

FlashQuation: Vergelijkingen vinden aan de hand van voorbeelden

Jeroen Craps & Tom De Groote

Introductie

Probleemgebied:

Gegeven: Set van getallen

X	Y	G
2	2	4
3	4	12



Gezocht: Passende vergelijking
 $\rightarrow X * Y = G$

Hypothese:

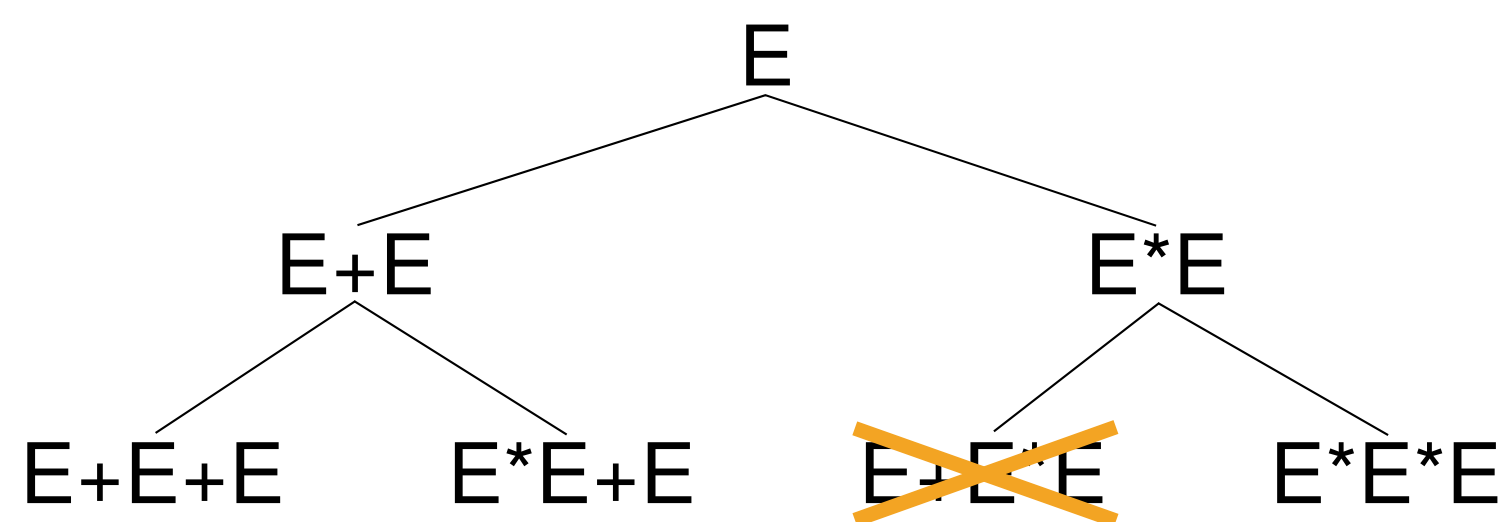
Er kan een passende vergelijking gevonden worden binnen beperkte tijd.

Methoden

STAPPEN:

- Opstellen van de bewerkingsboom.
- Verwijderen van redundante knopen.
 \rightarrow Optimale bewerkingsboom.

vb.



- Efficiënt zoeken naar een vergelijking.

Vraag 1: Hoeveel kleiner is het aantal knooppunten van de optimale boom ten opzichte van de originele boom?

Vraag 2: Worden er meer unieke vergelijkingen gevonden, in beperkte tijd, door gebruik van een optimale boom?

Vraag 3: Wordt er door toevoeging van constanten vaker een passende vergelijking gevonden?

Experimenten:

- 4 variabelen
- 500 uitvoeringen
- Bewerkingsboom met diepte 5
- Geen tijdslimiet

Contextvrije grammatica (CFG)

Productieregels van de vorm:

$E \rightarrow T$
 $E \rightarrow E O T$
 $T \rightarrow 1..9 | X | Y | \dots$

Voorbeeld:

