = \perp^\#$



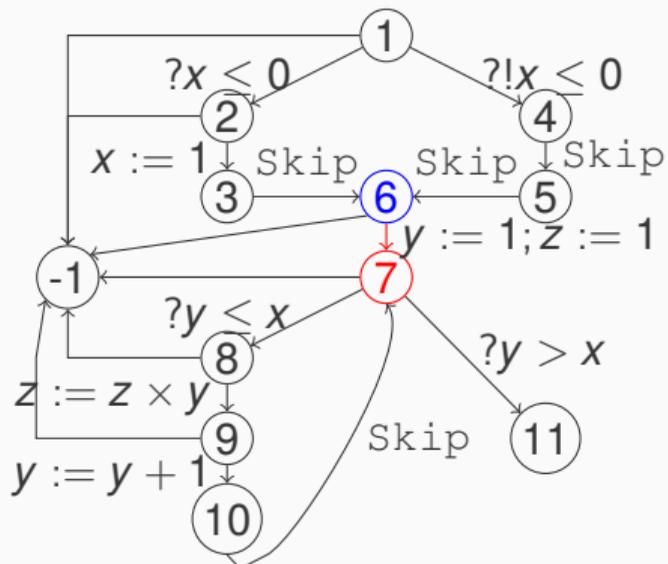
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = \perp \#$
- $R(s_7) = \perp \#$
- $R(s_8) = \perp \#$
- $R(s_9) = \perp \#$
- $R(s_{10}) = \perp \#$
- $R(s_{11}) = \perp \#$
- $R(s_{-1}) = \perp \#$



- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = \perp^\#$
- $R(s_8) = \perp^\#$
- $R(s_9) = \perp^\#$
- $R(s_{10}) = \perp^\#$
- $R(s_{11}) = \perp^\#$
- $R(s_{-1}) = \perp^\#$



- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = \perp^\#$
- $R(s_8) = \perp^\#$
- $R(s_9) = \perp^\#$
- $R(s_{10}) = \perp^\#$
- $R(s_{11}) = \perp^\#$
- $R(s_{-1}) = \perp^\#$



$$R(s_1) = x \leftarrow [-\infty; +\infty]$$

$$R(s_2) = x \leftarrow [-\infty; 0]$$

$$R(s_3) = x \leftarrow [1; 1]$$

$$R(s_4) = x \leftarrow [1; +\infty]$$

$$R(s_5) = x \leftarrow [1; +\infty]$$

$$R(s_6) = x \leftarrow [1; +\infty]$$

$$R(s_7) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$$

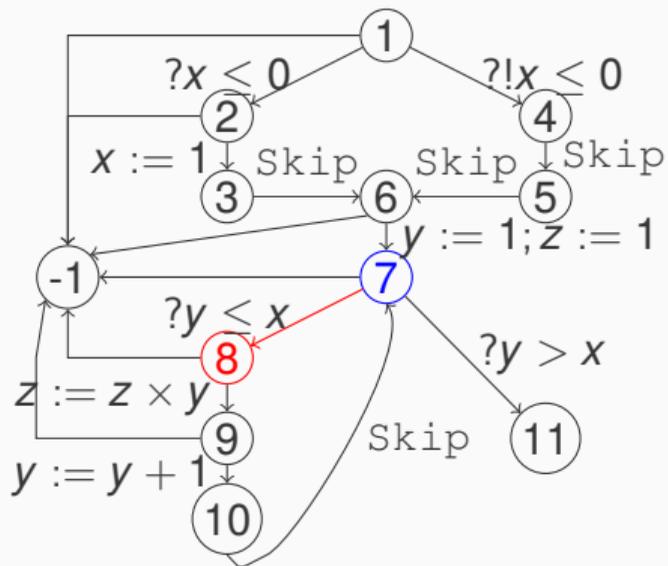
$$R(s_8) = \perp^\#$$

$$R(s_9) = \perp^\#$$

$$R(s_{10}) = \perp^\#$$

$$R(s_{11}) = \perp^\#$$

$$R(s_{-1}) = \perp^\#$$



$$R(s_1) = x \leftarrow [-\infty; +\infty]$$

$$R(s_2) = x \leftarrow [-\infty; 0]$$

$$R(s_3) = x \leftarrow [1; 1]$$

$$R(s_4) = x \leftarrow [1; +\infty]$$

$$R(s_5) = x \leftarrow [1; +\infty]$$

$$R(s_6) = x \leftarrow [1; +\infty]$$

$$R(s_7) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$$

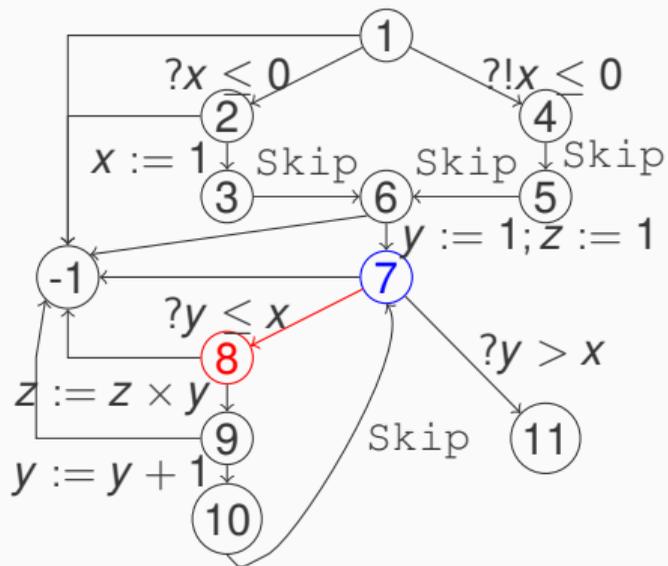
$$R(s_8) = \perp \#$$

$$R(s_9) = \perp \#$$

$$R(s_{10}) = \perp \#$$

$$R(s_{11}) = \perp \#$$

$$R(s_{-1}) = \perp \#$$



$$R(s_1) = x \leftarrow [-\infty; +\infty]$$

$$R(s_2) = x \leftarrow [-\infty; 0]$$

$$R(s_3) = x \leftarrow [1; 1]$$

$$R(s_4) = x \leftarrow [1; +\infty]$$

$$R(s_5) = x \leftarrow [1; +\infty]$$

$$R(s_6) = x \leftarrow [1; +\infty]$$

$$R(s_7) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$$

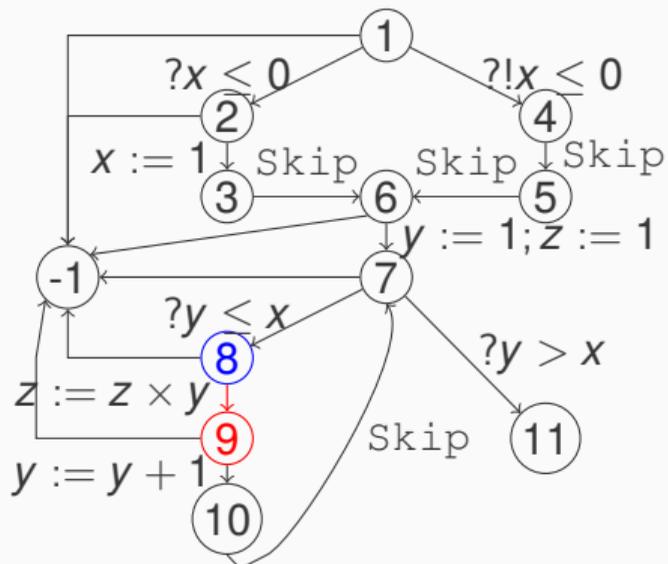
$$R(s_8) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$$

$$R(s_9) = \perp^\#$$

$$R(s_{10}) = \perp^\#$$

$$R(s_{11}) = \perp^\#$$

$$R(s_{-1}) = \perp^\#$$



$$R(s_1) = x \leftarrow [-\infty; +\infty]$$

$$R(s_2) = x \leftarrow [-\infty; 0]$$

$$R(s_3) = x \leftarrow [1; 1]$$

$$R(s_4) = x \leftarrow [1; +\infty]$$

$$R(s_5) = x \leftarrow [1; +\infty]$$

$$R(s_6) = x \leftarrow [1; +\infty]$$

$$R(s_7) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$$

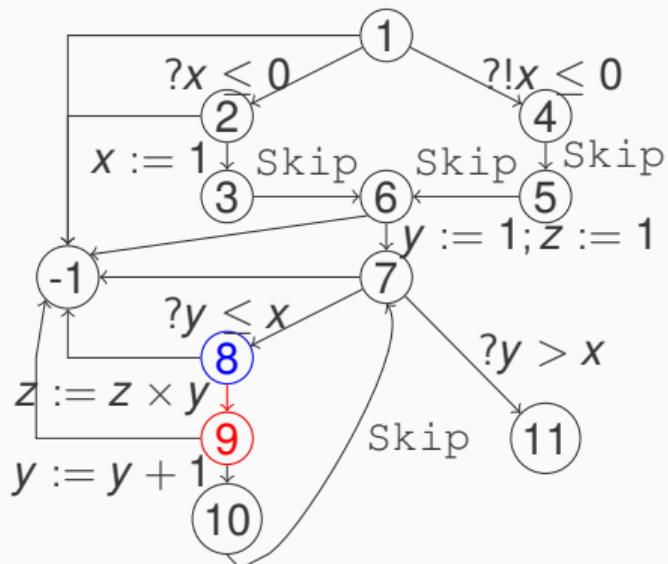
$$R(s_8) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$$

$$R(s_9) = \perp^\#$$

$$R(s_{10}) = \perp^\#$$

$$R(s_{11}) = \perp^\#$$

$$R(s_{-1}) = \perp^\#$$



$$R(s_1) = x \leftarrow [-\infty; +\infty]$$

$$R(s_2) = x \leftarrow [-\infty; 0]$$

$$R(s_3) = x \leftarrow [1; 1]$$

$$R(s_4) = x \leftarrow [1; +\infty]$$

$$R(s_5) = x \leftarrow [1; +\infty]$$

$$R(s_6) = x \leftarrow [1; +\infty]$$

$$R(s_7) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$$

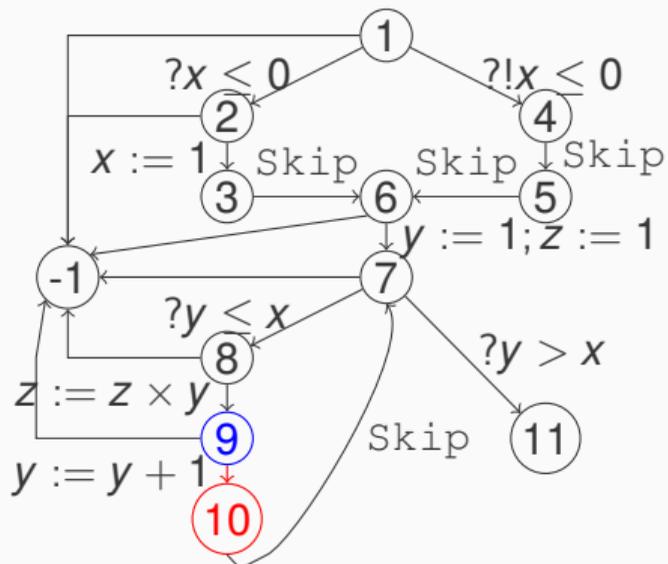
$$R(s_8) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$$

$$R(s_9) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$$

$$R(s_{10}) = \perp^\#$$

$$R(s_{11}) = \perp^\#$$

$$R(s_{-1}) = \perp^\#$$



$$R(s_1) = x \leftarrow [-\infty; +\infty]$$

$$R(s_2) = x \leftarrow [-\infty; 0]$$

$$R(s_3) = x \leftarrow [1; 1]$$

$$R(s_4) = x \leftarrow [1; +\infty]$$

$$R(s_5) = x \leftarrow [1; +\infty]$$

$$R(s_6) = x \leftarrow [1; +\infty]$$

$$R(s_7) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$$

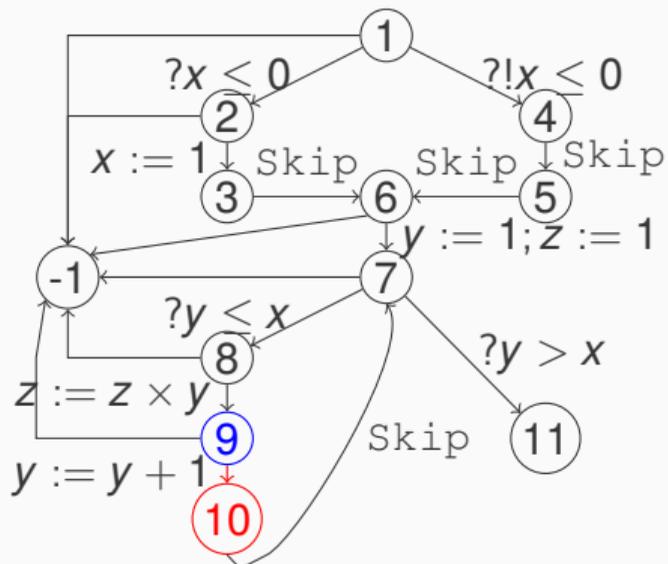
$$R(s_8) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$$

$$R(s_9) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$$

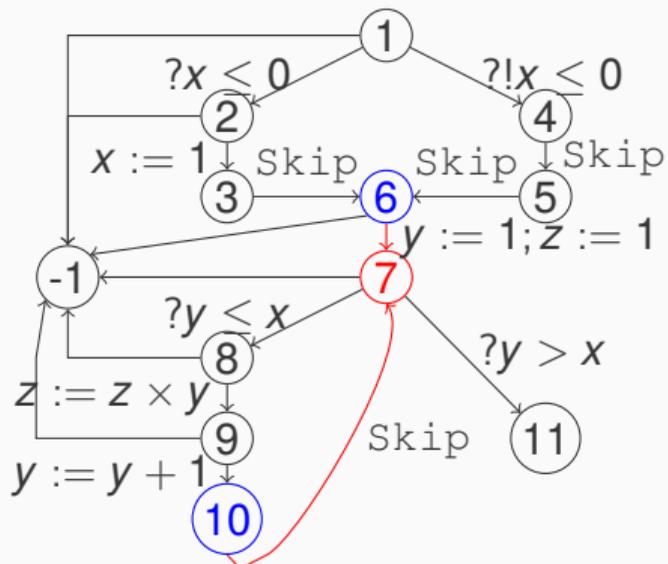
$$R(s_{10}) = \perp^\#$$

$$R(s_{11}) = \perp^\#$$

$$R(s_{-1}) = \perp^\#$$



- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_8) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_9) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_{10}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{11}) = \perp^\#$
- $R(s_{-1}) = \perp^\#$



$$R(s_1) = x \leftarrow [-\infty; +\infty]$$

$$R(s_2) = x \leftarrow [-\infty; 0]$$

$$R(s_3) = x \leftarrow [1; 1]$$

$$R(s_4) = x \leftarrow [1; +\infty]$$

$$R(s_5) = x \leftarrow [1; +\infty]$$

$$R(s_6) = x \leftarrow [1; +\infty]$$

$$R(s_7) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$$

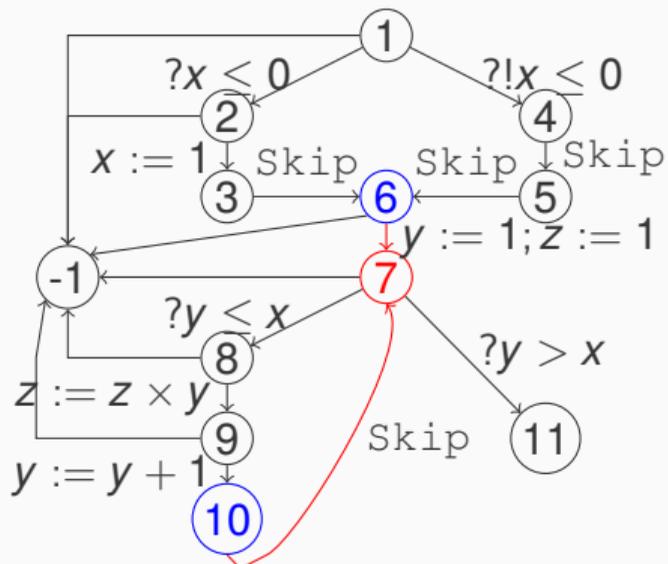
$$R(s_8) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$$

$$R(s_9) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$$

$$R(s_{10}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$$

$$R(s_{11}) = \perp^\#$$

$$R(s_{-1}) = \perp^\#$$



$$R(s_1) = x \leftarrow [-\infty; +\infty]$$

$$R(s_2) = x \leftarrow [-\infty; 0]$$

$$R(s_3) = x \leftarrow [1; 1]$$

$$R(s_4) = x \leftarrow [1; +\infty]$$

$$R(s_5) = x \leftarrow [1; +\infty]$$

$$R(s_6) = x \leftarrow [1; +\infty]$$

$$R(s_7) = y \leftarrow [1; 2]; z \leftarrow [1; 1]$$

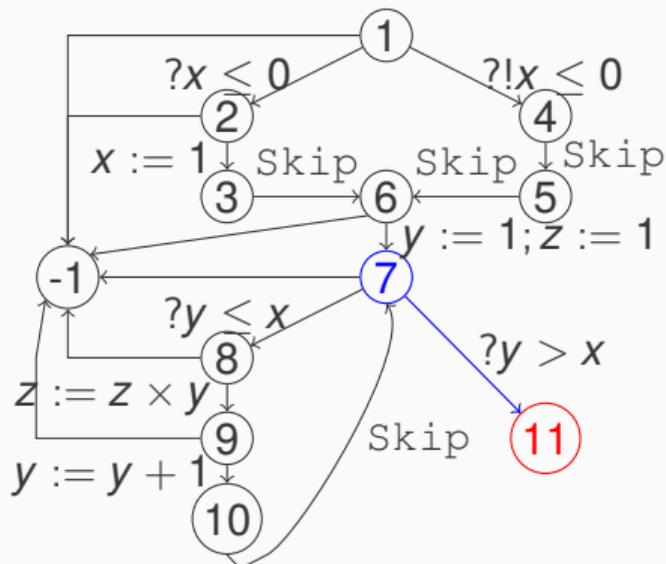
$$R(s_8) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$$

$$R(s_9) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$$

$$R(s_{10}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$$

$$R(s_{11}) = \perp^\#$$

$$R(s_{-1}) = \perp^\#$$



$$R(s_1) = x \leftarrow [-\infty; +\infty]$$

$$R(s_2) = x \leftarrow [-\infty; 0]$$

$$R(s_3) = x \leftarrow [1; 1]$$

$$R(s_4) = x \leftarrow [1; +\infty]$$

$$R(s_5) = x \leftarrow [1; +\infty]$$

$$R(s_6) = x \leftarrow [1; +\infty]$$

$$R(s_7) = y \leftarrow [1; 2]; z \leftarrow [1; 1]$$

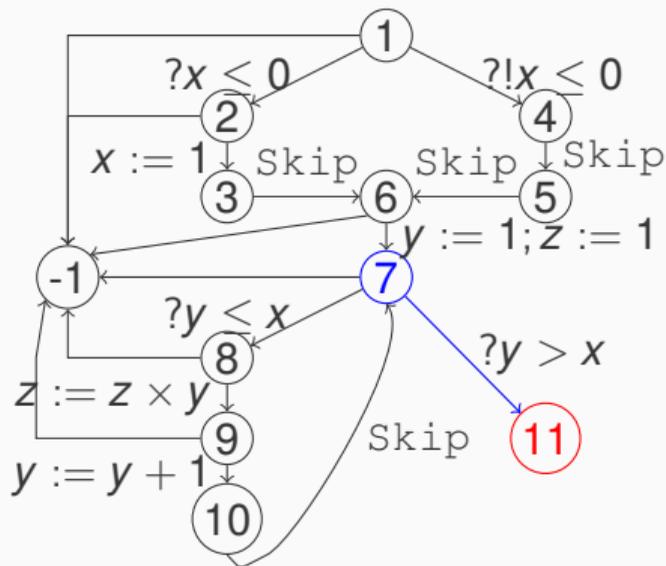
$$R(s_8) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$$

$$R(s_9) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$$

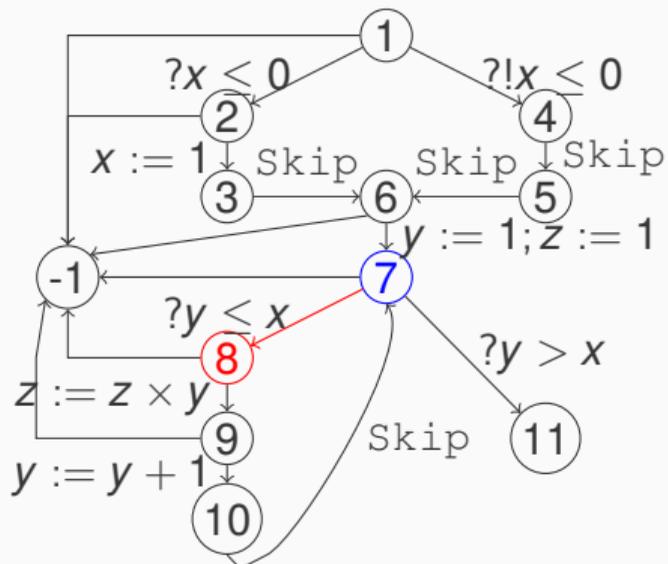
$$R(s_{10}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$$

$$R(s_{11}) = \perp^\#$$

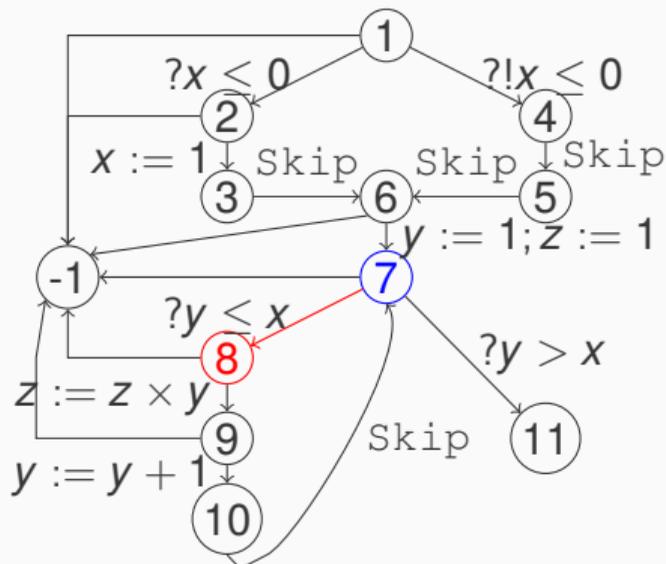
$$R(s_{-1}) = \perp^\#$$



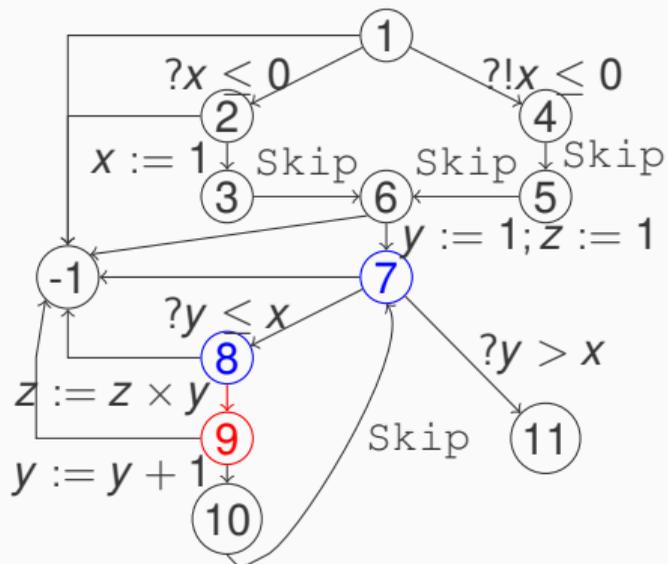
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; 2]; z \leftarrow [1; 1]$
- $R(s_8) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_9) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_{10}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{11}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{-1}) = \perp^\#$



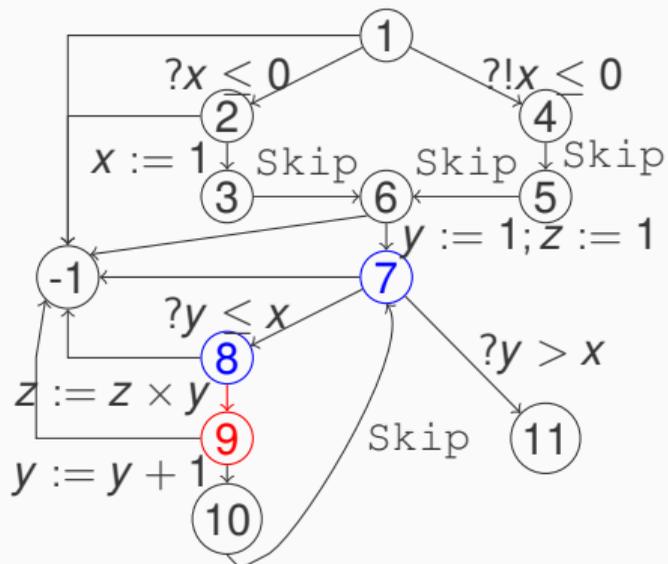
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; 2]; z \leftarrow [1; 1]$
- $R(s_8) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_9) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_{10}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{11}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{-1}) = \perp^\#$



- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; 2]; z \leftarrow [1; 1]$
- $R(s_8) = y \leftarrow [1; 2]; z \leftarrow [1; 1]$
- $R(s_9) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_{10}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{11}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{-1}) = \perp^\#$



- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; 2]; z \leftarrow [1; 1]$
- $R(s_8) = y \leftarrow [1; 2]; z \leftarrow [1; 1]$
- $R(s_9) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_{10}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{11}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{-1}) = \perp^\#$



$$R(s_1) = x \leftarrow [-\infty; +\infty]$$

$$R(s_2) = x \leftarrow [-\infty; 0]$$

$$R(s_3) = x \leftarrow [1; 1]$$

$$R(s_4) = x \leftarrow [1; +\infty]$$

$$R(s_5) = x \leftarrow [1; +\infty]$$

$$R(s_6) = x \leftarrow [1; +\infty]$$

$$R(s_7) = y \leftarrow [1; 2]; z \leftarrow [1; 1]$$

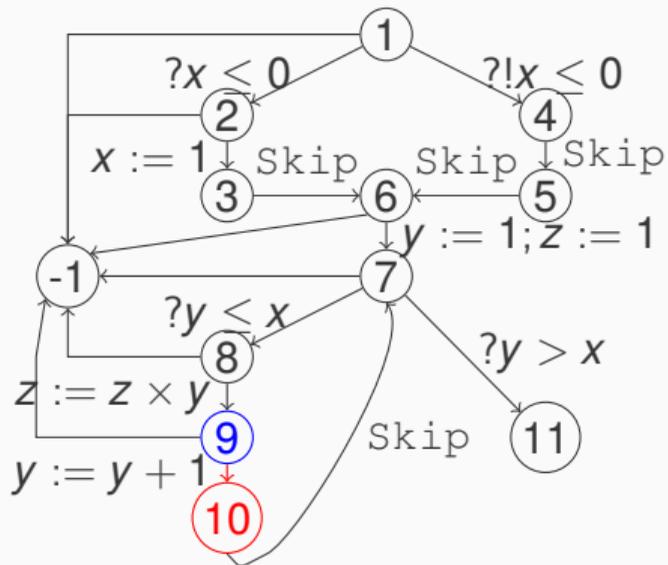
$$R(s_8) = y \leftarrow [1; 2]; z \leftarrow [1; 1]$$

$$R(s_9) = y \leftarrow [1; 2]; z \leftarrow [1; 2]$$

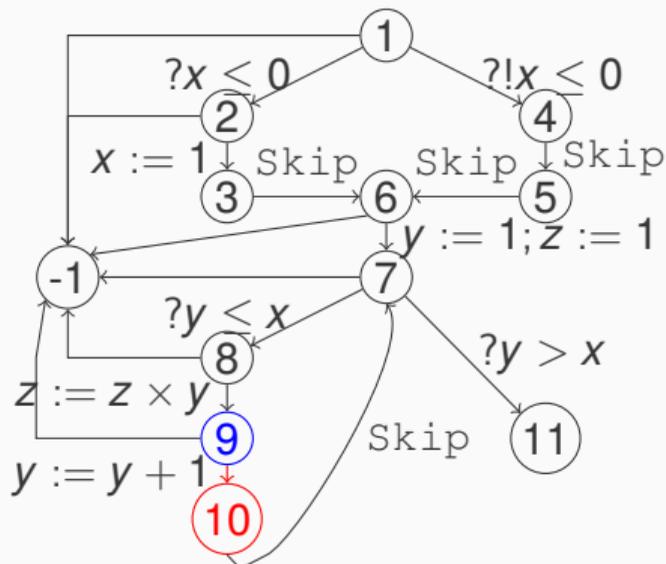
$$R(s_{10}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$$

$$R(s_{11}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$$

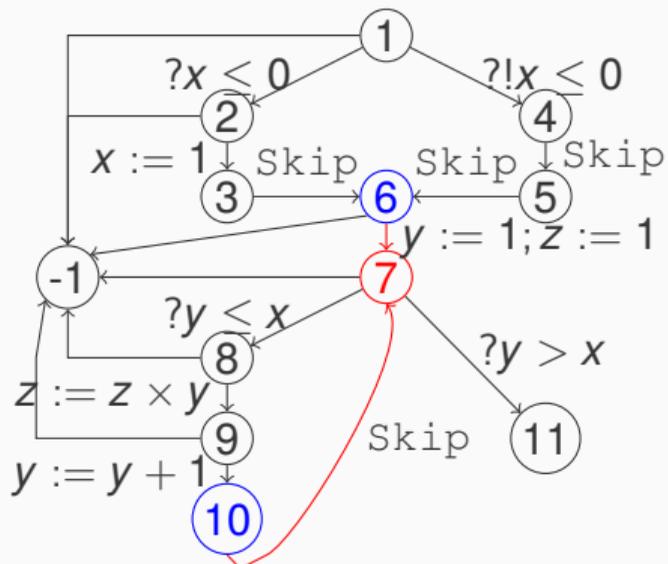
$$R(s_{-1}) = \perp^\#$$



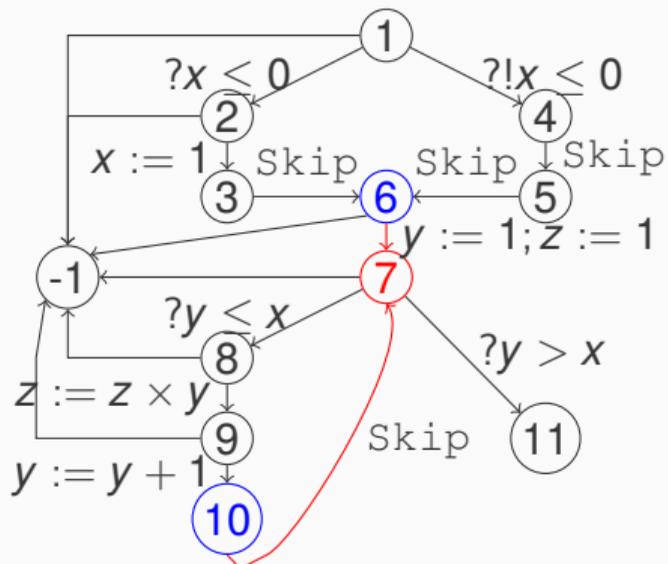
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; 2]; z \leftarrow [1; 1]$
- $R(s_8) = y \leftarrow [1; 2]; z \leftarrow [1; 1]$
- $R(s_9) = y \leftarrow [1; 2]; z \leftarrow [1; 2]$
- $R(s_{10}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{11}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{-1}) = \perp^\#$



- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; 2]; z \leftarrow [1; 1]$
- $R(s_8) = y \leftarrow [1; 2]; z \leftarrow [1; 1]$
- $R(s_9) = y \leftarrow [1; 2]; z \leftarrow [1; 2]$
- $R(s_{10}) = y \leftarrow [2; 3]; z \leftarrow [1; 2]$
- $R(s_{11}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{-1}) = \perp^\#$



- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; 2]; z \leftarrow [1; 1]$
- $R(s_8) = y \leftarrow [1; 2]; z \leftarrow [1; 1]$
- $R(s_9) = y \leftarrow [1; 2]; z \leftarrow [1; 2]$
- $R(s_{10}) = y \leftarrow [2; 3]; z \leftarrow [1; 2]$
- $R(s_{11}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{-1}) = \perp^\sharp$



$$R(s_1) = x \leftarrow [-\infty; +\infty]$$

$$R(s_2) = x \leftarrow [-\infty; 0]$$

$$R(s_3) = x \leftarrow [1; 1]$$

$$R(s_4) = x \leftarrow [1; +\infty]$$

$$R(s_5) = x \leftarrow [1; +\infty]$$

$$R(s_6) = x \leftarrow [1; +\infty]$$

$$R(s_7) = y \leftarrow [1; 3]; z \leftarrow [1; 2]$$

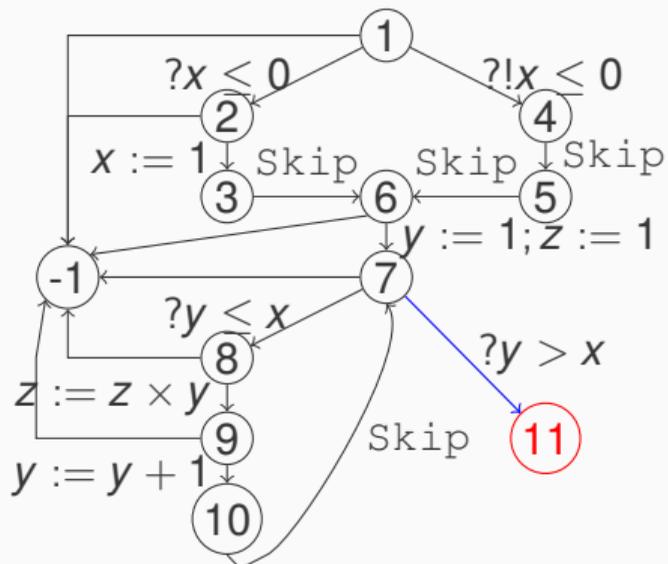
$$R(s_8) = y \leftarrow [1; 2]; z \leftarrow [1; 1]$$

$$R(s_9) = y \leftarrow [1; 2]; z \leftarrow [1; 2]$$

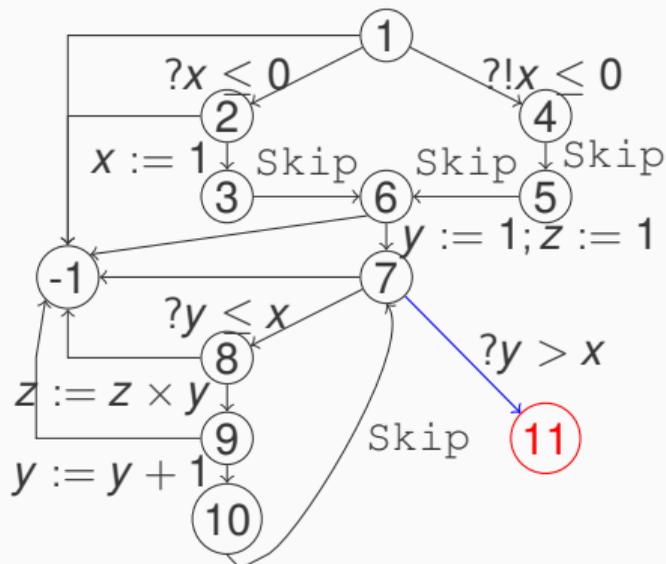
$$R(s_{10}) = y \leftarrow [2; 3]; z \leftarrow [1; 2]$$

$$R(s_{11}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$$

$$R(s_{-1}) = \perp^\sharp$$



- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; 3]; z \leftarrow [1; 2]$
- $R(s_8) = y \leftarrow [1; 2]; z \leftarrow [1; 1]$
- $R(s_9) = y \leftarrow [1; 2]; z \leftarrow [1; 2]$
- $R(s_{10}) = y \leftarrow [2; 3]; z \leftarrow [1; 2]$
- $R(s_{11}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{-1}) = \perp^\#$



$$R(s_1) = x \leftarrow [-\infty; +\infty]$$

$$R(s_2) = x \leftarrow [-\infty; 0]$$

$$R(s_3) = x \leftarrow [1; 1]$$

$$R(s_4) = x \leftarrow [1; +\infty]$$

$$R(s_5) = x \leftarrow [1; +\infty]$$

$$R(s_6) = x \leftarrow [1; +\infty]$$

$$R(s_7) = y \leftarrow [1; 3]; z \leftarrow [1; 2]$$

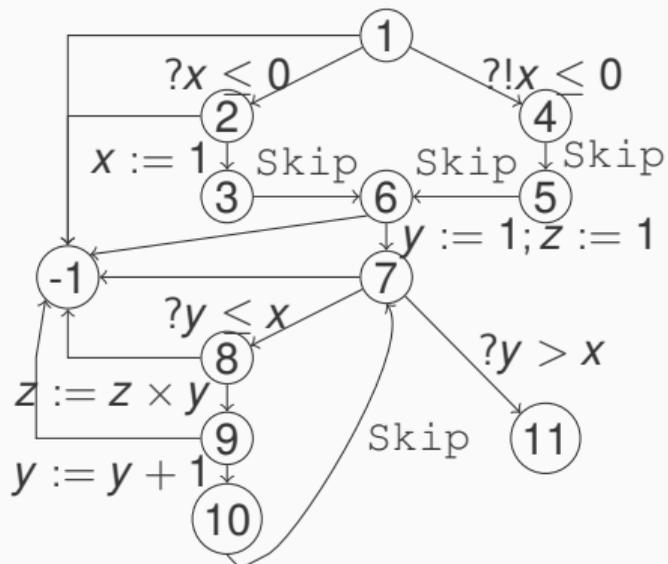
$$R(s_8) = y \leftarrow [1; 2]; z \leftarrow [1; 1]$$

$$R(s_9) = y \leftarrow [1; 2]; z \leftarrow [1; 2]$$

$$R(s_{10}) = y \leftarrow [2; 3]; z \leftarrow [1; 2]$$

$$R(s_{11}) = y \leftarrow [2, 3]; z \leftarrow [1; 2]$$

$$R(s_{-1}) = \perp^\#$$



- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_8) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_9) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_{10}) = y \leftarrow [2; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_{11}) = y \leftarrow [2; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_{-1}) = \perp^\#$

On veut obtenir une analyse qui **termine toujours**, même quand le treillis ne possède pas la propriété de chaîne ascendante.

Pour cela, on va utiliser une **sur-approximation** de la fonction de transition abstraite F^\sharp , construite de manière à garantir la terminaison.

Notre analyse consiste à calculer les itérations successives de F^\sharp en partant de \perp^\sharp .
Avec $R_n^\sharp = F^{\sharp n}(\perp^\sharp)$, on a $R_n^\sharp \sqsubseteq^\sharp R_{n+1}^\sharp$, et il vient :

$$R_{n+1}^\sharp = F^\sharp(R_n^\sharp)$$

Notre analyse consiste à calculer les itérations successives de F^\sharp en partant de \perp^\sharp .
Avec $R_n^\sharp = F^{\sharp n}(\perp^\sharp)$, on a $R_n^\sharp \sqsubseteq^\sharp R_{n+1}^\sharp$, et il vient :

$$\begin{aligned} R_{n+1}^\sharp &= F^\sharp(R_n^\sharp) \\ &= \begin{cases} R_n^\sharp & \text{si } F(R_n^\sharp) \sqsubseteq^\sharp R_n^\sharp \\ R_n^\sharp \sqcup^\sharp F^\sharp(R_n^\sharp) & \text{sinon} \end{cases} \end{aligned}$$

Notre analyse consiste à calculer les itérations successives de F^\sharp en partant de \perp^\sharp .
Avec $R_n^\sharp = F^{\sharp n}(\perp^\sharp)$, on a $R_n^\sharp \sqsubseteq^\sharp R_{n+1}^\sharp$, et il vient :

$$\begin{aligned} R_{n+1}^\sharp &= F^\sharp(R_n^\sharp) \\ &= \begin{cases} R_n^\sharp & \text{si } F(R_n^\sharp) \sqsubseteq^\sharp R_n^\sharp \\ R_n^\sharp \sqcup^\sharp F^\sharp(R_n^\sharp) & \text{sinon} \end{cases} \end{aligned}$$

Retrouver la terminaison

Dans cette présentation, on dispose d'une nouvelle possibilité pour faire une approximation du point fixe : remplacer \sqcup^\sharp par un opérateur donnant un autre majorant. Un tel opérateur s'appelle **Opérateur d'élargissement**.

Définition : opérateur d'élargissement ∇ sur un treillis (L, \sqsubseteq)

Opérateur binaire vérifiant les propriétés suivantes :

- > $\forall x, y \in L, x \sqsubseteq x \nabla y \wedge y \sqsubseteq x \nabla y$ (majorant de ses arguments)
- > Pour toute chaîne ascendante x_n , la chaîne $y_n = y_{n-1} \nabla x_n$ est stationnaire.

Théorème

Pour ∇ et une fonction monotone F , la chaîne R_n définie par

$$R_0 = \perp^\sharp$$
$$R_{n+1} = \begin{cases} R_n & \text{si } F(R_n) \sqsubseteq R_n \\ R_n \nabla F(R_n) & \text{sinon} \end{cases}$$

est stationnaire à partir d'un certain rang N . De plus $\text{lfp}(F) \sqsubseteq R_N$

Pour le treillis des intervalles, un exemple simple d'opérateur d'élargissement est le suivant :

$$\begin{aligned}\perp^\# \nabla v^\# &= v^\# \\ v^\# \nabla \perp^\# &= v^\# \\ \top^\# \nabla v^\# &= \top^\# \\ v^\# \nabla \top^\# &= \top^\#\end{aligned}$$

$$[i_{min}^1; i_{max}^1] \nabla [i_{min}^2; i_{max}^2] = [i_{min}; i_{max}] \quad \text{avec} \quad \begin{cases} i_{min} = i_{min}^1 & \text{si } i_{min}^1 \leq i_{min}^2 \\ = -\infty & \text{sinon} \\ i_{max} = i_{max}^1 & \text{si } i_{max}^1 \geq i_{max}^2 \\ = +\infty & \text{sinon} \end{cases}$$

Remarque

Un opérateur d'élargissement n'est en général **pas commutatif**.

On étend cet opérateur aux **environnements abstraits** :

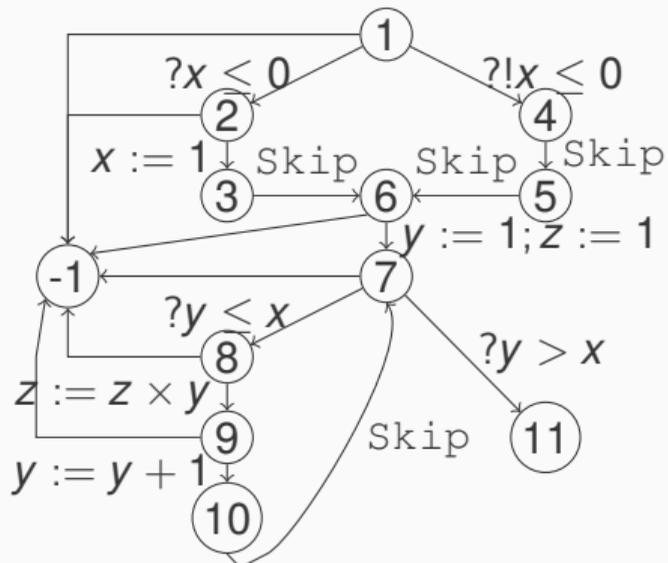
$$\rho_1^\# \nabla' \rho_2^\# = \lambda x. \rho_1^\#(x) \nabla \rho_2^\#(x)$$

On étend cet opérateur aux **environnements abstraits** :

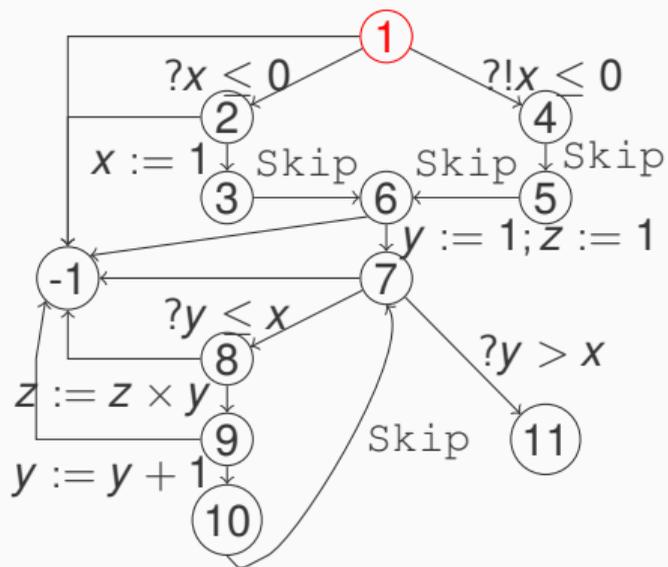
$$\rho_1^\# \nabla' \rho_2^\# = \lambda x. \rho_1^\#(x) \nabla \rho_2^\#(x)$$

puis à l'ensemble d'un **programme**. Pour cela, on peut remarquer qu'il suffit de faire l'élargissement aux **points d'entrée d'une boucle** pour préserver la terminaison. Avec $\mathcal{B} \subseteq \mathcal{S}$ les points correspondants, on a :

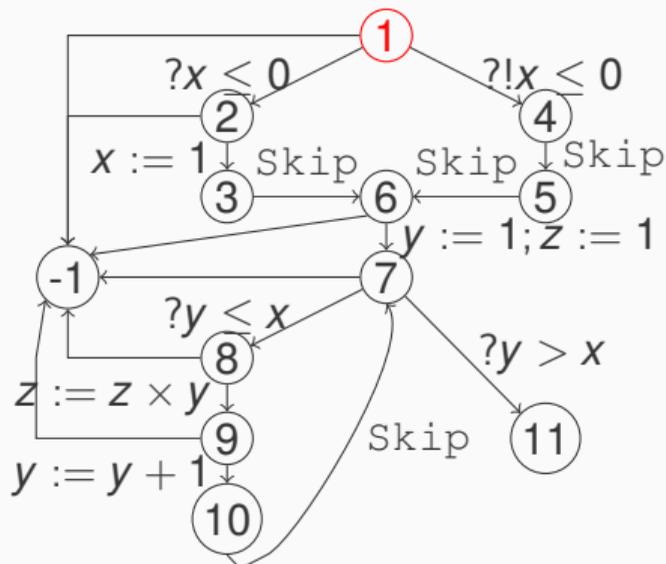
$$R_1^\# \nabla'' R_2^\# = \lambda s. \begin{cases} R_1(s) \sqcup^\# R_2(s) & \text{si } s \notin \mathcal{B} \\ R_1(s) \nabla R_2(s) & \text{si } s \in \mathcal{B} \end{cases}$$



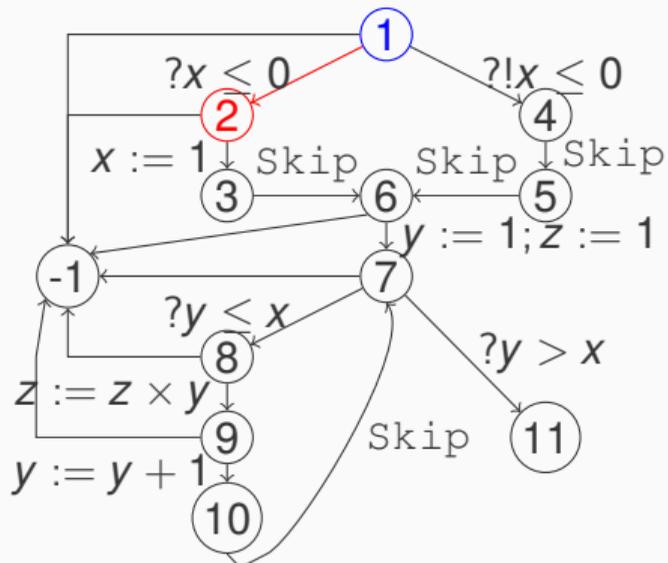
$R(s_1) = \perp \#$
 $R(s_2) = \perp \#$
 $R(s_3) = \perp \#$
 $R(s_4) = \perp \#$
 $R(s_5) = \perp \#$
 $R(s_6) = \perp \#$
 $R(s_7) = \perp \#$
 $R(s_8) = \perp \#$
 $R(s_9) = \perp \#$
 $R(s_{10}) = \perp \#$
 $R(s_{11}) = \perp \#$
 $R(s_{-1}) = \perp \#$



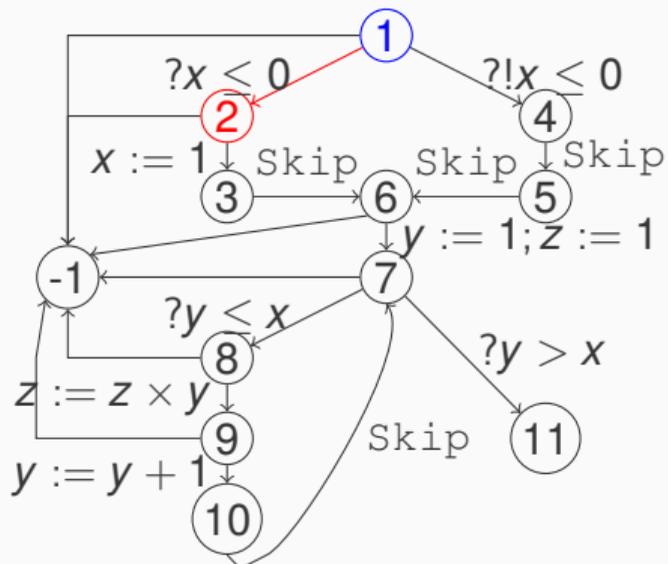
$R(s_1) = \perp \#$
 $R(s_2) = \perp \#$
 $R(s_3) = \perp \#$
 $R(s_4) = \perp \#$
 $R(s_5) = \perp \#$
 $R(s_6) = \perp \#$
 $R(s_7) = \perp \#$
 $R(s_8) = \perp \#$
 $R(s_9) = \perp \#$
 $R(s_{10}) = \perp \#$
 $R(s_{11}) = \perp \#$
 $R(s_{-1}) = \perp \#$