E
 $\rightarrow E O T$
 $\rightarrow T O T$
 $\rightarrow X * Y$

met E een niet-terminaal symbool, T een terminaal symbool en O een operand

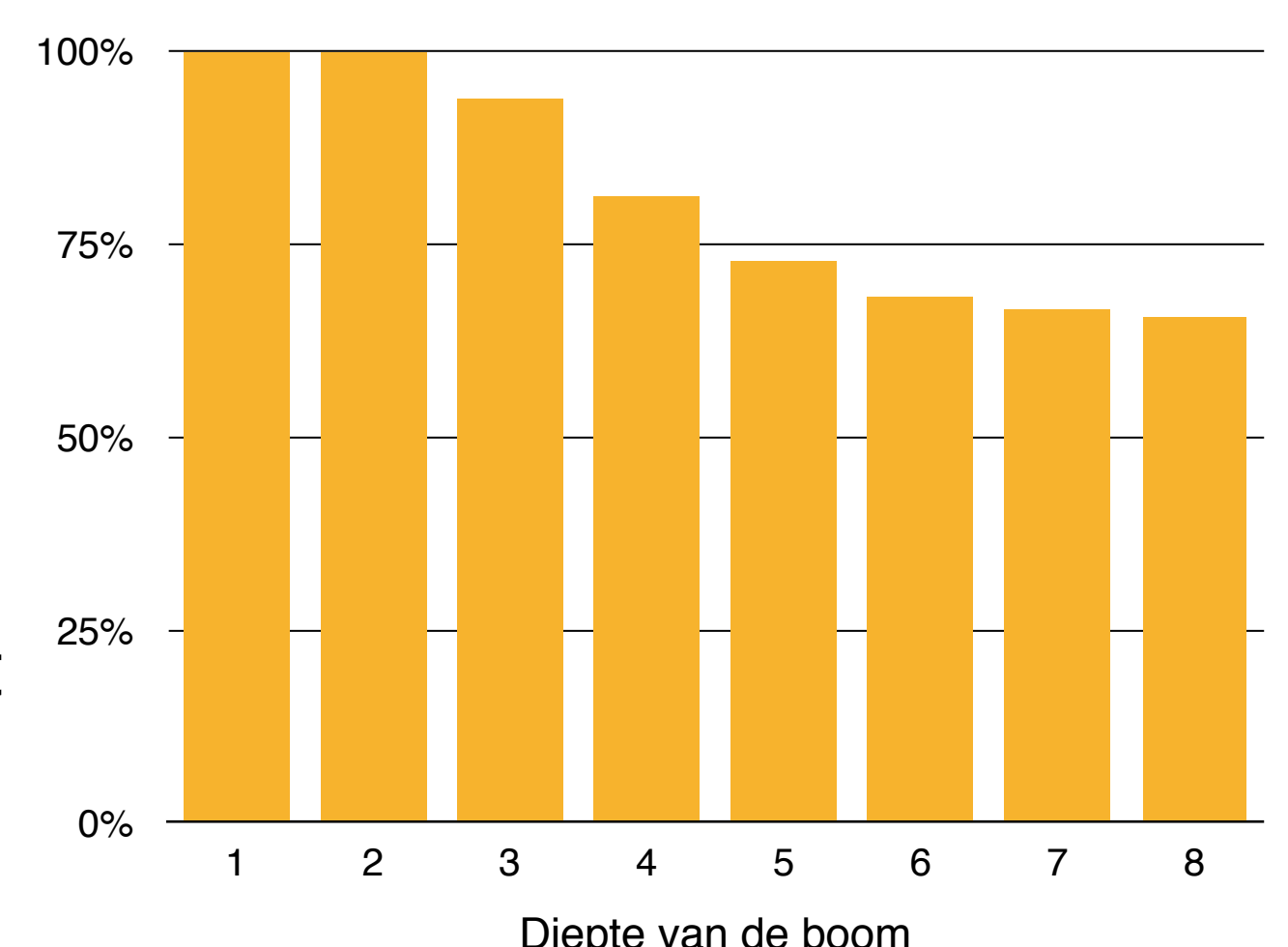
Resultaten

Antwoord 1:

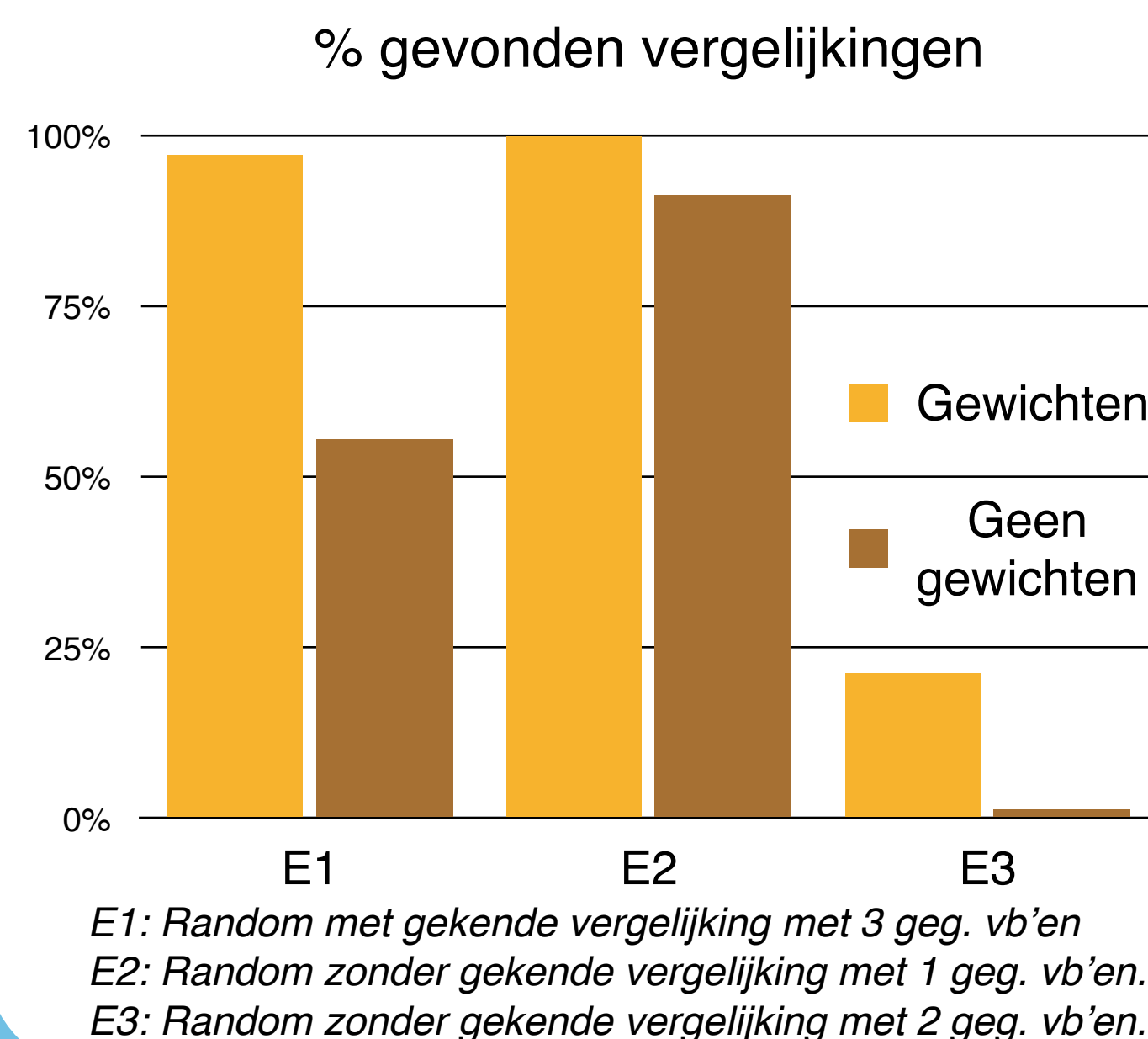
Diepte van de boom ↗
↓
verwijderde knopen ↗

Antwoord 2: (Toekomst)

Verwacht wordt dat de tijdswinst dezelfde curve zal volgen.



Procentuele weergave van het aantal knooppunten van de optimale boom in vergelijking met de volledige boom.



E1: Random met gekende vergelijking met 3 geg. vb'en
E2: Random zonder gekende vergelijking met 1 geg. vb'en.
E3: Random zonder gekende vergelijking met 2 geg. vb'en.

Antwoord 3:

Contextvrije grammatica
+
Gebruik van gewichten
||
Meer vergelijkingen

Conclusie

In ongeveer **70%** van de gevallen wordt er een passende vergelijking gevonden. (Indien gebruikersvergelijkingen gelijkaardig zijn aan deze uit E1 en E3)

Toekomstig onderzoek:

- De correlatie tussen de oplossingsgraad en welke constanten (of operanden) gebruikt worden.
- Implementeren van een efficiënt evaluatiealgoritme.

Referenties:

- Spreadsheet Table Transformations from Examples, CACM, W.R. Harris, S. Gulwani
- Example-Based Learning in Computer-Aided STEM Education, CACM 2014, Sumit Gulwani
- Equation Discovery, Encyclopedia of Machine Learning



Begeleiders: Prof. L. De Raedt
Dr. A. Kimmig

Contact: jeroen.craps@student.kuleuven.be
tom.degroote@student.kuleuven.be

KU LEUVEN