$R(s_1) = x \leftarrow [-\infty; +\infty]$
 $R(s_2) = \perp \#$
 $R(s_3) = \perp \#$
 $R(s_4) = \perp \#$
 $R(s_5) = \perp \#$
 $R(s_6) = \perp \#$
 $R(s_7) = \perp \#$
 $R(s_8) = \perp \#$
 $R(s_9) = \perp \#$
 $R(s_{10}) = \perp \#$
 $R(s_{11}) = \perp \#$
 $R(s_{-1}) = \perp \#$



$R(s_1) = x \leftarrow [-\infty; +\infty]$
 $R(s_2) = \perp \#$
 $R(s_3) = \perp \#$
 $R(s_4) = \perp \#$
 $R(s_5) = \perp \#$
 $R(s_6) = \perp \#$
 $R(s_7) = \perp \#$
 $R(s_8) = \perp \#$
 $R(s_9) = \perp \#$
 $R(s_{10}) = \perp \#$
 $R(s_{11}) = \perp \#$
 $R(s_{-1}) = \perp \#$



$$R(s_1) = x \leftarrow [-\infty; +\infty]$$

$$R(s_2) = x \leftarrow [-\infty; 0]$$

$$R(s_3) = \perp \#$$

$$R(s_4) = \perp \#$$

$$R(s_5) = \perp \#$$

$$R(s_6) = \perp \#$$

$$R(s_7) = \perp \#$$

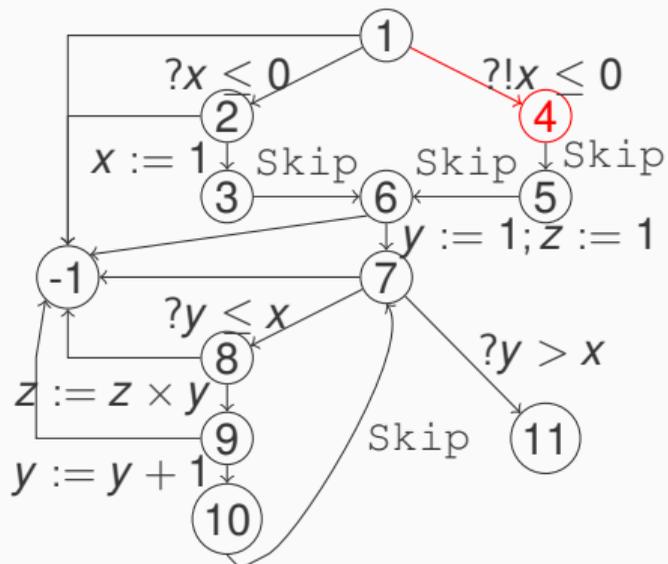
$$R(s_8) = \perp \#$$

$$R(s_9) = \perp \#$$

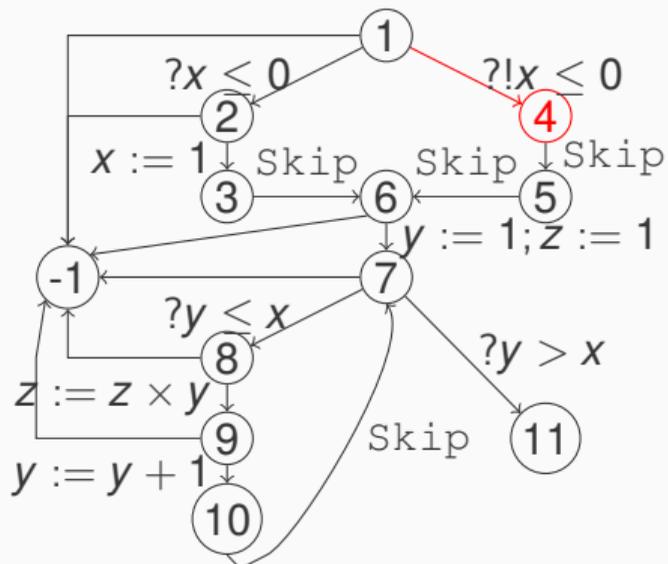
$$R(s_{10}) = \perp \#$$

$$R(s_{11}) = \perp \#$$

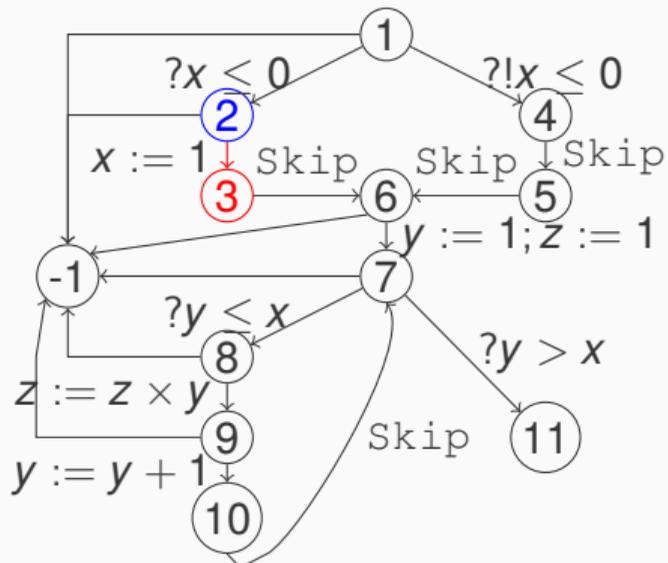
$$R(s_{-1}) = \perp \#$$



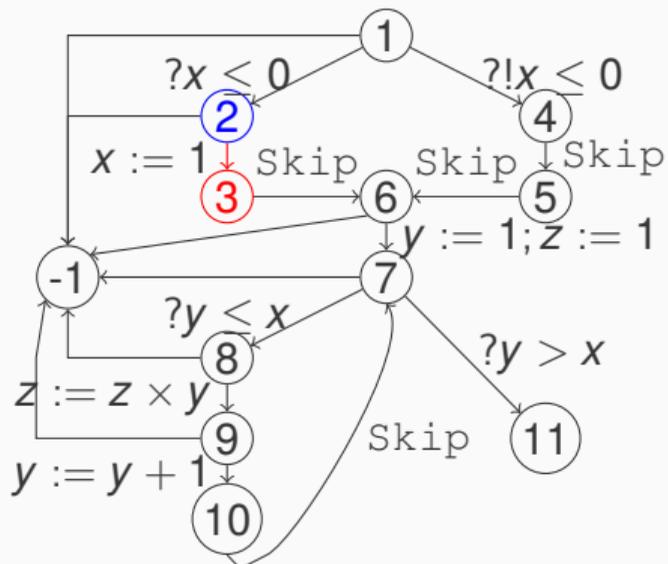
$R(s_1) = x \leftarrow [-\infty; +\infty]$
 $R(s_2) = x \leftarrow [-\infty; 0]$
 $R(s_3) = \perp \#$
 $R(s_4) = \perp \#$
 $R(s_5) = \perp \#$
 $R(s_6) = \perp \#$
 $R(s_7) = \perp \#$
 $R(s_8) = \perp \#$
 $R(s_9) = \perp \#$
 $R(s_{10}) = \perp \#$
 $R(s_{11}) = \perp \#$
 $R(s_{-1}) = \perp \#$



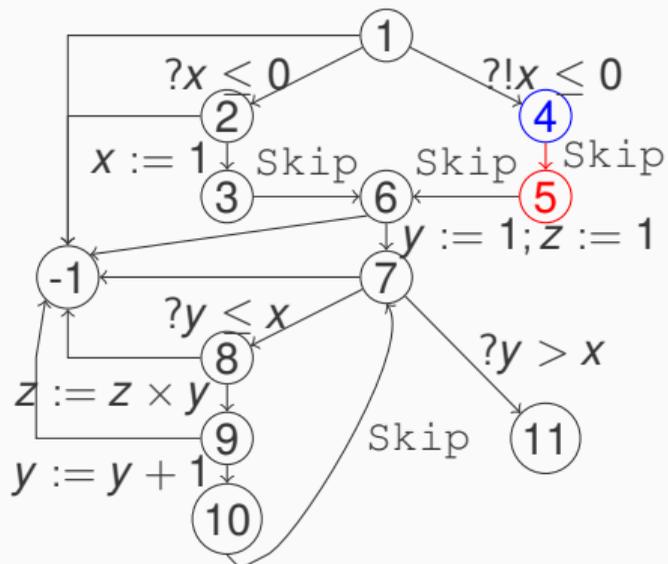
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = \perp \#$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = \perp \#$
- $R(s_6) = \perp \#$
- $R(s_7) = \perp \#$
- $R(s_8) = \perp \#$
- $R(s_9) = \perp \#$
- $R(s_{10}) = \perp \#$
- $R(s_{11}) = \perp \#$
- $R(s_{-1}) = \perp \#$



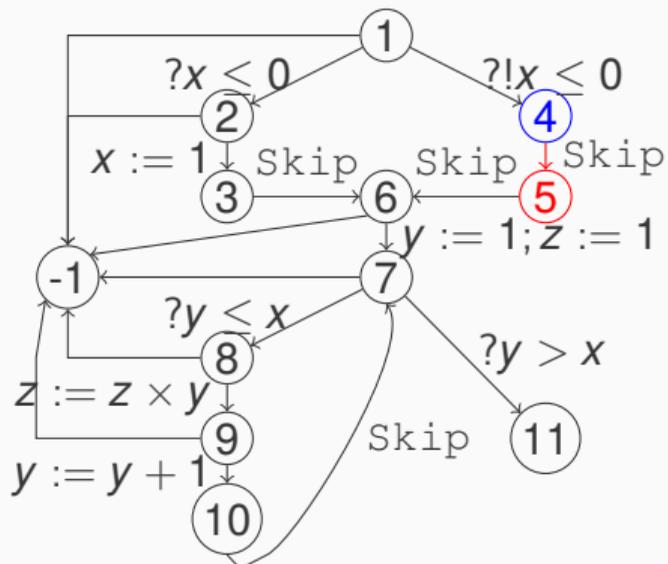
$R(s_1) = x \leftarrow [-\infty; +\infty]$
 $R(s_2) = x \leftarrow [-\infty; 0]$
 $R(s_3) = \perp \#$
 $R(s_4) = x \leftarrow [1; +\infty]$
 $R(s_5) = \perp \#$
 $R(s_6) = \perp \#$
 $R(s_7) = \perp \#$
 $R(s_8) = \perp \#$
 $R(s_9) = \perp \#$
 $R(s_{10}) = \perp \#$
 $R(s_{11}) = \perp \#$
 $R(s_{-1}) = \perp \#$



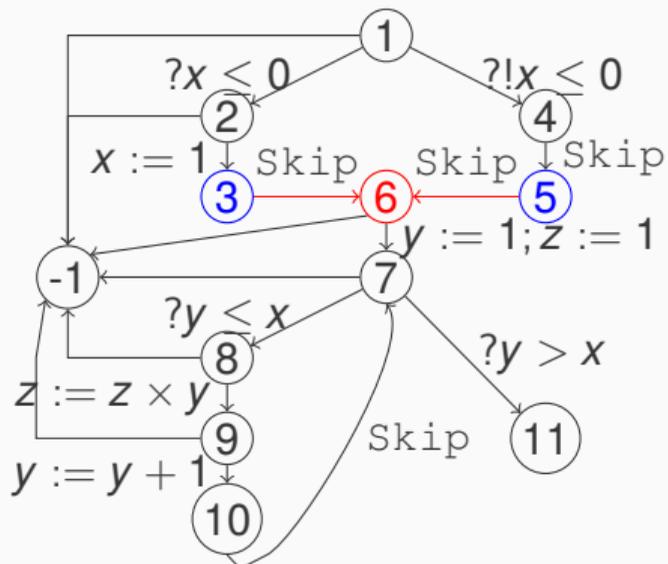
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = \perp \#$
- $R(s_6) = \perp \#$
- $R(s_7) = \perp \#$
- $R(s_8) = \perp \#$
- $R(s_9) = \perp \#$
- $R(s_{10}) = \perp \#$
- $R(s_{11}) = \perp \#$
- $R(s_{-1}) = \perp \#$



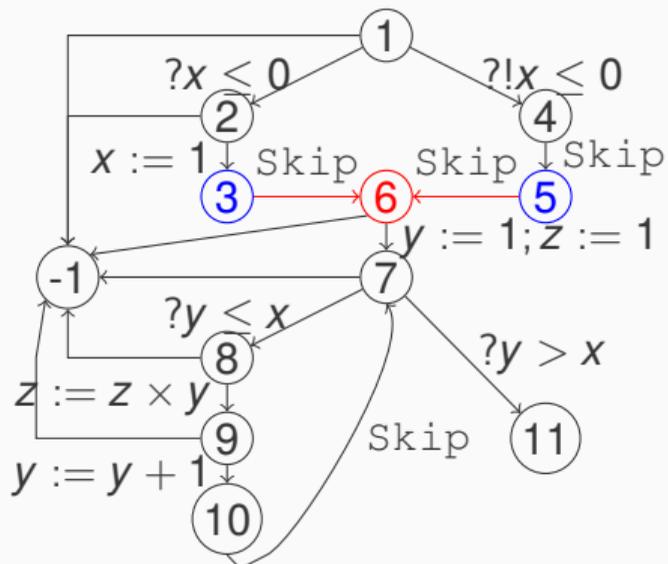
$R(s_1) = x \leftarrow [-\infty; +\infty]$
 $R(s_2) = x \leftarrow [-\infty; 0]$
 $R(s_3) = x \leftarrow [1; 1]$
 $R(s_4) = x \leftarrow [1; +\infty]$
 $R(s_5) = \perp \#$
 $R(s_6) = \perp \#$
 $R(s_7) = \perp \#$
 $R(s_8) = \perp \#$
 $R(s_9) = \perp \#$
 $R(s_{10}) = \perp \#$
 $R(s_{11}) = \perp \#$
 $R(s_{-1}) = \perp \#$



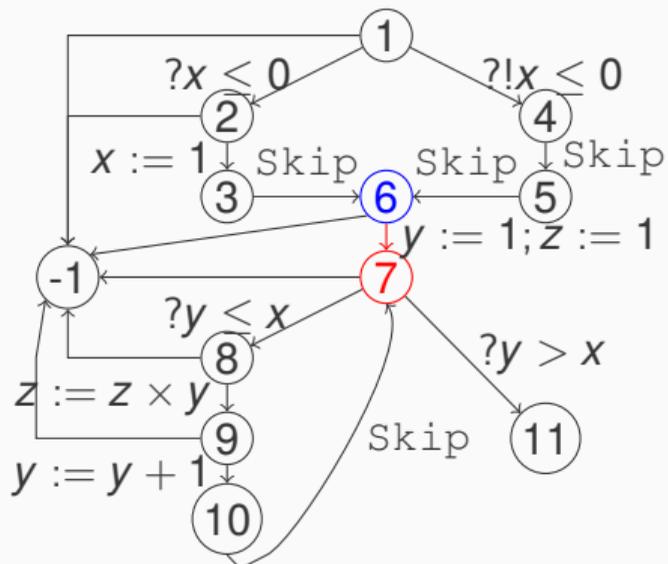
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = \perp \#$
- $R(s_7) = \perp \#$
- $R(s_8) = \perp \#$
- $R(s_9) = \perp \#$
- $R(s_{10}) = \perp \#$
- $R(s_{11}) = \perp \#$
- $R(s_{-1}) = \perp \#$



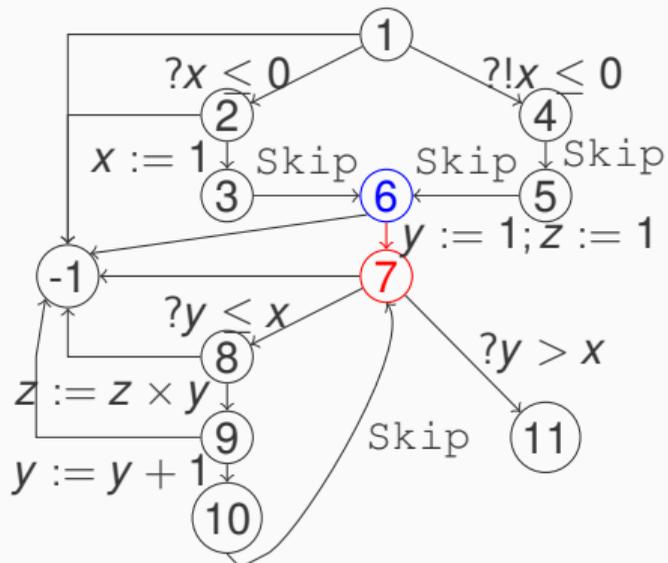
$R(s_1) = x \leftarrow [-\infty; +\infty]$
 $R(s_2) = x \leftarrow [-\infty; 0]$
 $R(s_3) = x \leftarrow [1; 1]$
 $R(s_4) = x \leftarrow [1; +\infty]$
 $R(s_5) = x \leftarrow [1; +\infty]$
 $R(s_6) = \perp \#$
 $R(s_7) = \perp \#$
 $R(s_8) = \perp \#$
 $R(s_9) = \perp \#$
 $R(s_{10}) = \perp \#$
 $R(s_{11}) = \perp \#$
 $R(s_{-1}) = \perp \#$



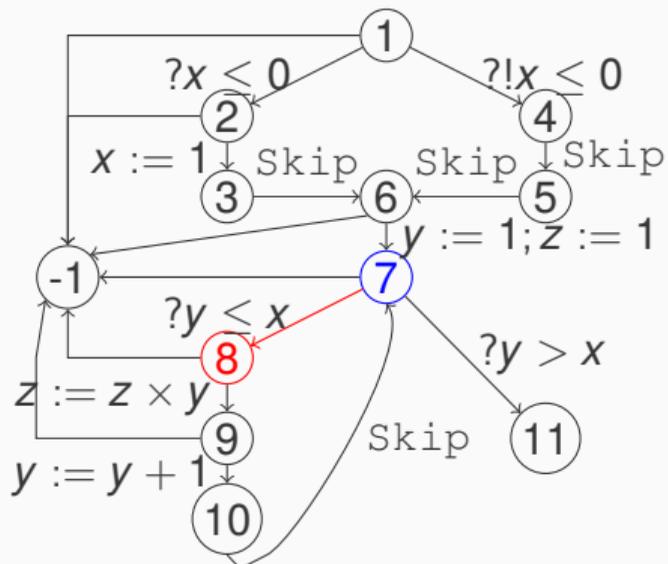
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = \perp^\#$
- $R(s_8) = \perp^\#$
- $R(s_9) = \perp^\#$
- $R(s_{10}) = \perp^\#$
- $R(s_{11}) = \perp^\#$
- $R(s_{-1}) = \perp^\#$



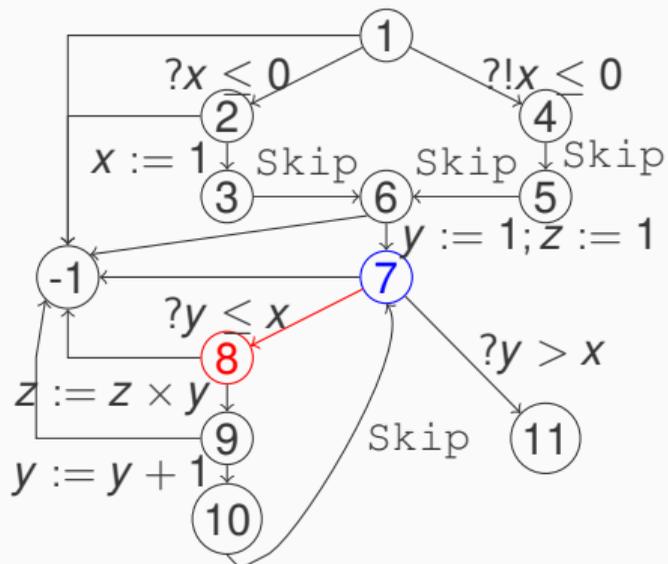
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = \perp^\sharp$
- $R(s_8) = \perp^\sharp$
- $R(s_9) = \perp^\sharp$
- $R(s_{10}) = \perp^\sharp$
- $R(s_{11}) = \perp^\sharp$
- $R(s_{-1}) = \perp^\sharp$



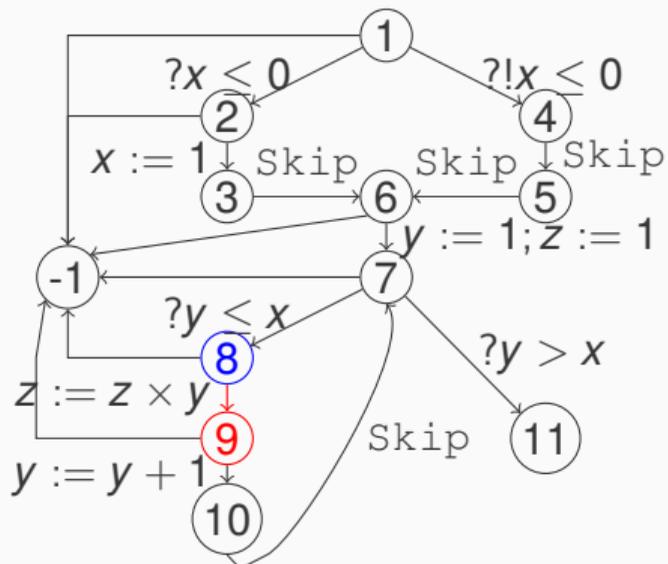
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_8) = \perp^\#$
- $R(s_9) = \perp^\#$
- $R(s_{10}) = \perp^\#$
- $R(s_{11}) = \perp^\#$
- $R(s_{-1}) = \perp^\#$



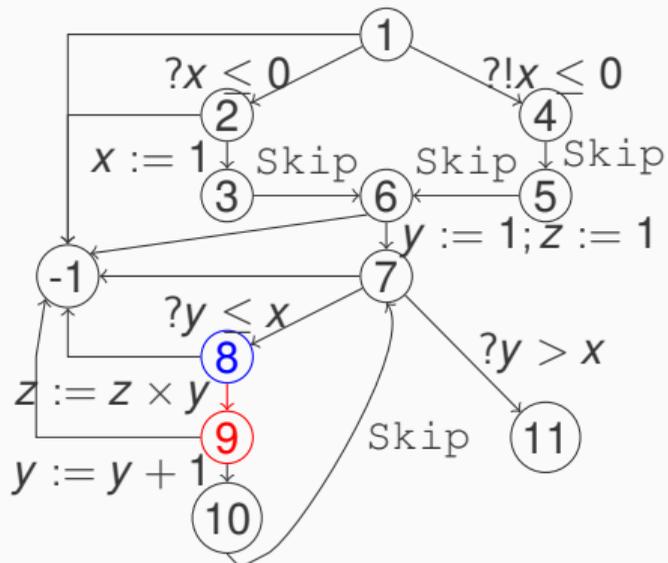
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_8) = \perp^\#$
- $R(s_9) = \perp^\#$
- $R(s_{10}) = \perp^\#$
- $R(s_{11}) = \perp^\#$
- $R(s_{-1}) = \perp^\#$



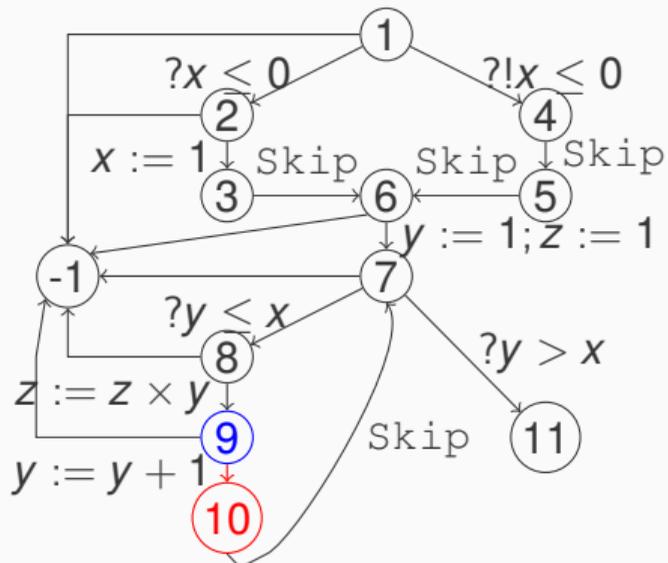
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_8) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_9) = \perp^\sharp$
- $R(s_{10}) = \perp^\sharp$
- $R(s_{11}) = \perp^\sharp$
- $R(s_{-1}) = \perp^\sharp$



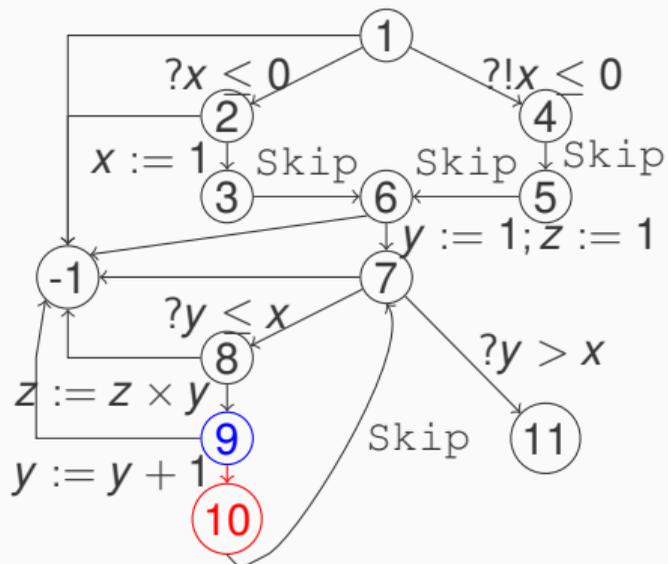
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_8) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_9) = \perp^\#$
- $R(s_{10}) = \perp^\#$
- $R(s_{11}) = \perp^\#$
- $R(s_{-1}) = \perp^\#$



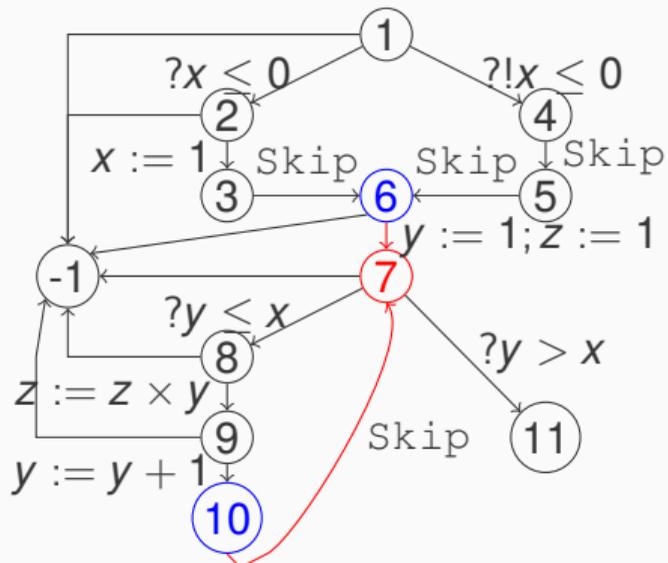
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_8) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_9) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_{10}) = \perp^\#$
- $R(s_{11}) = \perp^\#$
- $R(s_{-1}) = \perp^\#$



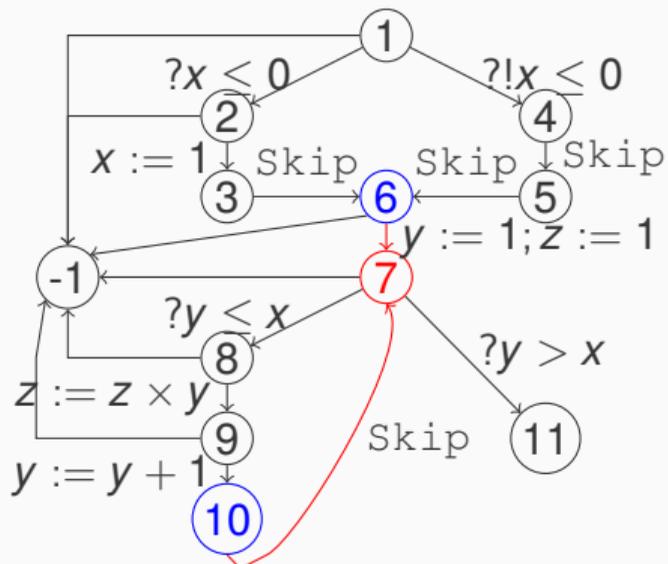
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_8) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_9) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_{10}) = \perp^\#$
- $R(s_{11}) = \perp^\#$
- $R(s_{-1}) = \perp^\#$



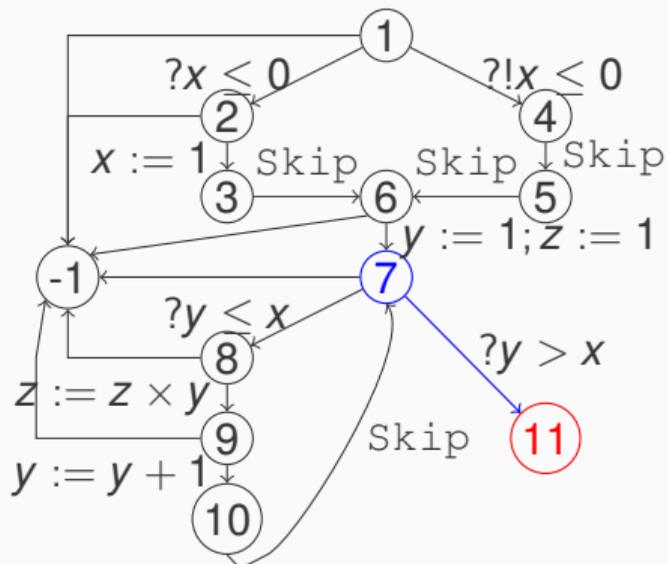
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_8) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_9) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_{10}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{11}) = \perp^\#$
- $R(s_{-1}) = \perp^\#$



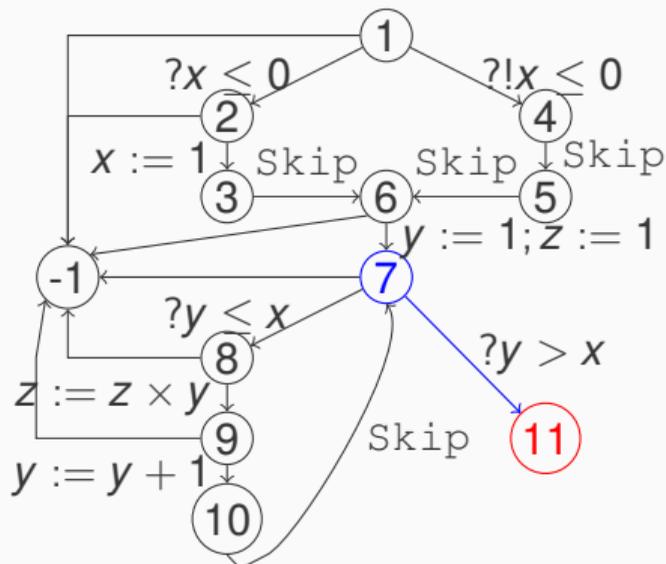
$R(s_1) = x \leftarrow [-\infty; +\infty]$
 $R(s_2) = x \leftarrow [-\infty; 0]$
 $R(s_3) = x \leftarrow [1; 1]$
 $R(s_4) = x \leftarrow [1; +\infty]$
 $R(s_5) = x \leftarrow [1; +\infty]$
 $R(s_6) = x \leftarrow [1; +\infty]$
 $R(s_7) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
 $R(s_8) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
 $R(s_9) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
 $R(s_{10}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
 $R(s_{11}) = \perp^\#$
 $R(s_{-1}) = \perp^\#$



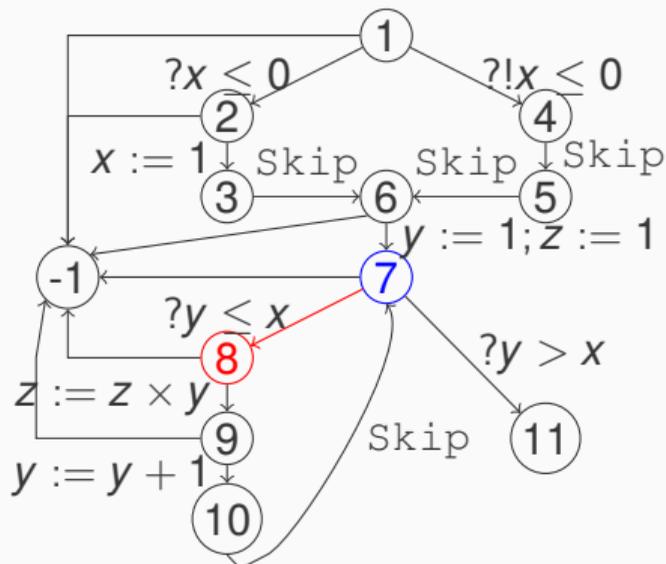
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; +\infty]; z \leftarrow [1; 1]$
- $R(s_8) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_9) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_{10}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{11}) = \perp^\#$
- $R(s_{-1}) = \perp^\#$



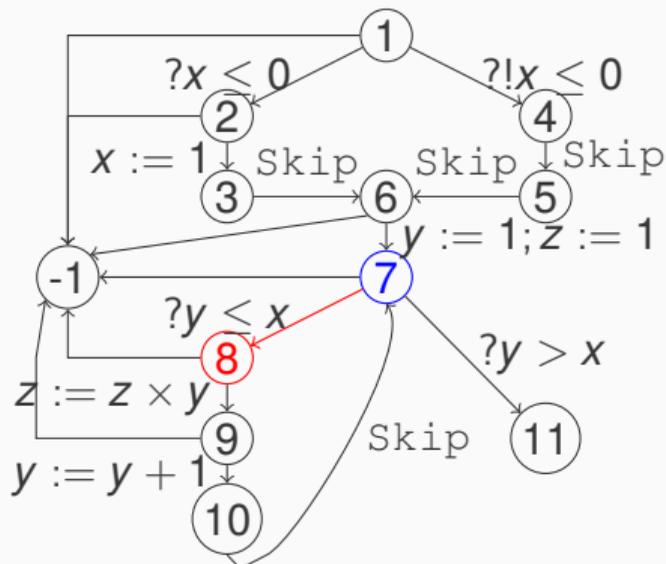
$R(s_1) = x \leftarrow [-\infty; +\infty]$
 $R(s_2) = x \leftarrow [-\infty; 0]$
 $R(s_3) = x \leftarrow [1; 1]$
 $R(s_4) = x \leftarrow [1; +\infty]$
 $R(s_5) = x \leftarrow [1; +\infty]$
 $R(s_6) = x \leftarrow [1; +\infty]$
 $R(s_7) = y \leftarrow [1; +\infty]; z \leftarrow [1; 1]$
 $R(s_8) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
 $R(s_9) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
 $R(s_{10}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
 $R(s_{11}) = \perp^\#$
 $R(s_{-1}) = \perp^\#$



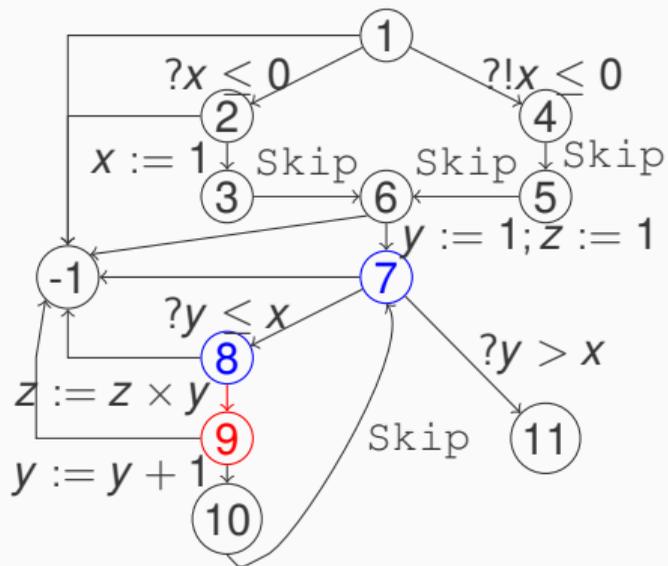
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; +\infty]; z \leftarrow [1; 1]$
- $R(s_8) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_9) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_{10}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{11}) = y \leftarrow [2; +\infty]; z \leftarrow [1; 1]$
- $R(s_{-1}) = \perp^\#$



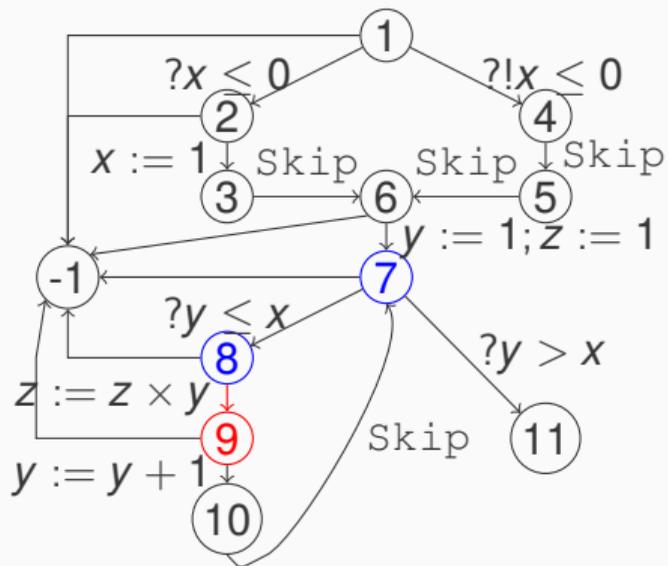
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; +\infty]; z \leftarrow [1; 1]$
- $R(s_8) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_9) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_{10}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{11}) = y \leftarrow [2; +\infty]; z \leftarrow [1; 1]$
- $R(s_{-1}) = \perp^\#$



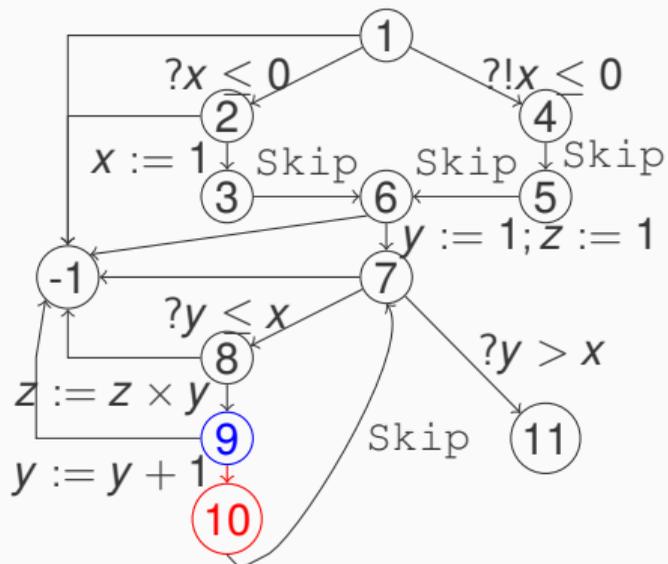
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; +\infty]; z \leftarrow [1; 1]$
- $R(s_8) = y \leftarrow [1; +\infty]; z \leftarrow [1; 1]$
- $R(s_9) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
- $R(s_{10}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{11}) = y \leftarrow [2; +\infty]; z \leftarrow [1; 1]$
- $R(s_{-1}) = \perp^\#$



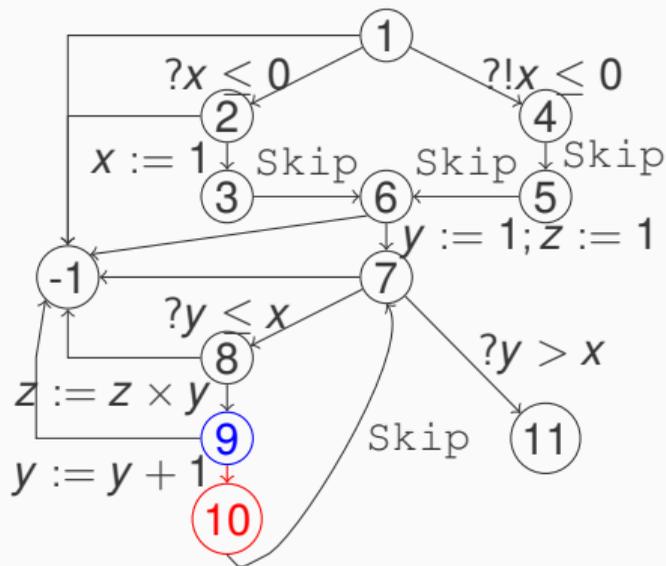
$R(s_1) = x \leftarrow [-\infty; +\infty]$
 $R(s_2) = x \leftarrow [-\infty; 0]$
 $R(s_3) = x \leftarrow [1; 1]$
 $R(s_4) = x \leftarrow [1; +\infty]$
 $R(s_5) = x \leftarrow [1; +\infty]$
 $R(s_6) = x \leftarrow [1; +\infty]$
 $R(s_7) = y \leftarrow [1; +\infty]; z \leftarrow [1; 1]$
 $R(s_8) = y \leftarrow [1; +\infty]; z \leftarrow [1; 1]$
 $R(s_9) = y \leftarrow [1; 1]; z \leftarrow [1; 1]$
 $R(s_{10}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
 $R(s_{11}) = y \leftarrow [2; +\infty]; z \leftarrow [1; 1]$
 $R(s_{-1}) = \perp^\#$



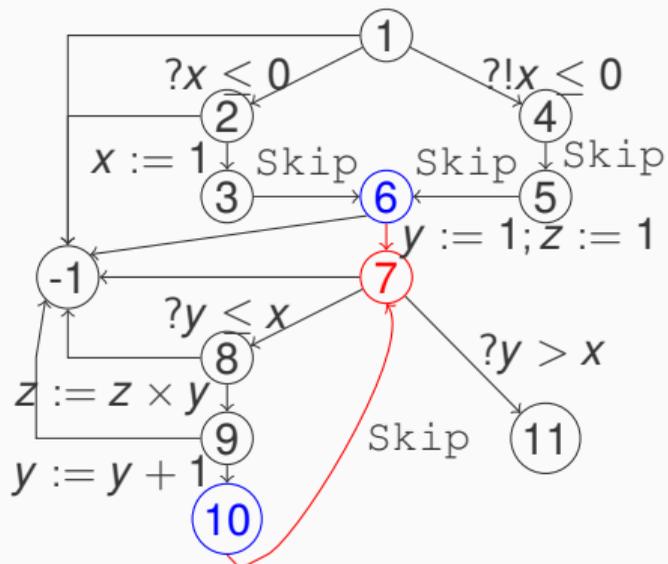
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; +\infty]; z \leftarrow [1; 1]$
- $R(s_8) = y \leftarrow [1; +\infty]; z \leftarrow [1; 1]$
- $R(s_9) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_{10}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{11}) = y \leftarrow [2; +\infty]; z \leftarrow [1; 1]$
- $R(s_{-1}) = \perp^\#$



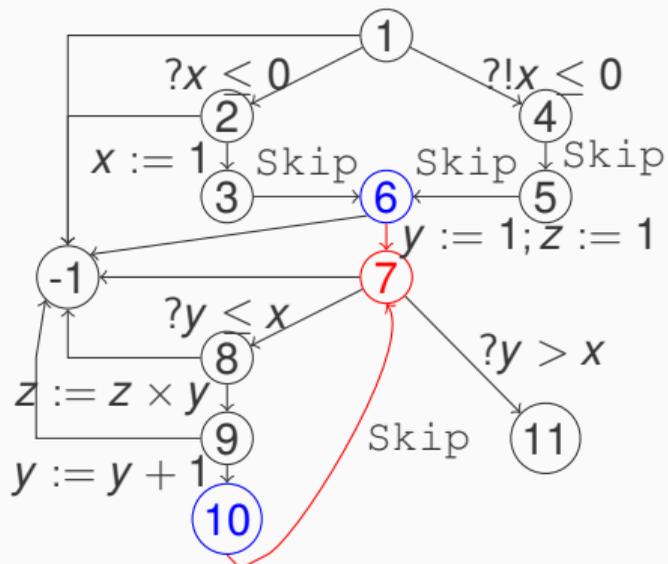
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; +\infty]; z \leftarrow [1; 1]$
- $R(s_8) = y \leftarrow [1; +\infty]; z \leftarrow [1; 1]$
- $R(s_9) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_{10}) = y \leftarrow [2; 2]; z \leftarrow [1; 1]$
- $R(s_{11}) = y \leftarrow [2; +\infty]; z \leftarrow [1; 1]$
- $R(s_{-1}) = \perp^\#$



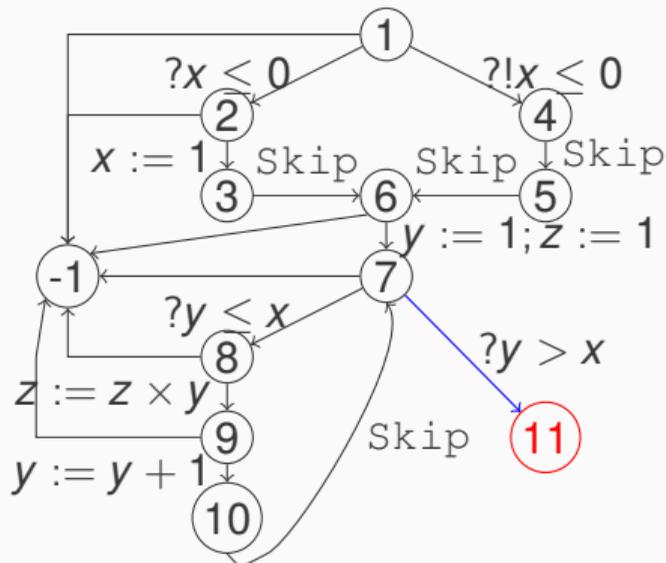
$R(s_1) = x \leftarrow [-\infty; +\infty]$
 $R(s_2) = x \leftarrow [-\infty; 0]$
 $R(s_3) = x \leftarrow [1; 1]$
 $R(s_4) = x \leftarrow [1; +\infty]$
 $R(s_5) = x \leftarrow [1; +\infty]$
 $R(s_6) = x \leftarrow [1; +\infty]$
 $R(s_7) = y \leftarrow [1; +\infty]; z \leftarrow [1; 1]$
 $R(s_8) = y \leftarrow [1; +\infty]; z \leftarrow [1; 1]$
 $R(s_9) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
 $R(s_{10}) = y \leftarrow [2; +\infty]; z \leftarrow [1; +\infty]$
 $R(s_{11}) = y \leftarrow [2; +\infty]; z \leftarrow [1; 1]$
 $R(s_{-1}) = \perp^\#$



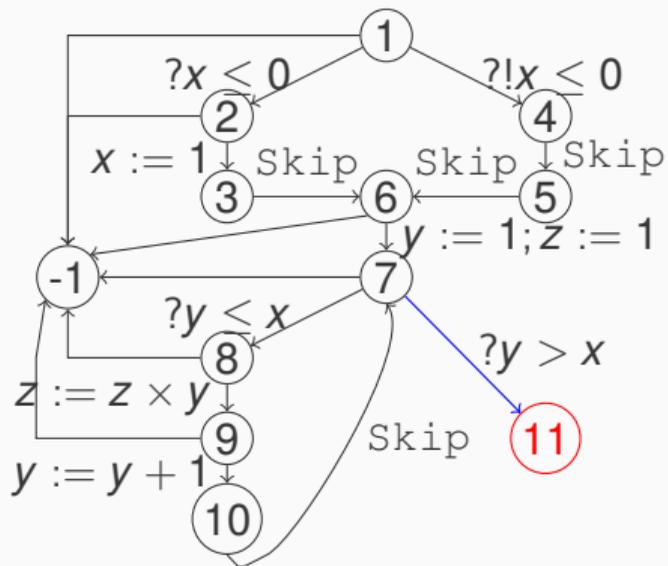
$R(s_1) = x \leftarrow [-\infty; +\infty]$
 $R(s_2) = x \leftarrow [-\infty; 0]$
 $R(s_3) = x \leftarrow [1; 1]$
 $R(s_4) = x \leftarrow [1; +\infty]$
 $R(s_5) = x \leftarrow [1; +\infty]$
 $R(s_6) = x \leftarrow [1; +\infty]$
 $R(s_7) = y \leftarrow [1; +\infty]; z \leftarrow [1; 1]$
 $R(s_8) = y \leftarrow [1; +\infty]; z \leftarrow [1; 1]$
 $R(s_9) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
 $R(s_{10}) = y \leftarrow [2; +\infty]; z \leftarrow [1; +\infty]$
 $R(s_{11}) = y \leftarrow [2; +\infty]; z \leftarrow [1; 1]$
 $R(s_{-1}) = \perp^\#$



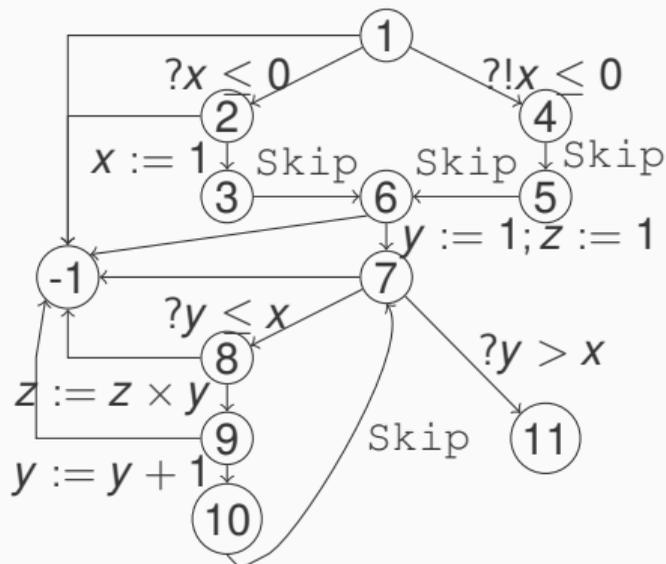
$R(s_1) = x \leftarrow [-\infty; +\infty]$
 $R(s_2) = x \leftarrow [-\infty; 0]$
 $R(s_3) = x \leftarrow [1; 1]$
 $R(s_4) = x \leftarrow [1; +\infty]$
 $R(s_5) = x \leftarrow [1; +\infty]$
 $R(s_6) = x \leftarrow [1; +\infty]$
 $R(s_7) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
 $R(s_8) = y \leftarrow [1; +\infty]; z \leftarrow [1; 1]$
 $R(s_9) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
 $R(s_{10}) = y \leftarrow [2; +\infty]; z \leftarrow [1; +\infty]$
 $R(s_{11}) = y \leftarrow [2; +\infty]; z \leftarrow [1; 1]$
 $R(s_{-1}) = \perp^\#$



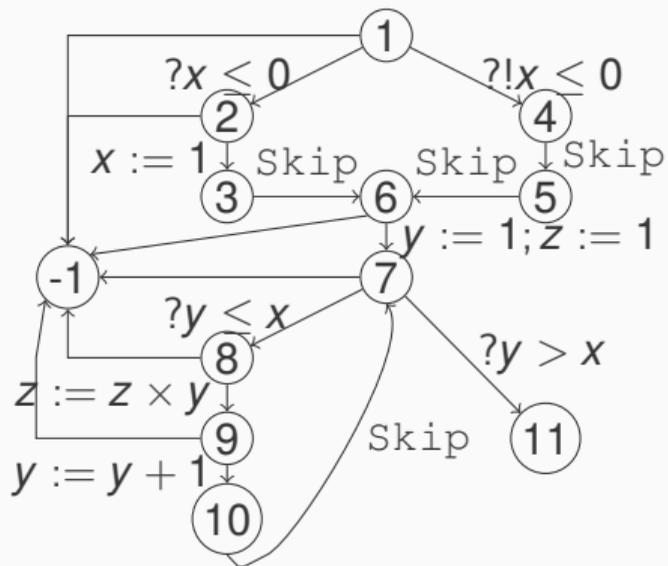
$R(s_1) = x \leftarrow [-\infty; +\infty]$
 $R(s_2) = x \leftarrow [-\infty; 0]$
 $R(s_3) = x \leftarrow [1; 1]$
 $R(s_4) = x \leftarrow [1; +\infty]$
 $R(s_5) = x \leftarrow [1; +\infty]$
 $R(s_6) = x \leftarrow [1; +\infty]$
 $R(s_7) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
 $R(s_8) = y \leftarrow [1; +\infty]; z \leftarrow [1; 1]$
 $R(s_9) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
 $R(s_{10}) = y \leftarrow [2; +\infty]; z \leftarrow [1; +\infty]$
 $R(s_{11}) = y \leftarrow [2; +\infty]; z \leftarrow [1; 1]$
 $R(s_{-1}) = \perp^\#$



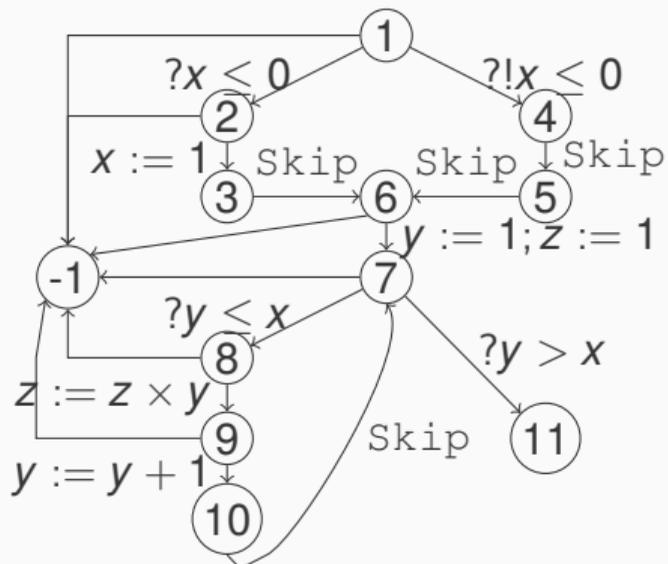
- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_8) = y \leftarrow [1; +\infty]; z \leftarrow [1; 1]$
- $R(s_9) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_{10}) = y \leftarrow [2; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_{11}) = y \leftarrow [2; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_{-1}) = \perp^\#$



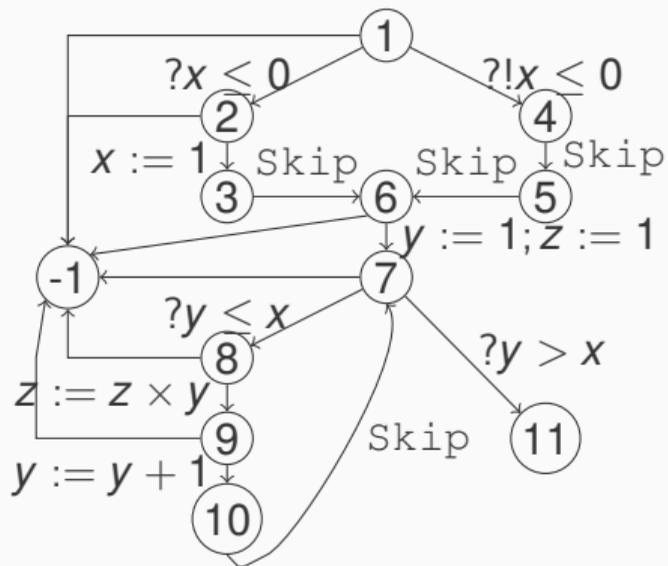
$R(s_1) = x \leftarrow [-\infty; +\infty]$
 $R(s_2) = x \leftarrow [-\infty; 0]$
 $R(s_3) = x \leftarrow [1; 1]$
 $R(s_4) = x \leftarrow [1; +\infty]$
 $R(s_5) = x \leftarrow [1; +\infty]$
 $R(s_6) = x \leftarrow [1; +\infty]$
 $R(s_7) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
 $R(s_8) = y \leftarrow [1; +\infty]; z \leftarrow [1; 1]$
 $R(s_9) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
 $R(s_{10}) = y \leftarrow [2; +\infty]; z \leftarrow [1; +\infty]$
 $R(s_{11}) = y \leftarrow [2; +\infty]; z \leftarrow [1; +\infty]$
 $R(s_{-1}) = \perp^\#$



- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_8) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_9) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_{10}) = y \leftarrow [2; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_{11}) = y \leftarrow [2; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_{-1}) = \perp^\#$



- $R(s_1) = x \leftarrow [-\infty; +\infty]$
- $R(s_2) = x \leftarrow [-\infty; 0]$
- $R(s_3) = x \leftarrow [1; 1]$
- $R(s_4) = x \leftarrow [1; +\infty]$
- $R(s_5) = x \leftarrow [1; +\infty]$
- $R(s_6) = x \leftarrow [1; +\infty]$
- $R(s_7) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_8) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_9) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_{10}) = y \leftarrow [2; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_{11}) = y \leftarrow [2; +\infty]; z \leftarrow [1; +\infty]$
- $R(s_{-1}) = \perp^\#$



$R(s_1) = x \leftarrow [-\infty; +\infty]$
 $R(s_2) = x \leftarrow [-\infty; 0]$
 $R(s_3) = x \leftarrow [1; 1]$
 $R(s_4) = x \leftarrow [1; +\infty]$
 $R(s_5) = x \leftarrow [1; +\infty]$
 $R(s_6) = x \leftarrow [1; +\infty]$
 $R(s_7) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
 $R(s_8) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
 $R(s_9) = y \leftarrow [1; +\infty]; z \leftarrow [1; +\infty]$
 $R(s_{10}) = y \leftarrow [2; +\infty]; z \leftarrow [1; +\infty]$
 $R(s_{11}) = y \leftarrow [2; +\infty]; z \leftarrow [1; +\infty]$
 $R(s_{-1}) = \perp^\#$

Ouvrages

- > Hanne Nielson, Flemming Nielson, et Chris Hankin. *Principles of Program Analysis*. Springer Verlag 1999
- > Neil Jones et Flemming Nielson. *Abstract Interpretation : a Semantics-Based Tool for Program Analysis*. Handbook of Logic in Computer Science (vol. 4) 1994

Fondations

- > Patrick et Radhia Cousot. *Abstract Interpretation : a Unified Lattice Model for Static Analysis of Programs by Construction or Approximation of Fixpoints*. PoPL'77

Partie VI

Méthode déductive

Comment prouver la correction ?

Quelles garanties ?

Procédure

- > **raisonne par déduction** sur le programme de manière automatique
- > prend en compte les **spécifications formelles**

Obligation de preuve

- > extraire les propriétés à vérifier (appelées **obligations de preuve** - OP)
- > la vérification des OP est laissée à la charge de l'utilisateur
- > **l'utilisation de prouveurs automatiques** ou semi-automatiques est indispensable

Garantie

Toutes les OP vérifiées = **programme correct** vis-à-vis de sa spécification

- > Un **triplet de Hoare** est un triplet noté $\{P\}i\{Q\}$ où P et Q sont des propriétés (formules logiques) et i est une instruction du programme.
- > P est appelé la **précondition** de i
- > Q est appelé la **postcondition** de i
- > Informellement, un triplet de Hoare est **valide** si la propriété suivante est vraie :
si la **précondition** P est valide avant d'exécuter l'instruction i , alors la **postcondition** Q sera valide après l'exécution de i

Les triplets de Hoare suivants sont-ils valides ?

> $\{x = 10\}x := x + 1\{x = 11\}$?

Les triplets de Hoare suivants sont-ils valides ?

> $\{x = 10\}x := x + 1\{x = 11\}$? ✓

> $\{c = 0\}\text{if } c \text{ then } x := c \text{ else } x = 1 \text{ fi}\{x = 0 \wedge c = 0\}$?

Les triplets de Hoare suivants sont-ils valides ?

- > $\{x = 10\}x := x + 1\{x = 11\}$? ✓
- > $\{c = 0\}\text{if } c \text{ then } x := c \text{ else } x = 1 \text{ fi}\{x = 0 \wedge c = 0\}$? ✗
- > $\{c = 2\}\text{if } c \text{ then } x := c \text{ else } x = 1 \text{ fi}\{x = 2 \wedge c = 2\}$?

Les triplets de Hoare suivants sont-ils valides ?

- > $\{x = 10\}x := x + 1\{x = 11\}$? ✓
- > $\{c = 0\}\text{if } c \text{ then } x := c \text{ else } x = 1 \text{ fi}\{x = 0 \wedge c = 0\}$? ✗
- > $\{c = 2\}\text{if } c \text{ then } x := c \text{ else } x = 1 \text{ fi}\{x = 2 \wedge c = 2\}$? ✓
- > $\{\text{false}\}x := 0; y := 1\{x = 0 \wedge y = 2\}$?

Les triplets de Hoare suivants sont-ils valides ?

- > $\{x = 10\}x := x + 1\{x = 11\}$? ✓
- > $\{c = 0\}\text{if } c \text{ then } x := c \text{ else } x = 1 \text{ fi}\{x = 0 \wedge c = 0\}$? ✗
- > $\{c = 2\}\text{if } c \text{ then } x := c \text{ else } x = 1 \text{ fi}\{x = 2 \wedge c = 2\}$? ✓
- > $\{\text{false}\}x := 0; y := 1\{x = 0 \wedge y = 2\}$? ✓
- > $\{i = 1\}\text{while } i \leq 10 \text{ do } i := i + 1 \text{ end}\{i = 10\}$?

Les triplets de Hoare suivants sont-ils valides ?

- > $\{x = 10\}x := x + 1\{x = 11\}$? ✓
- > $\{c = 0\}\text{if } c \text{ then } x := c \text{ else } x = 1 \text{ fi}\{x = 0 \wedge c = 0\}$? ✗
- > $\{c = 2\}\text{if } c \text{ then } x := c \text{ else } x = 1 \text{ fi}\{x = 2 \wedge c = 2\}$? ✓
- > $\{\text{false}\}x := 0; y := 1\{x = 0 \wedge y = 2\}$? ✓
- > $\{i = 1\}\text{while } i \leq 10 \text{ do } i := i + 1 \text{ end}\{i = 10\}$? ✗
- > $\{i = 1\}\text{while } !(i \leq 0) \text{ do } i := i + 1 \text{ end}\{i = 0\}$?

Les triplets de Hoare suivants sont-ils valides ?

- > $\{x = 10\}x := x + 1\{x = 11\}$? ✓
- > $\{c = 0\}\text{if } c \text{ then } x := c \text{ else } x = 1 \text{ fi}\{x = 0 \wedge c = 0\}$? ✗
- > $\{c = 2\}\text{if } c \text{ then } x := c \text{ else } x = 1 \text{ fi}\{x = 2 \wedge c = 2\}$? ✓
- > $\{\text{false}\}x := 0; y := 1\{x = 0 \wedge y = 2\}$? ✓
- > $\{i = 1\}\text{while } i \leq 10 \text{ do } i := i + 1 \text{ end}\{i = 10\}$? ✗
- > $\{i = 1\}\text{while } !(i \leq 0) \text{ do } i := i + 1 \text{ end}\{i = 0\}$? ✓

- > La **sémantique axiomatique** d'une instruction i peut être définie par l'**ensemble des triplets de Hoare valides** portant sur i .
- > Avec une définition informelle, il peut être néanmoins difficile de savoir si un triplet de Hoare est valide ou non, même sur un langage simple
- > **la sémantique axiomatique doit être définie formellement** pour éviter toute ambiguïté.
- > On utilise un **système de règles de déduction** dans ce but
- > Ce système est la **logique de (Floyd-)Hoare**

$$\begin{array}{c}
 \frac{}{\{P\} \text{ skip } \{P\}} \qquad \frac{\{P\} \ i_1 \ \{R\} \qquad \{R\} \ i_2 \ \{Q\}}{\{P\} \ i_1; i_2 \ \{Q\}} \qquad \frac{}{\{P[x \leftarrow e]\} \ x := e \ \{P\}} \\
 \\
 \frac{\{P \wedge e \neq 0\} \ i_1 \ \{Q\} \qquad \{P \wedge e = 0\} \ i_2 \ \{Q\}}{\{P\} \ \text{if } e \ \text{then } i_1 \ \text{else } i_2 \ \text{fi} \ \{Q\}} \\
 \\
 \frac{\{I \wedge e \neq 0\} \ i \ \{I\}}{\{I\} \ \text{while } e \ \text{do } i \ \text{end} \ \{I \wedge e = 0\}} \qquad \frac{P \implies P' \qquad \{P'\} \ i \ \{Q'\} \qquad Q' \implies Q}{\{P\} \ i \ \{Q\}}
 \end{array}$$

- > Ces règles d'inférence définissent une sémantique, **pas un algorithme**
- > **Séquence** : comment établir le prédicat R intermédiaire ?
- > **Boucle** : comment établir l'**invariant de boucle** I ?
- > **Règle d'affaiblissement (la dernière)** : quand l'appliquer ?
- > **Erreurs à l'exécution** : comment sont-elles prises en compte ?
- > **Affectation** : nature "arrière" de cet axiome, le secret de la réussite du calcul de plus-faible précondition qui va suivre

Mais au tableau parce que c'est un peu pénible dans une slide

But

Étant donné, une précondition Pre , une postcondition $Post$ et un programme p , $\{Pre\}p\{Post\}$ est-il un **triplet de Hoare valide** ?

Comment ?

- Le programme est vu comme un **transformeur de prédicat**.
- **On part de la postcondition $Post$** à la fin du programme et on raisonne en effectuant les étapes de calcul “à l’envers”.
- À chaque étape de calcul, on connaît la postcondition Q à vérifier et l’instruction i concernée et **on déduit la propriété P minimale** (la plus faible) telle que $\{P\}i\{Q\}$ soit un triplet de Hoare valide
- Lorsque le début du programme est atteint, **on doit prouver $Pre \implies P$**

Définition par induction sur la structure du programme

$$WP(\text{Skip}, Q) = Q$$

$$WP(i_1; i_2, Q) = WP(i_1, WP(i_2, Q))$$

$$WP(x := e, Q) = Q[x \leftarrow e] \quad \text{si } e \text{ est évaluée sans erreur}$$

$$WP(\text{if } e \text{ then } i_1 \text{ else } i_2 \text{ fi}, Q) = e \neq 0 \implies WP(i_1, Q) \\ \wedge e = 0 \implies WP(i_2, Q)$$

$$WP(\text{while } e \text{ do } i \text{ end}, Q) = I \\ \wedge (e \neq 0 \wedge I \implies WP(i, I)) \\ \wedge (e = 0 \wedge I \implies Q)$$

- > **Obligation de preuves** pour la gestion des erreurs
- > **Boucle** : comment établir l'**invariant** / ?
 - > Indiqué **manuellement** par l'utilisateur
 - > Idéalement, déduit automatiquement (par exemple, par interprétation abstraite)
- > **WP définit un algorithme**, modulo le problème des invariants de boucle

Correction partielle

Garantie que **si** le programme p termine, **alors** p satisfait la spécification S

Correction totale

Garantie que le programme p termine **et** p satisfait la spécification S

Terminaison

Prouver la terminaison nécessite la donnée d'une mesure décroissante pour chaque boucle (un **variant**), le plus souvent indiquée **manuellement par l'utilisateur**

Choix

Dépend du contexte : la correction partielle peut être suffisante.

Quand et pourquoi sont-elles générées ?

- > **pour garantir la sûreté du programme**, dès qu'une opération peut entraîner une erreur à l'exécution
 - > divisions par zéro
 - > déréférencement des pointeurs
 - > *overflow* des opérations arithmétiques
- > **pour vérifier les spécifications utilisateurs**
 - > postconditions
 - > assertions
 - > invariants et variants de boucles
 - > ...

Prouver ? Oui, mais comment ?

- > **Prouver à la main**
 - > impossible en pratique
- > **Prouver à l'aide d'un assistant de preuves** : outil pour assister un humain dans l'activité de preuves en le guidant, en résolvant les parties triviales ou calculatoires, et en vérifiant leur validité
 - > Coq, Isabelle, PVS, ...
- > **Prouver avec des prouveurs automatiques** : outil "presse-bouton" effectuant les preuves de manière automatique
 - > Alt-Ergo, CVC4, Z3, ...

Contexte de la preuve de programmes (de taille réelle)

- > les obligations de preuve sont **générées automatiquement** par des outils
- > plusieurs (dizaine, centaines, ... de) milliers d'obligations de preuve
- > plusieurs dizaines d'hypothèses par obligation de preuve (voire plus)

Faisabilité des preuves

- > infaisable par un humain
- > reste très lourd avec un assistant de preuves
- > emploi de **prouveurs automatiques indispensables**
- > **les prouveurs automatiques ne savent pas tout faire**

- > outil de vérification déductive de Frama-C
- > basé sur (une variante de) le calcul de plus faible précondition
- > vérification modulaire
- > génère des POs ensuite envoyées à des prouveurs automatiques
- > si toutes les POs sont prouvées **le programme respecte sa spécification**

- > **But** : spécification de fonctions
- > **Approche** : indiquer des assertions à propos des fonctions
 - > **Précondition** supposée vraie à l'entrée (garanti par l'appelant)
 - > **Postcondition** doit être vraie en sortie (garanti par la fonction)
- > quand la précondition est fausse, rien n'est garanti

But d'un contrat :

- > refléter la spécification informelle
- > ne devrait pas être modifié juste pour permettre à la preuve de passer

Spécifions la fonction valeur absolue (Démonstration)

- > WP peut vérifier (comme Eva) l'absence d'erreurs à l'exécution
- > si cela n'est pas fait, rien ne garanti que le programme fera son travail
- > aucune obligation d'utiliser le même outil, on pourrait
 - > utiliser Eva pour garantir cela
 - > utiliser WP uniquement pour les propriétés fonctionnelles.

Autre source d'imprécision (Démon)

assigns `v1, v2, ... vN;`

- > fait partie de la postcondition
- > spécifie les positions mémoire modifiées par la fonction
- > rien à dire sur les variables locales
 - > une fonction est autorisée à les modifier
 - > une postcondition ne peut de toute façon pas en parler
- > le mot clé `\nothing` indique un ensemble vide

Et si on a des pointeurs ? (Démonstration)

Spécification par cas

- > la précondition (**requires**) globales s'applique à tous les cas
- > la postcondition (**assigns**, **ensures**) globales s'applique à tous les cas
- > les comportements définissent des contrats pour des cas particuliers
- > pour chaque cas (**behavior**)
 - > le sous domaine est défini par la clause **assumes**
 - > la précondition du comportement est définie par **requires**
 - > la postcondition du comportement est définie par **assigns**, **ensures**
- > **complete behaviors** indique que les comportements couvrent les entrées
- > **disjoint behaviors** indique que les comportements ne se recouvrent pas

Spécifions la fonction valeur absolue avec des comportements (Démonstration)

Le calcul de plus faible précondition est une technique modulaire

En particulier ...

- > Chaque fonction est prouvée est isolation des autres
- > Lors d'un appel de fonction, l'analyseur ne va pas voir son code, on :
 - > prouve que les préconditions sont validées,
 - > admet que les postconditions sont validées.

Prouvons la correction de la fonction “maximum des valeurs absolues” (Démonstration)

Les boucles ont un nombre d'itérations variable

- > seule manière de traiter les boucles : la preuve par induction
- > nous avons donc besoin d'une hypothèse d'induction
 - > vraie juste avant la boucle
 - > vraie après $k + 1$ itérations quand elle est vraie après $k \geq 0$ itérations

Ce type de propriété est appelée un invariant de boucle

- > on génère pour cela deux obligations de preuves :
 - > l'invariant de boucle est vrai initialement
 - > l'invariant de boucle est préservé par toute itération

Comment trouver un bon invariant ?

- > Identifier **les variables modifiées par la boucle**
 - > Ne pas connaître le nombre d'itérations empêche la déduction de leur valeur
 - > Définir leur intervalle de valeur après k itérations
 - > Lister les variables modifiées avec **loop assigns**
- > Identifier les actions réalisées ou les **propriétés garanties par la boucle**
 - > Quelle **partie du travail** est déjà réalisée après k itérations ?
 - > Pourquoi la prochaine itération peut procéder comme elle le fait ?

Nécessite de l'expérience en preuve lorsque les programmes sont complexes

Remise à zéro d'un tableau (Démo)

La terminaison de programme est indécidable

- > un outil ne peut trouver une borne supérieure à un nombre d'itérations
- > en revanche, **si on fournit une borne supérieure**, on peut la vérifier
- > On appelle cette borne un **variant de boucle**

Variant de boucle

- > **expression** (pas un prédicat)
- > un variant n'est pas unique, si V est un variant, $V + 1$ aussi
- > la borne n'a pas besoin d'être précise
- > astuce simple, regarder la condition de boucle
 - > si l'expression est de la forme $e_1 < e_2$, $e_2 - e_1$ est un bon candidat

Remise à zéro d'un tableau (Démo)

Les boucles sont vues comme :

```
/*@ loop invariant I ;  
   loop variant   V ; */  
while (1) {  
    if (!cond) break ;  
}
```

On vérifie l'(in)variant quand on retombe sur l'entrée du **while**.
Donc **pas** lors d'un **break**.

Recherche dans un tableau (Démo)

Prédicats et fonctions utilisateur

```
/*@ logic integer diff{L1, L2}(int* a integer i) =  
    \at(a[i], L1) - \at(a[i], L2) ;  
  
predicate eq{L1, L2}(int* a, integer i) =  
    diff{L1, L2}(a, i) == 0;  
*/
```

Lemmes et axiomes

```
/*@ lemma transitive_eq{L1, L2, L3}:  
    \forall int* a, integer x ;  
        eq{L1, L2}(a, i) ==> eq{L2, L3}(a, i) ==>  
            eq{L1, L3}(a, i);  
  
    axiomatic Answer {  
        logic integer attltuae reads \nothing ;  
        axiom the_answer: attltuae == 42 ;  
    }  
*/
```

Représentation de la mémoire

- > En pratique, un programme (en particulier un programme C) contient des **opérations sur la mémoire** de nature complexe (manipulations de pointeurs, de tableaux, structures, ...)
- > Le calcul de WP doit en tenir compte

Différents modèles mémoires

- > **Modèle mémoire bas-niveau** : mémoire vue comme un tableau de bits ou d'octets
 - > simple, modélise bien la réalité
 - > possibilité de vérifier plein de propriétés (y compris très bas niveaux)
 - > obligations de preuve très compliquées à vérifier
- > **Modèle mémoire haut-niveau** : vue plus abstraite de la mémoire
 - > obligations de preuve plus faciles à prouver
 - > plus on est haut-niveau, moins les propriétés de bas-niveaux peuvent être prouvées
- > Exemple : **modèle de Burstall-Bornat** pour gérer les structures. Chaque champ est considéré comme séparé des autres.
- > Plus généralement, on peut **séparer la mémoire en différentes régions** pour faciliter la gestion des pointeurs et des tableaux.
- > **Limitations** : par exemple non-gestion des casts de pointeurs

- > Robert W. Floyd. *Assigning meanings to program.*
Mathematical Aspects of Computer Science 1967.
- > C. A. R. (Tony) Hoare. *An axiomatic basis for computer programming.*
Communication of the ACM 1969.
- > Edsger W. Dijkstra. *A Disciple of Programming.*
Prentice-Hall 1976.
- > Richard Bornat. *Proving pointer programs in Hoare logic.*
Mathematics of Program Construction 2000.

- > **Boogie**. Programmes C#. Langage de spécification : Spec#. <http://research.microsoft.com/en-us/projects/specsharp>
- > **Spark-2014**. Programmes Ada. <https://www.adacore.com/about-spark>
- > **ESC/Java2**. Programmes Java. Langage de spécification : JML. <https://www.cs.ucf.edu/~leavens/JML/index.shtml>
- > **Frama-C (WP)**. Programmes C. Langage de spécification : ACSL. <http://frama-c.com>

Partie VII

Combinaison d'outils

Intégration d'une méthodologie plus large

Tirer le meilleur de tous les mondes ... si on peut

Méthodes de spécification

- > Spécification informelle
- > Spécification formelle (ACSL ici)

Méthodes de vérification

- > Test (**précis**, mais **incomplet**)
- > Runtime monitoring (comme le test, plus **robuste**, plus **coûteux**)
- > Interprétation abstraite (**complet**, mais **imprécis**)
- > Vérification déductive (**complet**, mais **coûteux**)

Loin d'être en concurrence, ces techniques se complètent.

On n'a pas un budget infini pour le projet ...

Démarche

- > Montez progressivement en confiance
- > Trouvez le juste milieu entre coût de vérification et confiance
- > Trouvez le bon moment pour introduire un outillage

A garder en tête

- > Un patron va souvent vouloir tirer le coût vers le bas
- > L'industrie n'est pas toujours au courant de l'outillage existant en R&D
- > La R&D n'est pas toujours prête pour l'industrie

Aide au debugging de spécification

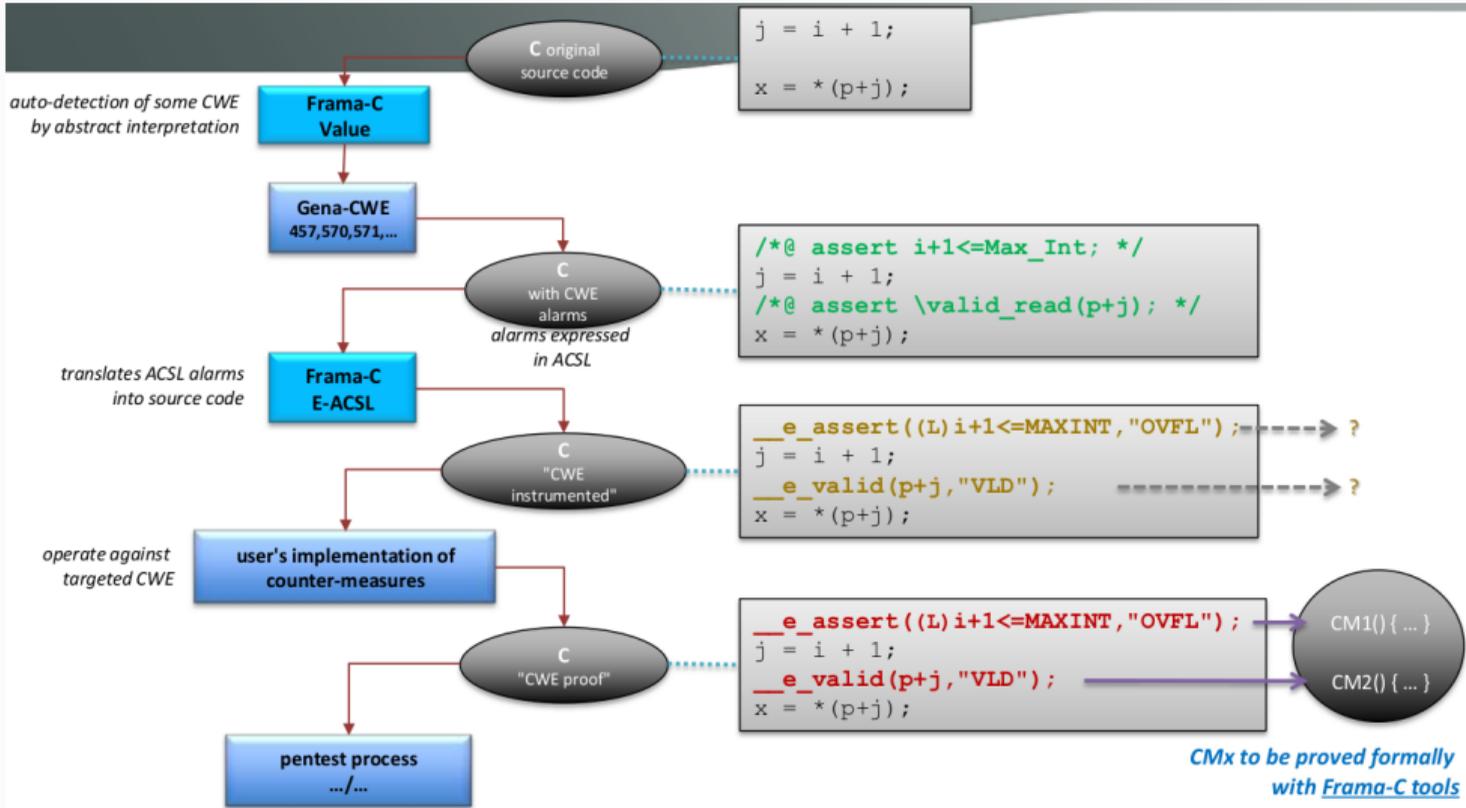
- Une spécification formelle peut être difficile à bien écrire
- On peut utiliser du runtime monitoring pour explorer la spécification

Vérifier ce qui reste à vérifier

- Des fois certaines propriétés sont trop dures à vérifier
- On peut utiliser du runtime monitoring pour construire un argumentaire

Mais on ne veut pas explorer ce qui est déjà vérifié

Voyons ça ...



Pour aider la méthode déductive ...

- > Vérifier l'absence d'erreur à l'exécution peut être long avec WP
 - > Mais Eva est très efficace pour ce travail
- ⇒ Eva pour vérifier l'absence d'erreur à l'exécution puis WP pour le reste

Pour aider l'interprétation abstraite

- > Parfois Eva n'arrive pas à tout vérifier
 - > Mais WP peut souvent avec un peu plus d'informations
- ⇒ Quand Eva n'arrive pas à prouver, utiliser WP pour aider

Voyons ça ...

Journey to a RTE-free X.509 parser

Arnaud Ebalard, Patricia Mouy, and Ryad Benadjila
`prenom.nom@ssi.gouv.fr`

ANSSI

Abstract. C programming language is a security nightmare. It is error-prone and unsafe, but, year after year, the conclusion remains the same: no credible alternative will replace C in a foreseeable future; all the more in low-level developments or for constrained environments

<https://github.com/ANSSI-FR/x509-parser>

Partie VIII

Conclusion

A-t-on nos réponses ?

Aller plus loin ?

But initial

avoir confiance dans les logiciels embarqués critiques

Qu'avons nous mis en place ?

- > du test systématique
- > de l'analyse statique correcte
- > de la vérification déductive
- > de la vérification à l'exécution

Loin d'être en concurrence, ces techniques se complètent.

Peut on avoir confiance dans un programme vérifié ?

- > à 100%, non !
- > les méthodes correctes restent un niveau d'assurance très élevé

Nombreuses possibilités de bugs

- > Implémentation des analyseurs
- > Prouveurs et assistants de preuve
- > Compilateur générant du code erroné
- > Matériel

Solutions possibles

- > noyau des assistants de preuve vérifié (Coq)
- > noyau des prouveurs automatiques vérifié (Alt-Ergo)
- > possibilité de vérifier les preuves des prouveurs (Zénon)
- > compilateur certifié (CompCert)
- > analyseur certifié (Verasco)

Autres risques

- > Non-vérification des **hypothèses implicites** des analyseurs
- > **Erreur dans les spécifications**
- > **Erreur dans une axiomatique** utilisée (théorie incohérente)

Solutions possibles : réduire les risques

- > vérifier toutes les hypothèses implicites
- > être prudent, mettre ceintures et bretelles

Propriétés du code

- > Propriétés fonctionnelles
- > Absence d'erreur d'exécution
- > Terminaison
- > Dépendances
- > Non-interférence
- > Propriétés temporelles
- > ...

Périmètre de la vérification

- > Quelle partie du code sous analyse ?
- > Que contexte initial ?

Base de confiance

- > Axiomes ACSL
- > Hypothèses des analyseurs
- > Doublures de fonctions
- > Frama-C et prouveurs externes
- > Et aussi : le compilateur, le runtime

Informatique et mathématiques

- > Les outils que nous avons vus sont basés sur des fondements mathématiques forts !
- > Ils font le lien entre informatique et mathématiques
- > Conseil de cours vidéos :

`https://www.college-de-france.fr/chaire/xavier-leroy-sciences-du-logiciel-chaire-statutaire/events`

Sur la preuve de programmes avec Frama-C

- > ACSL by Example
`https://github.com/fraunhoferfokus/acsl-by-example`
- > Introduction à la preuve de programmes C
`https://allan-blanchard.fr/frama-c-wp-tutorial.html`

Adoption of formal methods in various areas has dramatically improved the quality of computer systems. We anticipate that formal methods can provide similar improvement in the security of computer systems. ... Without broad use of formal methods, security will always remain fragile.

(Formal Methods for Security, 2016)

Mastering the complexity of software systems is a formidable intellectual challenge and Computer Science graduates need to understand the powerful formal methods and tools that are available. Then they can choose the right technique and tool for each task and with suitable humility deserve the title of Software Engineer.

(Formal Methods in Industry, 2024)

Contribuer à Framac

- > Open-Source classique : remonter des bugs, aider à la doc, etc.
- > Mais pas que !
 - > Opportunités de stages, de thèses ...

Contribuer à Framac

- > Open-Source classique : remonter des bugs, aider à la doc, etc.
- > Mais pas que !
 - > Opportunités de stages, de thèses ...

Merci de m'avoir écouté et bonne continuation !